ASTES

Advances in Science, Technology & Engineering Systems Journal

VOLUME 2-ISSUE 6|NOV-DEC 2017

www.astesj.com
ISSN: 2415-6698

Editorial

Advances in Science, Technology and Engineering Systems Journal (ASTESJ) is an online-only journal dedicated to publishing significant advances covering all aspects of technology relevant to the physical science and engineering communities. The journal regularly publishes articles covering specific topics of interest.

Current Issue features key papers related to multidisciplinary domains involving complex system stemming from numerous disciplines; this is exactly how this journal differs from other interdisciplinary and multidisciplinary engineering journals. This issue contains 36 accepted papers in Computer and Telecommunication domain.

Editor-in-chief
Prof. Passerini Kazmersk

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 2 Issue 6 **November-December 2017**

CONTENTS

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 1-12 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

Clean Energy Use for Cloud Computing Federation Workloads

Yahav Biran[*,1], Sudeep Pasricha[2], George Collins[1], Joel Dubow[3]

[1]*Colorado State University, Department of Systems Engineering, Fort Collins, 80523, CO, USA*

[2]*Colorado State University, Department of Electrical and Computer Engineering, Fort Collins, 80523, CO, USA*

[3]*Fulcrum Co., Cyber Security, Centreville, 20120, VA, USA*

ARTICLEINFO

Article history:
Received: 01 July, 2017
Accepted: 10 August, 2017
Online: 19, August, 2017

Keywords:
power consumption
resource utilization
cloud computing
cloud federation

ABSTRACT

Cloud providers seek to maximize their market share. Traditionally, they deploy datacenters with sufficient capacity to accommodate their entire computing demand while maintaining geographical affinity to its customers. Achieving these goals by a single cloud provider is increasingly unrealistic from a cost of ownership perspective. Moreover, the carbon emissions from underutilized datacenters place an increasing demand on electricity and is a growing factor in the cost of cloud provider datacenters. Cloud-based systems may be classified into two categories: serving systems and analytical systems. We studied two primary workload types, on-demand video streaming as a serving system and MapReduce jobs as an analytical systems and suggested two unique energy mix usage for processing that workloads. The recognition that on-demand video streaming now constitutes the bulk portion of traffic to Internet consumers provides a path to mitigate rising energy demand. On-demand video is usually served through Content Delivery Networks (CDN), often scheduled in backend and edge datacenters. This publication describes a CDN deployment solution that utilizes green energy to supply on-demand streaming workload. A cross-cloud provider collaboration will allow cloud providers to both operate near their customers and reduce operational costs, primarily by lowering the datacenter deployments per provider ratio. Our approach optimizes cross-datacenters deployment. Specifically, we model an optimized CDN-edge instance allocation system that maximizes, under a set of realistic constraints, green energy utilization. The architecture of this cross-cloud coordinator service is based on Ubernetes, an open source container cluster manager that is a federation of Kubernetes clusters. It is shown how, under reasonable constraints, it can reduce the projected datacenter's carbon emissions growth by 22% from the currently reported consumption. We also suggest operating datacenters using energy mix sources as a VoltDB-based fast data system to process offline workloads such as MapReduce jobs. We show how cross-cloud coordinator service can reduce the projected datacenters carbon emissions growth by 21% from the currently expected trajectory when processing offline MapReduce jobs.

1 Introduction

Over the past decade, cloud-based systems have been required to serve an increasing demand from users work flows and data. Cloud-based systems may be classified into two categories: *serving* systems and *analytical* systems. The former provides low-latency read or write access to data. For example, a user requests a web page to load online or requests video or audio streaming. The latter provides batch-like compute tasks that process the data offline that are later sourced to the serving systems. The service level objectives (SLO) for serving jobs are on the order of fractions of a second, while the SLO for analytical jobs are on the order of hours, sometimes days.

[*]Corresponding Author, ybiran@colostate.edu

https://dx.doi.org/10.25046/aj020601

Today, public cloud service providers (CSP) attempt to process both of these workloads with a rich platform that guarantees cost and SLO to their clients. Cloud computing is an emerging infrastructure with limited regulation and compliance requirements [1]. Recently the Office of Management and Budget issued a Federal Data Center Optimization Initiative that promotes increasing use of Green Energy and increased utilization efficiency for all US Federal datacenters [2]. Specific target numbers are set for the end of fiscal year 2018. This publication addresses how those federal requirements may be attained and how federated cloud computing is a key enabler for attaining those performance targets.

Beginning in 2013, the US government initiated a carbon-tax on IT organizations to encourage major CSPs to pursue green energy opportunities for their datacenters operations [3]. US datacenters are projected to consume approximately 73 billion kWh by 2020 [3] with a corresponding increase greenhouse gases. Green energy generation growth is expected to triple by 2040 [4]. However, there is no cohesive system existing to coordinate the rising datacenter energy demand with rising green energy supply. This chapter utilizes Cloud-Federation as a multi-cloud resource coordination system that matches computational resource demands with available energy supply to maximize the utilization of green energy for processing cloud-workloads.

Most CSPs seek more market share in competition with other CSPs. One outcome of such competition is an ever-growing infrastructure in the form of new datacenters across the globe with no countervailing forces to meet user demand more efficiently and satisfy societal environmental and energy requirements. This sub optimum use of infrastructure increases the carbon footprint attributable to cloud computing services and also drives up costs to CSP's. The following sections investigate two types of workloads, serving and analytical systems. This chapter will focus on two workload types, On-demand streaming as a representation of *serving* systems. It will also investigate *analytical* systems and present a unique model that allows optimal clean energy usage.

1.1 Enabling Green Content Distribution Network by Cloud Orchestration

On-demand streaming constitutes up to 85% of Internet traffic consumption [5]. On-demand streaming content is managed and distributed by content service providers. It then cached and distributed by Content Delivery Networks(CDN) located at the edges of the Internet network close to the consumers. Because streaming constitutes such a large fraction of Internet resource consumption, this paper will, of necessity, focus on methods to employ green energy to better operate CDN instances of on-demand streaming jobs, which include both video and audio content.

Meeting the Federally mandated approach of maximizing the utilization of green energy to operate CDN instances (for government with recommendations for private sector use as well) requires an energy source-demand coupling scheme that insures SLO levels of power availability but is structurally biased towards green energy sources over hydrocarbon fueled energy sources. A system to accomplish this will have to provide seamless failover in the case of sudden interruption of green energy to grid-energy sources or vice versa i.e., fallback from grid-energy to green energy when surplus green energy is available.

User expectations in on-demand streaming requires different service level requirements than other serving systems workloads. Serving systems workloads are comprised of interactive sessions that pivots on minimum latency. However, low latency is less critical in analytical on-demand streaming since application clients use buffering techniques to mitigate long latency effects. Therefore, on-demand streaming workloads fits, more closely than interactive workloads, with the observed intermittent and varying green energy availability characteristics.

Green energy supply is unpredictable and requires a complex, adaptable, resource allocation system to provide CDN services with steady energy supplies while concurrently seeking minimal carbon footprint. This dynamic availability of green energy resources in a smart grid requires real time communication of both short term and predictive energy needs from cloud service providers to green energy providers. The green energy providers need to disclose availability dynamically to CSPs, who, in turn, disclose their changing energy demands for near term computing. SPs can then better maximize the use of green energy for on-demand streaming processing.

This is a classical resource management and coordination problem [6, 7]. The following approach builds upon prior work that was done in this area [8, 9], specifically that done on alleviating the sudden lack of green energy to meet low-latency workloads. The approach herein employs an application-buffering scheme that better allows for opportunistic, green, on-demand streaming processing. It requires an extended, cohesive, federated system that aggregates supply and demand across multiple geographic locations employing the smart grid command and control infrastructure to achieve an optimal dynamic matching of green energy sources and computing loads.

This chapter proposes an implementation that utilizes a control component in a federated cloud that coordinates and optimizes the resource allocation among the participant CDN providers. It treats the volatile nature of green energy resources as a resource allocation problem, the solution of which is a resource orchestration system that is optimized with the goal to operate increasingly near to the limit of supply by green energy sources constrained by SLO reliability requirements. This system will be demonstrated by modeling a prototype that simulates resource allocation in a micro federated cloud eco-system to achieve an energy supply-computation demand match optimized within seconds.

Since the focus of this work is on green energy utilization in a federated cloud, the scheduling algorithms and resource management issues, while important, are discussed only to the extent necessary to help the reader understand the required architecture for heterogeneous energy compute clustering. This work is meant both as a case study in energy utilization, and a presentation of a novel method of coordinating high-velocity data streams, and extended to a unified orchestration system, to optimize the performance of federated cloud systems.

The chapter starts, with the on-demand streaming economics, increased green energy utilization and anticipated smart grid progressions as applied to on-demand streaming. Then it discusses the green energy utilization problem is analyzed. Finally, we present a cloud coordinator prototype that is built on Kubernetes [10, 11], an open source cluster manger, and extend that prototype to discusses the need for and requirements of a unified system that orchestrates cluster compute resources in a federated cloud.

1.1.1 On-Demand Streaming and CDN

Over the last decade, video and audio traffic became the dominant segment of consumer internet traffic. Cloud service-providers such as Netflix, Amazon Instant and YouTube disrupted the prior linear TV data distribution model. Also, video streams delivered by mobile terminals grew as mobile connectivity improved [12]. Video streaming is expected to constitute up to 85% of Internet consumers traffic [5] within a few years. The US portion of video streaming is 14%[1] and the number of US Unique IPv4 connected addresses is 17% [12]. The streaming workload is comprised of live streaming and on-demand streaming, with the relative fractions of 6% and 94% respectively [5]. Other predictions support similar ratios, 12% live-streaming and 88% for on-demand.

A key driver for the rapid expansion of streaming video was the shift from specialized streaming protocols and infrastructures such as RTSP, and RTMP [13] to a simple HTTP progressive download protocol. This led to a shift from proprietary streaming appliances to commodity servers. In turn, this change removed a barrier for CDN's to process on-demand workloads. Most present day, CDN service providers support a seamless integration with cloud-based object storage that pipelines the digital content from the organization site to the CDN instance that runs at the Internet provider edge[2]. Furthermore, the HTTP chunk-based streaming protocol support in a CDN allows the client application sufficient time to detect the optimal CDN instance to handle user workload. The optimal CDN instance assignment is done by the cloud control plane resource manager. The prototype described below will demonstrate such optimal resource allocation.

We used server utilization and power metrics from [3, 12] to design the prototype. Most of these sources we considered have limited utilization rates and server utilization distribution. Also, utilization and power consumption do not scale linearly [14]. However, for clarity, in the interest of maintaining focus on the larger goal of the paper, CDN resource management systems for green energy utilization, we assumed linear relationships and accepted the risk of loss of accuracy in our estimates. High fidelity simulation accuracy is not critical for the goal of this chapter.

1.1.2 Energy Saving Potential in Operating a Distributed CDN Resource Management System

The approach is to aggregate the required traffic for on-demand workload processing, and use standard compute device specification to assess the electrical energy and carbon footprint that will be required by that workload[3].

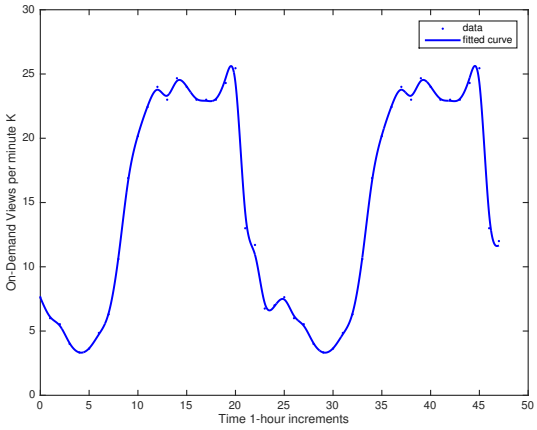


Figure 1: On-Demand Video views observed throughout 48 hours with 1-hour increments. Data was fitted with smoothingspline for curve and surface fitting [5]

The estimated data rate for streaming is given as $S_{total} = 63000PB/mo$ (where PB is Petabytes). Figure 1 shows users workload pattern of 9 busy hours in which the workload spans throughout 13 hours a day which yield $126PB/sec/mo/$ The on-demand streaming portion is estimated as 78% across four main US regions denoted by $k = 4$, the number of region used, for the purposes of this paper although k can be varied depending upon the degree of granularity desired in the simulations. $S_{on-demand}$ denote the on-demand portion.

$$S_{on-demand} = S_{total} \cdot 78\% = 49140PB/mo$$

$n_{hours/day}$ denotes the number of effective hours in a day for streaming.

$$n_{hours/day} = n_{busy} + k = 13$$

$$D_{rate} = \frac{S_{on-demand}}{n_{hours/day}} = 126PB/sec/mo$$

N_{max} denotes the estimated number of required servers in maximum CPU capacity. N_{max} is bounded by the maximum network throughput a single server can ingest. Standard commodity servers can handle up to 8.5Gbps i.e. 1.026GB/sec.

$$N_{max} = \frac{D_{rate}}{T_{max}} = \frac{126 \cdot 10^6 GB/sec}{1.026GB/sec} = 118588235.3$$

u_{opt} denotes the CPU utilization factor so servers has sufficient capacity to handle management tasks. We estimate 60% utilization factor $u_{opt} = 100/60$.

$$N_{opt} = N_{max} \cdot u_{opt} = 118588235.3 \cdot \frac{100}{60} = 197647058.8$$

E_s denotes the midrange server energy consumption for various server types s. We consider three types of servers:(1) compute server(5kWh/server), (2)

[1]Internet Statistics retrieved from https://www.statista.com/chart/2647/global-internet-usage-by-the-numbers/
[2]Cloud Front reference retrieved from https://aws.amazon.com/cloudfront/
[3]OpenCompute Project, Servers Specification guide retrieved from http://www.opencompute.org/wiki/Server

digital storage server(1.7kWh/server) and (3) network server(1kWh/server). Storage server acts the digital storage controller. The network server acts as the router and switch. The compute server is the server that processes the on-demand streaming.

$$E_y = N_{opt} \cdot \sum_{s \in S} E_s \quad (1)$$

$$= N_{opt} \cdot (5kWh + 1.7kWh + 1kWh)$$

$$= 1.521 \cdot 10^6 kWh/mo = 18.26GWh/y$$

The saving potential from running on-demand video streaming using green energy resources is 18.26GWh a year based on current on-demand consumption and expected to grow 89% by 2019 [4]. i.e. 34.5GWh per year for on-demand streaming. The next sections will explain the challenges in utilizing green energy followed by a method that addresses some of these challenges and thereby maximizes the utilization of green energy.

1.2 Coordinating Green Clouds as Data-Intensive Computing

The following paragraph focuses on *analytical* systems workloads that typically comprise 48% of the cloud workload [15]. Also, some of the workloads patterns can be predicted as recurring jobs [16] and the deadlines for *analytical* jobs are more liberal by an order of magnitude than the *serving* jobs. Finally, one of the ways to handle offline workloads is to process the data streams offline through a highly scalable data-parallel frameworks like MapReduce by submitting jobs to a control plane [17].

This chapter proposes a component in a federated cloud that will optimize the resource allocation and coordination among the participant CSPs. It will also focus on addressing the time and power volatile nature of renewable energy resources as a fast date problem. It will present a resource management system with a goal to increase the utilization of data-centers, and to operate increasingly to the limit of supply by renewable energy sources. Finally, a prototype is built to simulate resources allocation in a micro federated cloud eco-system to achieve a supply demand optimized match within seconds.

Clean energy sources are time and power volatile and require complex resource allocation and coordination systems to maintain highly available data-center service with steady energy at a minimal carbon footprint. Further, availability of clean energy resources in a smart grid can be pushed from a variety of energy sources deployed across the nation as can the pull needs of data-service centers. Thus, CSPs will publish their compute resources availability, and service providers (SP), will publish their changing energy demands for near term computing. These streams of data include both high-volume and high-velocity characteristics termed a fast-data problem [18]. Our proposal uses as a base previous work that was done in this area [8, 19, 20]. However, previous work treated clean energy within a single data-center operated by a CSP. The proposal below suggests an extended cohesive system that aggregates supply and demand across multiple geographic locations employing the smart grid sense command and control to achieve an optimal match.

1.2.1 Optimum Clean Energy Utilization is a Fast-Data Problem!

This study suggests that efficient, clean energy utilization requires three consecutive steps: (1) Clean energy resource availability signals. (2) Exploration and analysis of prior years seasonal solar data and current weather reports and evolving green energy capacity availability planning. (3) Acting fast enough based on the predictive analysis. The last step is the key component in solving the optimum allocation problem.

We judge that clean energy utilization is different from the classic big-data problem such as the Hadoop MapReduce Method. Big-data solutions solve the case where the data is at rest, not fully consistent (aka eventual consistency) and require liberal SLO that is later provided as computing trends, and other business intelligence applications. However, our case handles data in motion that requires consistent, real-time aggregation, and transaction processing. That is per-event decision making using both real-time contextual and historical data as dual guides for proper algorithms. Finally, we argue that based on the clean energies' volatile nature, processing streams of compute demand and clean energy supply requires that a need for a compute resources can be addressed with a currently available demand.

1.3 Why is Green Energy Utilization Hard?

The following section describes why utilizing green energy for compute purposes, while a justifiable goal, is limited by SLO reliability. It will present a scenario where balancing time-varying energy generation patterns with changing dynamic energy demands of cloud computing sometimes conflict. The green energy time varying generation patterns considered by us focuses on wind and solar generation. Figure 4 show historical data on dynamic nature of green energy sources and Figure 5 shows the dynamic cloud energy demand.

1.3.1 Frequency Stability in Wind-Power Generation

The daily wind power variation characteristics will be employed as a metric that illustrates the duration and level for a given amount if wind energy availability. The electricity generation process from wind is comprised of a wind turbine extracting a kinetic energy from the air flow. The wind is rarely steady; it is influenced by the weather system and the ground surface conditions, which are often turbulent [21]. Also, the generation process must happen at the same instant it is consumed [22] unless it is stored in grid level battery banks. Unfortunately, grid level energy storage technology is not keeping up with grid level energy generation technology.

Sample wind and power generation data were obtained from NREL [22]. We used datasets from 2006-2012 across different regions in the US and aggregated more than 600

[4] https://www2.deloitte.com/content/dam/Deloitte/in/Documents/technology-media-telecommunications/in-tmt-rise-of-on-demand-content.pdf

observations. Finally, the data were fit using smoothing splines[5]. The usable power generated from a wind turbine is generally described by a Rayleigh distribution [21]. It defines three main points in the wind power generation process: (1) the cut-in is the minimum viable wind speed for electricity generation from a wind turbine. (2) the rated level, describes the point where the power reached its local maximal capacity without adverse effects on the turbine life by too strong a wind. (3) Cut-Off, is the term for the local minimum for the generation cycle. Beneath that speed, there is not enough power for viable electricity generation. Thus, if the wind velocity is too low, the data-server gets no wind energy. Figure 2 shows wind generation variations that crudely fit a Rayleigh distribution with $b = 300$ assuming the form of the Rayleigh Probability Distribution Function is:

$$f(x|b) = \frac{x}{b^2} e^{\frac{-x^2}{2b^2}}$$

The measured generation cycles range between 140 and 180 minutes per cycle. Equation 2 expresses the generated power by a wind turbine, given a wind velocity. The function g describes a viable electricity generation given a wind power. The wind power availability indications will be generated by a wind turbine and fed into the coordinator database as a potential power source to datacenters in a region. Our prototype will assume wind power availability indications as the wind tuple {*region,cut-in,rated,cut-off*} indications.

$$P_{output}(wind_v) = \begin{cases} 0 & : \text{ if } wind_v \leq rated \\ g(wind_v) & : \text{ if } rated < wind_v \leq c_{out} \\ 0 & : \text{ if } wind_v \leq c_{out} \end{cases} \quad (2)$$

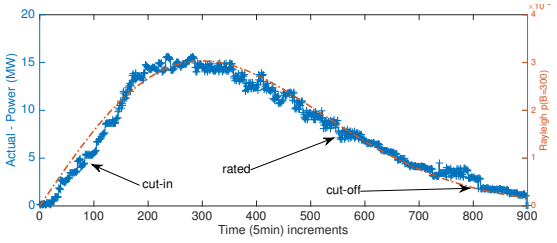


Figure 2: Aggregated wind power measurements between 2007-2012 that fits Rayleigh Distribution with b=300

1.3.2 Efficiency and Daily/Hourly Availability in Solar-Power Generation

Photovoltaic solar (PV) energy availability is defined by the solar power intensity denoted by s(*Watts/m2*), which varies with local daylight hours and the clear or cloudy sky conditions [22]. Moreover, the PV cells are most effective at lower temperatures [23]. The PV cells electrical power generation, defined by Equation 3, is a function of the solar intensity denoted by η_{solar}. Solar power generation also depends on the PV power efficiency denoted by s. It encapsulates both the predicated temperature,the sky conditions, the solar cell efficiency, and the DC to AC inverter efficiency. The solar cell area denoted by $a(m^2)$.

$$P_{output}(s) = \eta_{solar} \cdot s \cdot a \quad (3)$$

Our prototype will assume solar power availability indications as the solar-tuple {*region,power-efficiency*}. Based on the solar generation pattern presented in figure 3, the generation prediction utilizes the the local time in a region and the given power-efficiency

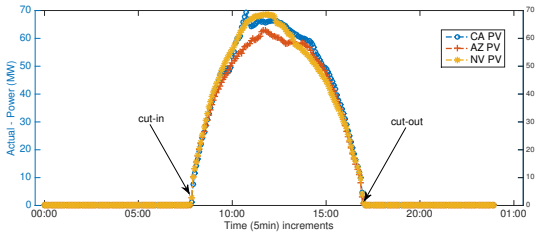


Figure 3: Aggregated solar power generation between 2008-2011 taken in Palm Springs CA, Prescott Airport CPV, AZ and Nevada Solar One, NV, indicating on a stable and fixed solar-based power [22]

1.3.3 Optimum Green Energy Utilization for On-Demand Streaming is a Resource Management Problem!

This study suggests that efficient; green energy utilization for on-demand video streaming workloads has three main requirements: (1) efficient compute resource discovery, (2) efficient load balancing among the provisioned compute resources and (3) smart failover mechanism that mask failover events from *green* CDN-edge instance to *grid* CDN-edge instance while end-users stream on-demand video [24]. These are discussed below.

Compute Resources Discovery. This assessment comprises both internal and external discovery. Internal discovery refers to CDN-edge instances that run in compute pods that must be able to be easily discovered and connected to control-plane endpoints consistently regardless of which cloud-service-provider is hosting the CDN-edge. External discovery refers to the ability of end-users discovering CDN-edge instances through DNS services for HTTP(S) on-demand video streaming.

Optimal Load Balancing is the seeking of the ”best” CDN-edge, based on optimization criteria, for any given the workload processing. After initial discover and connection, clients should be served by the optimal instances based on proximity from the end-users, current load factor, and the availability of green energy resources. e.g., session requests originated from New Jersey should be served by US East as oppose to US West to avoid latency and signal loss.

Efficient Failover is a main component for on-demand video streaming based on green energy. If the endpoint becomes unavailable, in this case due to a sudden lack of green energy, the system must failover the client to another available endpoint that manages the streamed content. Also, failover must be completely automatic i.e. the clients end of the connection remains intact, and the end-user oblivious to the failover event, which means that the end-user's client software requires no support handling failover events. Finally, multiple CDN-edge instances co-located in a region should be accessible by end-users through Domain Naming

[5]Matlab Smoothing Splines retrieved from https://www.mathworks.com/help/curvefit/smoothing-splines.html?requestedDomain=www.mathworks.com

Service (DNS), as most clients-streaming (browsers) software supports DNS resolutions for finding available CDN-edges.

The green energy utilization model for processing on-demand video streaming is different than the classic scheduling problem where classical optimal resource allocation techniques are applied [25, 26]. We argue that based on the green energies' volatile nature and the on-demand video streaming workload characteristics, the optimal resource allocation approach should be opportunistic. It requires an effective resource management system for processing on-demand video streaming workloads. Our prototype will employ a Kubernetes flavor "Ubernetes" that implements the three main requirements above.

2 Clean Energy Mix Evaluation for Online On-Demand System

In the following section we evaluate a compute load coordination system component that harmonizes *on-demand streaming* job demands with available compute resources, with priority given to those powered by green energy sources. Such resources will be published to the coordination system through a resource availability tuple {*region,cut-in,rated,cut-off,power-efficiency*}, where *region* indicates the geographic availability region. *power-efficiency* indicates solar or wind based energy power efficiency. *cut-in, rated* and *cut-off* the values appropriate to those energies.

On-demand streaming job demand includes the specific *region, total-job workload, load-factor*, as well as *contract deadline* SLA. The *load-factor* indicates the required number of CPU cores per the *total-job-workload*. The geographic *region* indication will be used to optimize the match between the supply and demand. Also, the *total-job-workload* and the *deadline* will be compared against the *cut-in,rated, cutoff* time for wind or *power-efficiency* for solar, based on the published *load-factor*.

We suggest a hybrid datacenter that does not deviate from the common datacenter architecture. The core difference lies on an automatic transfer switch (ATS) that switches between different available power sources: generator, grid or green energy when available. In both cases the datacenter design does not change and requires incremental changes only by adding green energy power sources to the datacenter's ATS's (Figure 6).

We suggest two types of compute clusters, green-clusters powered by green energy and grid-clusters powered by the electrical grid. Figure 6 shows a simplified datacenter power distribution that supports green energy sources. In such datacenter, both *serving* and *analytical* systems deployed in grid clusters. Further, for incoming *analytical* workloads, few clusters use green resources when there are a viable green energy and mostly standby. As a mitigation strategy, a compute live migration procedure will be available in case of unpredicted lack of renewable resources during a workload processing which presents a risk for SLO violation.

2.1 Experiment Planning

Below is simulation of a cross-regional platform that is comprised of control-plane, workload-plane and coordinating components. This will be embodied in a resource allocation system (Kubernetes). This system will: (1) provision resources to be neared users; (2) optimize utilization by prioritizing the use of underutilized resources; and (3) seamlessly remove malfunctioning hardware from the system. The control-plane will enable an effective compute resource provisioning system that spans across different public cloud providers and regions. The coordinating components will accept user-workload demands as well as green energy availability from various regions and opportunistically seek to process streaming workloads using compute resources provisioned by green energy resources. The workload-plane will be comprised of edge streaming servers that process the end-user on-demand video streaming. It will built of standard Apache HTTP[6] servers that runs on the edge location.

The control-plane software infrastructure is based on Kubernetes [10], it facilitates internal discovery between CDN instances so instances can connect across different cloud boundaries and regions. Further, end-users can discover the optimal CDN-edges that are (1) nearby, (2) less loaded and (3) healthy. Finally, the Kubernetes automation framework allows the failover mechanism with no dependency upon the end-user client. In particular, we will exploit the *livenessProbe* option that automatically removes green-compute pods, a set of CDN-edge instances, in case of a sudden lack of green energy.

The coordinator component accepts incoming supply and demand traffic, calculates a potential match, within minutes, and notifies back the CSP and the SP for transaction completion. We use Redis[7] as the in-memory data store as the database that stores the system supply and demand calls originated by the end-user workload. The workload is generated by Jmeter instances[8]. The workload generated based on the on-demand video views observed by [12] depicted in Figure 1. Green energy availability simulated based on the known regional patterns depicted by Figure 2 and Figure 3.

We count the number of matches i.e. on-demand video streaming processed by CDN-edge instances operating on green energy. Also, we measure the false-positive cases where a match was suggested but did not met the SLO's deadline due to a violation that caused by a sudden lack of green energy resources. We use the data to extrapolate the possible energy (kWh) that could be generated by the using green CDN-edge instances depicts in Equation 1.

2.2 Execution - The Preparation

The prototype experiment included the setup of three virtual datacenters deployed in different regions: (1) Central US, (2) West US and (3) East US. The clusters were sized

[6] Apache Web Server reference retrieved from https://httpd.apache.org
[7] http://redis.io
[8] http://jmeter.apache.org

based on US population distribution[9] by regions i.e. 20% for West US, 40% for East US and 40% Central US. The cluster sizes for West US, Central US, and East US are 3, 7 and 7 machines respectively. Each machine is standard 2-CPU cores with 7.5GB of memory. Also, the user demand simulation will rely on the US population distribution. Finally, the green energy supply simulation will be based on wind or solar availability observed in the various regions.

The control-plane is comprised of docker API server and controller-manager. The controller coordinator component will need to allocate resources across several geographic regions different cloud providers. The API server will run a new federation namespace dedicated for the experiment in a manner such that resources are provisioned under a single system. Since the single system may expose external IPs it needs to be protected by an appropriate level of asynchronous encryption[10]. For simplicity, we use a single cloud provider, Google Container Engine, as it provides a multi-zone production-grade compute orchestration system. The compute instances that process the user workloads are deployed as Docker containers that run Ubuntu 15 loaded with Apache HTTP server. For simplicity, we avoid content distribution by embedding the video content to be streamed in the Docker image. We run 52 Docker containers that span across the three regions and act as CDN-edges. Green CDN-edge instances differ from grid CDN-edge instances by Kubernetes labeling. The simulation of the hybrid datacenter is depicted in Figure 6.

A coordination database system that aggregates green energy, solar or wind, availability, was built in software. When energy sources manifest the cut-in patterns depicted by Equation 2 and Equation 3, the coordination system starts green CDN edges in the availability regions. Also, when green energy availability reaches cut-off rates, the coordination system turns off green CDN edge instances.

2.3 Baseline and Variability of Workloads

The baseline execution included data populations of both green energy availability and user demand for video streaming. The data population was achieved by the Kubernetes-based Jmeter batch jobs. The loader jobs goal is to populate the coordinator database with green energy supply based on using a Weibull distribution, which is a generalization of the Rayleigh distribution described above for wind and a normal distribution for solar. Also, the user demand was populated according to the observed empirical patterns depicted by Figure 1.

We simulated the availability and unavailability of green energy using Jmeter-based workload plan against the coordination system. Our implementation starts green CDN-edge instances opportunistically upon green energy availability. Once a CDN-edge instance declares its availability it processes live workloads.

We use the Kubernetes *livenessProbe* for communication between CDN-edge instance pool and its load-balancer that divert traffic to its pool members. Finally, another workload Jmeter-based simulator generates on-demand streaming calls. This workload simulates end-user demand. It includes HTTP progressive download calls to pre-deployed video media in the CDN-edges.

2.4 Main Execution

In each of the three regional CDN-edge clusters the Kubernetes Jmeter batch jobs that generated green availability traffic to the coordination component were executed. The simulation is comprised of availability indication that are based on Figure 2 and Figure 3. We randomized solar production by using a factor of $\alpha = 0.2$ based on collected data between 2008-2011 in Palm Springs CA, Prescott Airport CPV, AZ and Nevada Solar One, NV [3]. Also, we randomized the wind production by a factor of $\beta = 0.4$ based on collected data between the years 2007-2012 [22]. The demand simulations included a set of calls to the coordinator component spread across 48 hours. The calls originated from three different timezones. The supply simulations consist of wind and solar-based energy time and power windows.

The experiment executions generated two main data traces that we used for the resulting generation computation. The first trace is the simulators logs. The simulator logs includes the demand and supply records. Demand records stored in the Redis key-value store under the key "DemandEvents" followed by timestamp, region and the required compute capacity. The supply calls were stored in the Redis key-value store under the key "SupplyEvents" follows by timestamp, region and supply phase i.e. cut_{in}, cut_{out} or *rated*. For query simplicity the loader ingested three types of records for each supply and demand the by the keys: (1) supply or demand (2) timestamp, and (3) by region. This approach optimized the coordinator queries by timestamp and regions for green CDN-edge instances allocation.

The second trace is the actual allocation logs. It is generated by the coordination system that invokes the Kuberenetes command for green CDN-edge instances initialization and disposal. This was used to determine the green energy utility translated into energy (kWh) that did not use grid energy sources.

2.4.1 Limitations

Every supply and demand was recorded three times to ease the query process. This approach was used since Redis provided limited query abilities by different keys. This approach might suffer data inconsistency issues where a supply metric was successfully committed to one key recode but missing on other key. Production systems should add extra safety gates when ingesting data. We used Redis because of its popularity in the Kubentese community. However, our approach is not limited to Redis or other database systems for that matter.

When measuring the green energy overall utility, we used the container initialization and disposal as indication that green energy utility was used. Specifically, we used the *'kubectl logs POD'* command based on the assumption that the coordination system invocation commands are tightly

[9]US Population Distribution retrieved from https://www.census.gov/popclock/data

[10]Simulation code and data retrieved from https://github.com/yahavb/green-content-delivery-network

coupled with green energy availability. It is likely that collecting the actual video streaming traces through the various Apache access logs of the CDN-edges will be more accurate.

In the case of a sudden lack of green energy while streaming video a failover occurs. Such failover event relies on domain naming services (DNS), the impact of DNS caching was not included since that might cause streaming delay on the user side. Also, when the coordinator algorithm determines there is enough green energy available it will take grid pods down and activate green pods up in a controlled fashion e.g., one at a time so that no requests are lost during the transition phase. For simplicity, the algorithm avoid that.

2.5 Analysis

The green energy supply simulation plotted in Figure 4 shows the energy generation in MW for both wind and solar sources. The simulated amounts were adjusted to the amount observed in the traces between 2007-2012 [22].

The user workload simulation plotted in Figure 4 follows the observed user patterns depicts in Figure 1. Also, it shows the aggregated green energy availability for each region. The cloud-coordinator uses these data sets to determine if there is enough green energy available before provisioning green-pods and possibly taking grid-pods down.

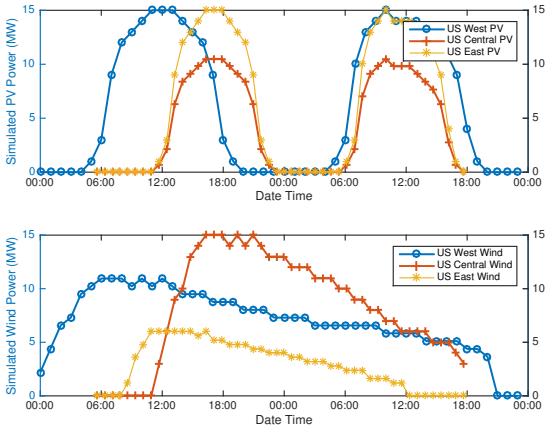


Figure 4: Green energy availability simulated in MW across three regions. Amounts are adjusted to NERL measurements i.e., wind generation in West US moderate, Central US outstanding, East US fair. For solar generation in West US strong, Central US moderate-high, and East US low.

The utility of the green energy was calculated based on the cases where sufficient green energy was available to run the green-CDN pods within the same region. Otherwise cross-regional latencies might degrade the on-demand video experience. The measurements in Figure 5 were adjusted to the estimates of required energy (kWh) for operating the green compute pods. The case where there was negative green energy available it was considered as a miss in the overall utility reckoning.

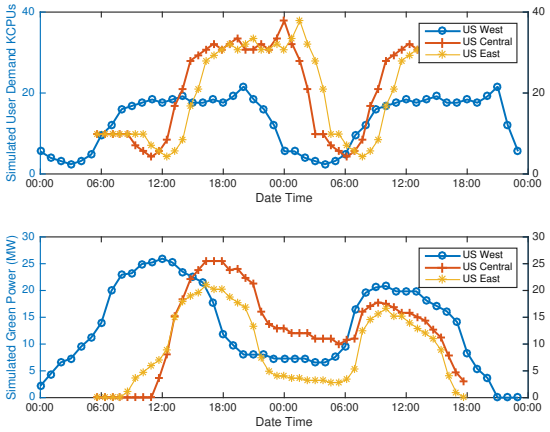


Figure 5: Video on-demand user workload per region adjusted with user population opposed to aggregated green energy availability, solar and wind energy.

2.6 Discussion on Energy Mix when Processing CDN Workloads

The scenario described above simulated the usage of green 801.3 *kWh* out of total 3642 *kWh* to process the video on-demand streaming workload. i.e. 22% by opportunistic matching. When counting the utility per region West US used 42% of the green energy. Central US used 28% of the simulated green energy. East US utilized only 18% as the initial ratio between user demand and green energy availability was relatively low. Although West US reached 40% utilization it contributed nationally only to the 20% portion it contains from the entire experiment test set. By way of comparison, Jeff Barr of AWS noted that their data centers utilize a 28% cleaner power mix[11]. Extrapolating the simulation results to the initial assessment in Equation 1 yields to a saving of:

$$(18.26(GWh) \cdot 1.89) \cdot 0.1(\$/kWh) \cdot 22\% = \$759,250.8/year$$

3 Clean Energy Mix Evaluation for Offline System

In the following section we evaluate a coordination component that harmonizes *analytics* jobs demands with available compute resources powered by green energy resources. Such resources will be published to the coordination system through a resource availability tuple {*region,cut-in,rated,cut-off,power-efficiency*}, where *region* and *power-efficiency* indicates solar or wind based energy and *region, cut-in, rated* and *cut-off* of those energies.

Analytics job demand includes the specific region, total-job workload, load-factor, as well as contract deadline stermed tuple. The *load-factor* indicates the required number of CPU cores per the *total-job-workload*. The region indication will optimize the match between the supply and demand. Also, the *total-job-workload* and the *deadline* will be checked against the *cut-in-rated, cutoff* time for wind or *power-efficiency* for solar, based on the published *load-factor*.

[11] Amazon Web Services Sustainability reference retrieved from https://aws.amazon.com/about-aws/sustainability/

We will suggest a hybrid data center structure that does not deviate from the common data-center architecture. The core difference lies on an automatic transfer switch (ATS) that switches between different available power sources: generator, grid or clean-energy when available. In both cases the data-center design does not change and requires incremental changes only by adding clean-energy power sources to the datacenter's ATS's (Figure 6).

Figure 6 shows a simplified data-center power distribution that supports clean energy sources. In a data-center with available clean-energy resources for both *serving* and *analytical* systems deployed in brown clusters. Further, few clusters use green resources when there are a viable clean energy and standby for incoming *analytical* workloads. As a mitigation strategy, a compute live migration procedure will be available in case of unpredicted lack of renewable resources during a workload processing with a risk for SLO violation.

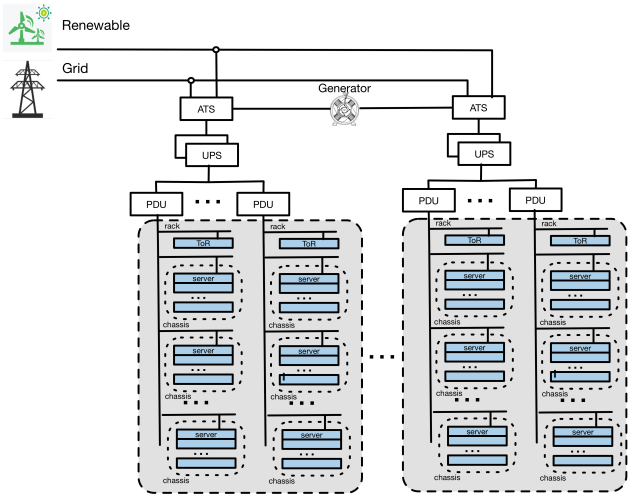


Figure 6: Datacenter Architecture - Compute server clusters aggregated by racks and chassis. Racks are fed by separate Power Distribution Unit (PDU) that connects to Uninterruptible Power Supply (UPS) that connects to the primary power source through Automatic Transfer Switch (ATS).

3.1 Experiment Planning

We wish to simulate an isolated group of computing resources so it can operate by various energy resources, especially available green ones. We use a group of leased resources from existing cloud providers to form a virtual-data-centers set that operates in a federated scheme. Each virtual-data-center includes with internal arbitrator component that collects and aggregates internal signals about its utilization and availability. The arbitrator then reports to the central coordination system. We use Apache Mesos[12] for the virtual-data-center abstraction.

Also, we build a highly available coordination component that accepts incoming supply and demand traffic, calculates a potential match, within minutes, and notifies back the cloud-service-provider and the service-provider for transaction completion. We use VoltDB[13] as the database and application server for the coordinator component.

[12]http://mesos.apache.org
[13]http://voltdb.com

Finally, we simulate customer's demand for compute resources through client simulator with a Java-based application that generates pseudo demand traffic to the coordinator service. The coordinator service will run on separate resources pool than the virtual-data-centers and the client simulator.(Figure 7)

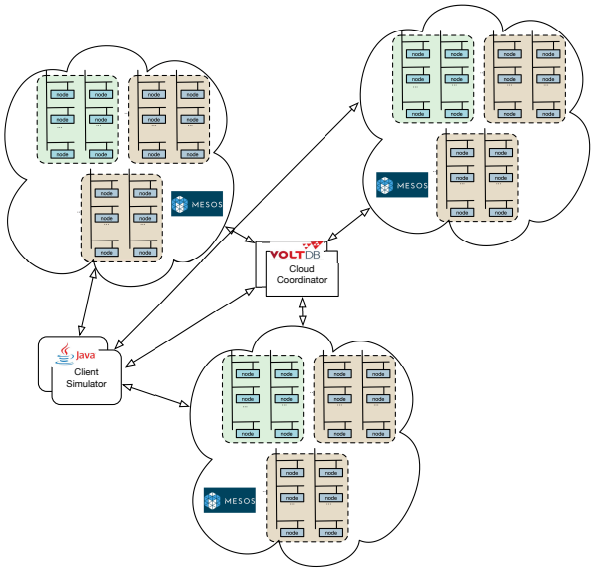


Figure 7: Experiment Architecture - Three virtual clusters deployed on a public cloud services located in a different regions grouped by Apache Mesos. A single VoltDB instance as the coordinator service. Finally, Client simulator that generates pseudo demand that sends data to process

The data collected includes customer workload for jobs processing and jobs deadline. Also, the number of matches found by the coordination component and processed by green clusters without SLO. Finally, we measure the false-positive cases where a match was suggested but not met the SLO's deadline due to a violation that caused by a sudden lack of clean energy resources. We use the data to extrapolate the possible carbon-footprint that could be generated by the used virtual clusters.

The research goal is to show a significant improvement in the carbon emission generation by data-centers. Let $MtCO_2e$ denote the carbon emission. Electrical usage can be consumed by the $Consumers$ set: {cooling, storage, servers, CPU, power} systems and denoted by EU measured in kWh. The average regional carbon dioxide emissions measured in lbs/kWh and denoted by RE_{co_2}. Therefore, the total electrical usage is:

$$EU_{Total}(kWh) = \sum_{c_i \in Consumers} EU(c_i) \tag{4}$$

and the carbon footprint generated by the workload is:

$$MtCO_2e = \frac{EU_{Total}(kWh) \cdot RE_{co_2}(lbs/kWh)}{2,204.6(lbs)} \tag{5}$$

As the experiment uses virtual data-centers, we do not have access to the power consumption by the cooling, power and storage system. The results capture compute jobs durations

measure in server $\frac{CPUcore}{time}$ and use the average Thermal Design Power (TDP) of 200W per core (0.2kW). c_i and t_i denotes the number of cores and time used per job respectively. The electric usage by server ($EU_{servers}$) will be the sum of the electric utilization of the executed jobs that was matched by the coordinator component (Equation 6)

$$EU_{servers}(kWh) = \sum_{i=1}^{jobs} c_i \cdot 0.2 \cdot t_i \quad (6)$$

We assume a linear relation between the power utilization of the cooling, power and storage systems with the servers power utilization. (Equation 4)

3.2 Execution - The Preparation

The Preparation for the execution included the setup of the three virtual data-centers each with 6 machines D4 series with 8 Intel(R) Xeon(R) CPU E5-2660 0 @ 2.20GHz cores, 10Gbps NICs, 28 GB memory and run Ubuntu 15.0. Each virtual data-center is located in a different geo-location. The virtual data-centers were simulated by Mesosphere cluster with three jobtracker(Mesos masters) and three tasktracker(Mesos slaves). We installed Hadoop Cloudera CDH 4.2.1-MR1 on the takstrackers. The Coordinator Database run on a separate resource pool with 1 machine D series similar to the virtual data-centers specification. Finally, the loaders run on a 3 D2 series machines with 2 cores Intel(R) Xeon(R) CPU E5-2660 0 @ 2.20GHz cores, 10Gbps NICs, 7 GB memory and run Ubuntu 15.0. Apache Hadoop ships with a pre-built sample app, the ubiquitous WordCount example. The input data file was created using */dev/urandom* on the takstrackers hosts[14]. The input data file was copied to the HDFS directory that was created as part of the Hadoop preparation (/user/foo/data).

Anticipated Required Deviations from the original job plan. The original plan was to simulate the scenario where the customer keeps his data at a different location than it might be processed. We plan on using a job migration scheme that was originally designed for workload migration across different geo-location[20]. The suggested method optimizes the bandwidth costs of transferring application state and data over the wide-area network. Our experiment generated the data file at the loader host and did not include the job migration. We believe that including the job migration aspect could impact the presented results. However, the job migrations proven efficiency and later studies minimize that deviation.

3.3 Baseline and Variability of Workloads

Baseline. The execution baseline included a load that runs without the coordinator component i.e. loaders generated load to the virtual data-centers resources for 48 hours. The load scenario included a single file generation that was submitted to one of the tasktrackers. The output of each executed jobs included the CPU time spent for each execution. The data collection included the execution log of the command: Hadoop jar loader.jar wordcount /user/foo/data /user/foo/out.

Variability of Workloads. The workload comprises of data files with words that need to be counted using the Hadoop WordCount. The load complexity depends on two factors, the file randomness level, and its size. We rely on the native operating system randomness, and our virtual data-centers and loaders are homogeneous. Therefore, the size is the remaining factor for differentiating workload types. We evaluated the federated-cloud coordinator by generating three load types simultaneously. The three types intend to cover the following cases (1) A match was found between workload and sufficient green energy resources. (2) A match was found, but there was not enough power to complete the job with no SLO deadline violation. (3) Like the former but with SLO deadline violation. The three types will be uniquely distributed across wind and solar based virtual data-centers.

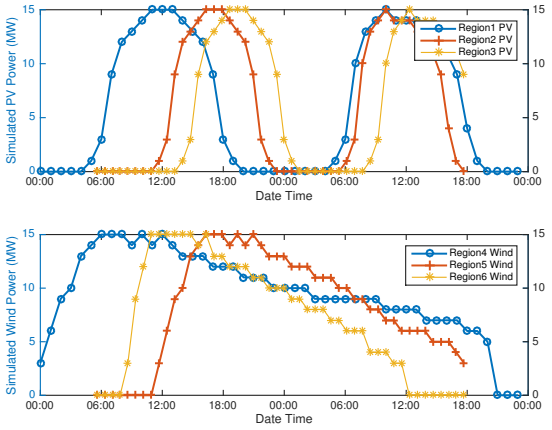


Figure 8: Green Energy availability simulated across six different timezones, three for PV power and the bottom three for wind. These indications are fed, in near-time, to the coordinator database that runs a temporal stored procedure that seeks for match with pending registered jobs.

3.4 Main Execution Issues

In each jobtracker host in a virtual data-center, we executed a simulator that generated green availability traffic to the coordination component. The simulation comprises of availability indication that are based on figure 2 and figure 3. We randomized solar production by using a missing factor of $\alpha = 0.2$ based on collected data between 2008-2011 in Palm Springs CA, Prescott Airport CPV, AZ and Nevada Solar One, NV [22]. Also, we randomized the wind production by a missing factor of $\beta = 0.4$ based on collected data between the years 2007-2012 [22]. The demand simulations included a set of calls to the coordinator component spread across 48 hours triggered by a Monte Carlo simulation inspired by [27]. We built a Rayleigh-based distribution model in the Monte Carlo simulation using Matlab Statistics and Machine Learning Toolbox[15]. The calls originated from three different timezones. Each call comprises of the tuple {region,total-job-workload, load-factor, contract deadlines}. The supply simulators are comprised of

[14]https://github.com/yahavb/GreenCloudCoordination/
[15]https://www.mathworks.com/products/statistics.html

the wind and solar-based energy time and power windows. Supply call to the coordinator includes the tuple {region,cut-in,rated,cut-off,power-efficiency}. The coordinator runs a temporal stored procedure that find a match between the supply and demand data. When a match found, the coordinator generates an assignment call to the jobtracker in the corresponded virtual data-center to execute the jobs. The execution will fetch the data to be processed and report temporal statuses to the coordinator. Every status call generates a check against the current demand levels in the particular region. If the demand changed, and the request cannot be fulfilled, the job considered as false-positive. If the jobs are completed successfully, the request counted as success along with the Core/hour saved.

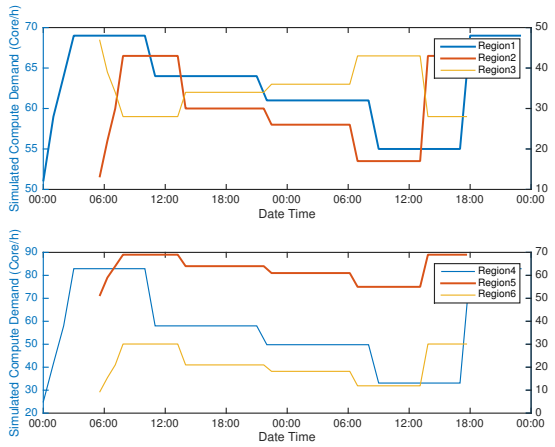


Figure 9: A Monte-Carlo-based simulated compute demand across six different timezones, three for PV power and the bottom three for wind. These indications are fed, in near-time, to the coordinator database that runs a temporal stored procedure that seeks for match with available green energy.

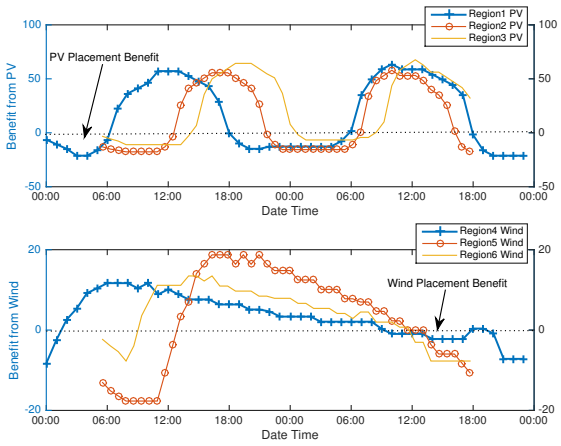


Figure 10: Job placement ranges denoted by the doted line in both the PV and the Wind plots. Any values above the zero levels indicate on potential benefits. However, assigned jobs might not be able to fully processed when unpredicted drop in the availability.

3.5 Analysis

The experiment executions generated three core data logs that we used for the result generation. The first data set is the simulators logs. The simulator logs comprise of the demand records. Demand records stored in the Demand VoltDB table. The supply calls were stored in the Supply VoltDB table. The match transactions stored in the Jobs VoltDB table with status and the number of cores/hour used for the job. Status values can be Success or False-Positive. Figure 9 shows the two supply indications. Both signals were generated based on Equation 2 and Equation 3. The demand data was generated based on known usage patterns that are spread across three timezones.

3.6 Discussion on Energy Mix when Processing Offline Workloads

The experiment simulated the usage of green 708 *kWh* out of total required 3,252 *kWh* for *analytical* systems workload processing i.e. 22% less carbon emission (Equation 5). 1822 *kWh*, 50% of the total workload consumption, was processed by brown energy because of false-positive events i.e. the coordinator assigned a job with no sufficient green recourses to process the job. We believe that optimizing the coordinator algorithm can improve the footprint reduction up to 50%.

Figure 10 shows the job placement ranges denoted by the dotted line in both the PV and the Wind plots. Any values above the zero levels indicate on potential benefits. However, such cases are subject to false positive events that can occur when an unpredicted drop in the availability. Further, such cases utilize the hybrid datacenter power scheme describes in figure 6.

4 Conclusions

The future growth of cloud computing will increase its energy consumption as a fraction of grid power and will cause a significant addition to the ever growing carbon emission since 70% of US power is generated by hydrocarbon fired power plants. Using rapidly emerging green energy for processing cloud computing workloads can limit the anticipated carbon emission growth. However, balancing time varying green energy utilization with time varying energy demands of cloud computing is a complex task that requires sophisticated command and control prediction algorithms beyond the scope of this paper but are emerging in the form of a smart grid system of systems [28]. Our study shows that green energy utilization for on-demand streaming workload is best described as a resource management problem. The solution presented demonstrates real time balance of green resource supply and cloud computing workload demand and utilizes Ubernetes, an open source container cluster manager. The results approximate within 21% those observed in a single cloud instance in the field. Our study also shows that green energy utilization for offline workload processing is a fast data problem. Its solution best utilizes VoltDB, an in-memory database, to allow near-time response for green resources supply and workload demand. Our future work will focus on optimizing the false-positive ratio, to further reducing the cloud computing carbon footprint by up to 50%.

Conflict of Interest The authors declare no conflict of interest.

References

[1] Mell, P.; Grance, T. *Department of Energy* **2011**,

[2] Rao, L.; Liu, X.; Xie, L.; Liu, W. Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment. INFOCOM, 2010 Proceedings IEEE. 2010; pp 1–9.

[3] Shehabi, A.; Smith, S.; Sartor, D.; Brown, R.; Herrlin, M.; Koomey, J.; Masanet, E.; Horner, N.; Azevedo, I.; Lintner, W. *Data Center* **2016**,

[4] Outlook, A. E. *Department of Energy* **2010**, *92010*, 1–15.

[5] Krishnan, S. S.; Sitaraman, R. K. Understanding the effectiveness of video ads: a measurement study. Proceedings of the 2013 conference on Internet measurement conference. 2013; pp 149–162.

[6] Liu, X.; Dobrian, F.; Milner, H.; Jiang, J.; Sekar, V.; Stoica, I.; Zhang, H. A case for a coordinated internet video control plane. Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication. 2012; pp 359–370.

[7] Biran, Y.; Collins, G.; Dubow, J.; Pasricha, S. Coordinating Green Clouds as Data-Intensive Computing. 2016: 3rd Smart Cloud Networks & Systems Conference (SCNS'16). 2016; pp 1–8.

[8] Lim, N.; Majumdar, S.; Ashwood-Smith, P. Engineering resource management middleware for optimizing the performance of clouds processing mapreduce jobs with deadlines. Proceedings of the 5th ACM/SPEC international conference on Performance engineering. 2014; pp 161–172.

[9] Biran, Y.; Collins, G.; Liberatore, J. Coordinating Green Clouds as Data-Intensive Computing. 2016 IEEE Green Technologies Conference (GreenTech). 2016; pp 130–135.

[10] Burns, B.; Grant, B.; Oppenheimer, D.; Brewer, E.; Wilkes, J. *Communications of the ACM* **2016**, *59*, 50–57.

[11] Biran, Y.; Collins, G.; M, B. J.; Dubow, J. *Advances in Science, Technology and Engineering Systems Journal* **2017**, *2*, 553–561.

[12] Nygren, E.; Sitaraman, R. K.; Sun, J. *ACM SIGOPS Operating Systems Review* **2010**, *44*, 2–19.

[13] Maier, G.; Feldmann, A.; Paxson, V.; Allman, M. On dominant characteristics of residential broadband internet traffic. Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference. 2009; pp 90–102.

[14] Barroso, L. A.; Clidaras, J.; Hölzle, U. *Synthesis lectures on computer architecture* **2013**, *8*, 1–154.

[15] Abolfazli, S.; Sanaei, Z.; Ahmed, E.; Gani, A.; Buyya, R. *IEEE Communications Surveys & Tutorials* **2014**, *16*, 337–368.

[16] Ferguson, A. D.; Bodik, P.; Kandula, S.; Boutin, E.; Fonseca, R. Jockey: guaranteed job latency in data parallel clusters. Proceedings of the 7th ACM european conference on Computer Systems. 2012; pp 99–112.

[17] Dean, J.; Ghemawat, S. *URL: http://static. googleusercontent. com/media/resear ch. google. com (diunduh pada 2015-05-10)* **2004**, 10–100.

[18] Lam, W.; Liu, L.; Prasad, S.; Rajaraman, A.; Vacheri, Z.; Doan, A. *Proceedings of the VLDB Endowment* **2012**, *5*, 1814–1825.

[19] Liu, Z.; Lin, M.; Wierman, A.; Low, S. H.; Andrew, L. L. *ACM SIGMETRICS Performance Evaluation Review* **2011**, *39*, 62–66.

[20] Buchbinder, N.; Jain, N.; Menache, I. *NETWORKING 2011* **2011**, 172–185.

[21] Deshmukh, M.; Deshmukh, S. *Renewable and Sustainable Energy Reviews* **2008**, *12*, 235–249.

[22] Energy, G. *NREL Subcontract Report* **2010**, 1–13.

[23] Tiwari, A.; Sodha, M. *Solar Energy* **2006**, *80*, 751–759.

[24] Abdelbaky, M.; Diaz-Montes, J.; Parashar, M.; Unuvar, M.; Steinder, M. Docker containers across multiple clouds and data centers. Utility and Cloud Computing (UCC), 2015 IEEE/ACM 8th International Conference on. 2015; pp 368–371.

[25] Beloglazov, A.; Abawajy, J. *Future generation computer systems* **2012**, *28*, 755–768.

[26] Zhang, Q.; Zhu, Q.; Boutaba, R. Dynamic resource allocation for spot markets in cloud computing environments. Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on. 2011; pp 178–185.

[27] Feijoo, A. E.; Cidras, J.; Dornelas, J. G. *IEEE transactions on Energy Conversion* **1999**, *14*, 1582–1588.

[28] Biran, Y.; Collins, G.; Azam, S.; Dubow, J. Federated Cloud computing as System of Systems. Computing, Networking and Communications (ICNC), 2017 International Conference on. 2017; pp 711–718.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 13-19 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

Design of Cognitive Radio Database using Terrain Maps and Validated Propagation Models

Anwar Mohamed Fanan*, Nick Riley, Meftah Mehdawi

School of Engineering, University of Hull, Cottingham Road, Hull, HU6 7RX, United kingdom.

ARTICLE INFO

Article history:
Received: 10 July, 2017
Accepted: 16 August, 2017
Online: 04 September, 2017

Keywords:
Geolocation database
DTED
Propagation Model
TVWS
Cognitive Radio

ABSTRACT

Abstract — Cognitive Radio (CR) encompasses a number of technologies which enable adaptive self-programing of systems at different levels to provide more effective use of the increasingly congested radio spectrum. CRs have potential to use spectrum allocated to TV services, which is not used by the primary user (TV), without causing disruptive interference to licensed users by using appropriate propagation modelling in TV White Spaces (TVWS). In this paper we address two related aspects of channel occupancy prediction for cognitive radio. Firstly, we continue to investigate the best propagation model among three propagation models (Extended-Hata, Davidson-Hata and Egli) for use in the TV band, whilst also finding the optimum terrain data resolution to use (1000, 100 or 30 m). We compare modelled results with measurements taken in randomly-selected locations around Hull UK, using the two comparison criteria of implementation time and accuracy, when used for predicting TVWS system performance. Secondly, we describe how such models can be integrated into a database-driven tool for CR channel selection within the TVWS environment by creating a flexible simulation system for creating a TVWS database.

1. Introduction

As radios in future wireless systems become more flexible and reconfigurable and available radio spectrum becomes scarce, there is the possibility of using TV white space devices (WSDs) as secondary users in the Broadcast Bands without causing harmful interference to licensed incumbents. Currently, one candidate method could be to utilise a geolocation database approach. The white space device should be able to determine available channel opportunities for a given location by accessing a database of TV White Space (TVWS) channels including data on each transmitter and each site, variable channels, transmitter power, and time of validation [1]. Therefore, the TV channel can be protected from harmful interference by accurate prediction of TVWS using an appropriate propagation model. Design of any wireless network depends on accurate prediction of radio propagation, which impacts deployment and management strategies. In this paper we extended the previous work by investigating the best propagation model among three propagation models (Extended-Hata, Davidson-Hata and Egli), using various terrain data resolutions (1000, 100 and 30 m) and comparing with the real measurements taken around Hull UK using two the performance comparison criteria of implementation time and accuracy. Agreement between the measured and predicted values of path loss has investigated, using MATLAB to analyse and compare the variation of path loss between the measured and predicted values. The terrain profile was extracted from terrain database Global1 and then taken into account in selected propagation models. The flexible cognitive TVWS database system was built using different propagation models to calculate available channels in each pixel of the selected area.

2. Digital Terrain Elevation Data (DTED)

DTED was developed by the US Defense Mapping Agency (DMA) and can be used to improve signal detection accuracy. Currently, the paper has selected only three DTED levels of spatial resolution, which are available to the public [2].

Table 1: Resolution levels of DTED.

DTED Level	Post Spacing	Ground Dist	Row x Column
0	30 arcsecond	~ 1 km	121 x 121
1	3.0 arcsecond	~ 100m	1200 x 1200
2	1.0 arcsecond	~ 30 m	3600 x 3600

*Corresponding Author: Anwar Mohamed Fanan, School of Engineering, University of Hull, Cottingham Road, Hull, HU6 7RX, UK,
Email: Anuarfanan@yahoo.com

The elevation data of the resolution from 30 arcsecond (arcsec) to 1 arcsec (level 0,1 and 2) are avaliable for public use and can be downloaded as different DTED extension files in each resolution. The USGS web application has been used to define the desired research region by specifying latitude and longitude as shown in Figure 1, which identifies the required terrain tiles. Figure 2 classifies all tiles, identified by the web application earthexplorer.usgs.gov .

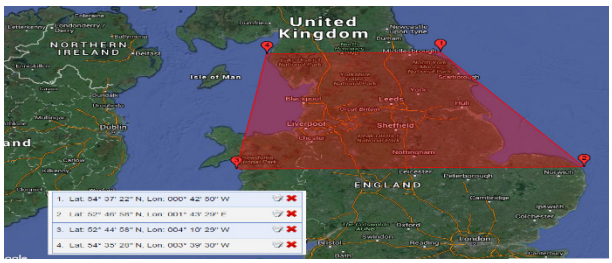


Figure 1. Region selected for terrain data acquisition

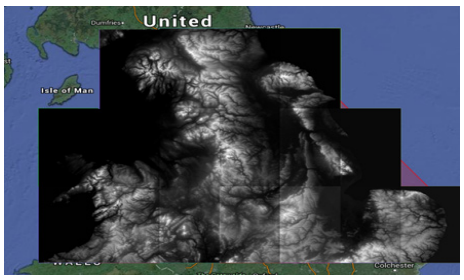


Figure 2. Research region of terrain elevation data by NASA's shuttle radar topography mission (SRTM) with resolution 1 arc sec (30 m)

3. TVWS Geolocation Database

The use of a TV white space geolocation database enables the most effective detection method for prediction of available channels and calculation of TV coverage maps for each pixel in the selected region by using an appropriate propagation model, selected for accuracy and efficiency. The technique can avoid signal detection problems caused by fading effects and shadowing. Construction of the geolocation database requires primary user information including frequency of operation, transmitted power, location, transmission time and height and type of transmit antenna. This information will protect spectrum incumbents from interference from secondary users who will access the database by sending a query to obtain available channels in a given area at a certain time. Furthermore, the geolocation database might have proxy to make queries and identify available channels for WSD [3].

4. Propagation Models

When planning wireless communication systems and designing wireless networks, the accuracy of the prediction of propagation characteristics of each environment should be taken into account. One of the most significant parameters, which can be provided by propagation prediction, is large-scale path loss, which affects directly the coverage of a base station placement and its performance. However, using field measurements to obtain these parameters without depending on propagation models is time-consuming and costly. The following subsections provide a brief explanation of several appropriate empirical propagation models [4] including the Extended Hata, Davidson-Hata and Egli models.

4.1. Extended Hata Model

The Extended Hata model was derived from Hata-Okumura which is widely used for signal prediction in urban areas, wholly based on measured data that have been collected in Tokyo Japan. Also, it does not have any analytical explanation, but includes empirical factors that depend on the type of environment. This model can be used in different environment by adding some correction factors to meet the requirements of ITU-R and also extending range up to 100 km as shown in the following equation:

$$PL_{Ex\_Hata} = A + B\log(d)^b + C \quad (1)$$

$$A = 69.55 + 26.16\log_{10} f_c - 13.82\log_{10} h_{tx} - a(h_{rx})$$

$$B = 44.9 - 6.55\log_{10} h_{tx}$$

Where f_c represents the carrier frequency (150 to 1500 MHz), h_{tx} is the height of the base station antenna (m) and h_{rx} is the height of receive antenna (m). The distance from transmitter to receiver is d km. The value of the correction factors a(h_{rx}) and C depend on the type of environment. In small and medium sized cities and metropolitan areas the value of C will be 0, while in other environments such as suburban and rural areas, different equations are used [5]. The factor b denotes an extended range up to 100 km as shown in the following equation.

$$b = \begin{cases} 1, & ,d_{Km} \le 20 \\ 1 + (0.14 + 0.000187 f_c + 0.00107 h_t')(\log_{10} d_{Km}/20))^{0.8} & ,d_{Km} > 20 \end{cases} \quad (2)$$

$$h_t' = \frac{h_t}{1 + 7\times 10^{-6} h_t^{\,2}}$$

4.2. Davidson-Hata Model

$$PL_{David} = PL_{Hata} + A(h_t, d_{km}) - S_1(d_{km}) - S_2(h_t, d_{km}) - S_3(f_{MHz}) - S_4(f_{MHz}, d_{km}). \quad (2)$$

$$PL_{Hata} = 69.55 + 26.16\log_{10} f_{MHz} - 13.82\log_{10} h_t$$

$$- a(h_r) + (44.9 - 6.55\log_{10} h_t)\log_{10} d_{Km} \quad (3)$$

where A is a factor extending distance up to 300 Km, S1 and S2 are correction factors for extending transmitter height up to 2500m, while the factors S3 and S4 extend the frequency range over 30 to1500 MHz [6]. *a(h$_r$)* is a correction factor for receiver antenna height.

4.3. Egli Model

The Egli model was introduced by John Egli in 1957. The model includes a terrain model and was derived from real measurements of UHF and VHF television transmissions conducted in several large cities. It then used a point-to-point model for predicting the total path loss [7]. Thus, this model is commonly used for point to point communications to predict path loss in an urban or rural area, where transmission has to go over an irregular terrain between a fixed transmitter and receiver in the frequency range 40 to 900 MHz. The following equation illustrates the path loss calculation of the Egli model:

$$PL_{Egli} = \begin{cases} 20\log_{10} f_{MHz} + p_0 + 76.3, & h_2 < 10 \\ 20\log_{10} f_{MHz} + p_0 + 83.9, & h_2 < 10 \end{cases} \quad (4)$$

$$p_0 = 40\log_{10} d_{Km} - 20\log_{10} h_t - 10\log_{10} h_r \quad (5)$$

where h_t denotes the height of the transmitter antenna (m), h_r is the height of the receiver antenna (m), the distance between transmitter and receiver is denoted by d km and the transmission frequency is f (MHz) [8].

5. Field Measurement and Data Collection

The main goal of selection of various positions at which to conduct measurements is to examine the signal strength behavior in different environments at various distances from the transmitter and to observe how the terrain affects the received signal. The measurements have been taken at 23 locations distributed randomly around Kingston-upon Hull, UK as shown in Figure 3.

Figure 3. Measurement equipment and geographical location in Hull and surrounding areas.

The measurement equipment used includes an omnidirectional antenna (covering the frequency range 174 to 230 MHz (VHF) and 470 to 790 MHz (UHF) with gain of 3.5 dBi) and spectrum analyser (Agilent E4407B, frequency range 9 kHz to 26.5 GHz) which was connected with a laptop computer by using a general purpose interface bus. A Matlab program on the laptop received raw data and stored them in (bin) files. In addition, the measurement locations were determined by using a mobile GPS application.

6. ANALYSIS of Models' Performance

6.1. Propagation Path Loss Analysis

The main criterion for model assessment is path loss. A simulation program was implemented in Matlab, using channel 33 to conduct the comparison between the three propagation models and the measured results. In order to compare the real measurements with different propagation models, the path loss should be extracted from the real measurements by using the following equation in each location [9].

$$PL = TX + TX_{Gain} + RR_{Gain} - RP \quad (7)$$

Where TX denotes the transmitted power, transmitting antenna gain is represented as TX_{gain}, PL is the path loss, receiving antenna gain is denoted as PR_{gain} and RP is the received power, dBm.

To evaluate the propagation models against real measurements, several parameters might be used to identify the most accurate propagation model. The error between predicted and measured path loss values was calculated by Equation 8 and the average error calculated by Equation 9.

$$e_i = E_{pi} - M_{pi} \quad (8)$$

$$\bar{e} = \frac{1}{N}\sum_{i=1}^{N} e_i \quad (9)$$

In which e_i is denoted as the difference between the calculated path loss E_{Pi} and measured path loss M_{pi} derived from measured received power in each location.

Equations 8 and 9 are then used to calculate the standard deviation, Equation 10, whilst Root Mean Square Error (RMSE) is calculated by Equation 11, which also depends on the average error calculated in Equation 9.

$$\sigma = \sqrt{\frac{1}{N-1}\left(\sum_{i=1}^{N} e_i^{\,2} - N.(\bar{e})^2\right)} \quad (10)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n} |e_i|^2} \quad (11)$$

To remove the influence of dispersion from the overall error, a significant metric can be derived from RMSE by subtracting the standard deviation of the absolute value of the error to obtain the Spread Corrected Root Mean Square Error (SCRMSE), as illustrated in Equations 12 and 13:

$$E_i = |e_i| - \sigma \quad (12)$$

$$SCRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n} |E_i|^2} \quad (13)$$

6.2. Diffraction Factor Based on Terrain Profile Database

A common phenomenon that exemplifies the wave property of EM waves and light is diffraction, which is the bending of EM waves around obstacles. Diffraction is considered as a non-line of sight (NLOS) propagation mechanism which may occur when the propagation path is obscured by a barrier such as a mountain or hill or man-made obstacles including buildings. Diffraction can

be the cause of significant signal weakness at the reception site, due to the presence of some of the aforementioned barriers between the transmitter and receiver. There are two types of diffraction. "Shadow diffraction", occurs when the received signal is blocked by obstacles and the received field strength will be decreased when the reception site is within the shadowed area. The second case, which occurs if the impediment is underneath the LOS, is called "lit diffraction," and commonly leads to multi-path interference. Shadow diffraction is the one of the main reasons for increased the path-loss. One of the common diffraction models is the single knife edge model, explained by Huygens's Principle, which states that when an electromagnetic wave is obstructed by a natural or man-made obstruction, the obstruction acts as a secondary source for creating a new wavefront which then propagates into the geometric shadow region of the obstruction [10].

6.3. Terrain Profile-based Diffraction Model

One of the main effects of the terrain profile is to cause diffraction or bending of EM waves around obstacles such as mountains, hills or man-made structures which obscure the direct path.

In previous work [11], we considered terrain resolutions of 1 and 30 arcsec and determined that each of them had advantages and disadvantages in terms of accuracy and calculation time.

In this work, we attempt to improve these results by investigating a third resolution value between 1 and 30 arcsec to improve the compromise between accuracy and implementation time. For example, whilst calculating diffraction using the three different resolutions and investigating its effect on the received signal, we noticed that in the location approximately 38 km along the path shown in Figure 4, in the 30 arcsec resolution the elevation value is 100 m, whilst when using 3 arcsec and 1 arcsec resolution, the elevation values are approximately 86 and 83 m respectively.

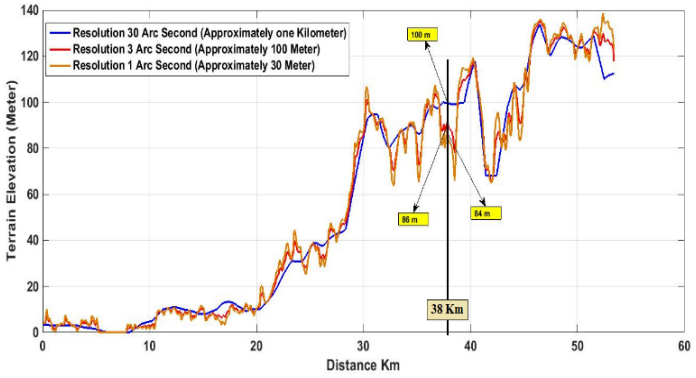
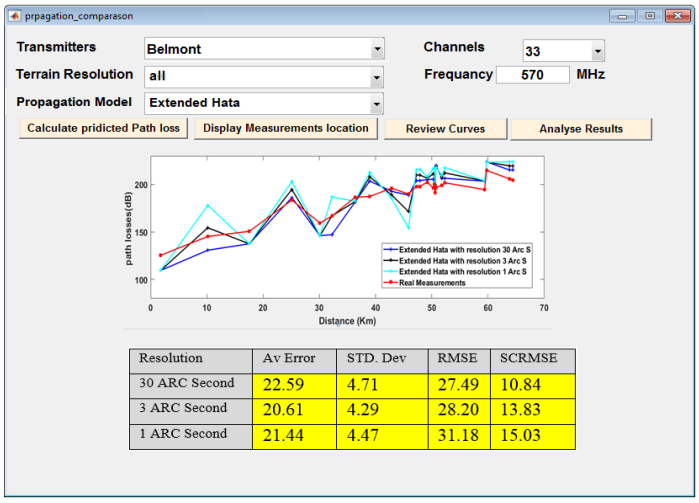


Figure 4. Terrain elevation data for the path from University of Hull to Belmont TV Transmitter in different resolutions

Using 30 arcsec resolution takes a short time for the implementation process but has less accuracy. When using 1 arcsec, we have good accuracy but a long time for the implementation process. However, using 3 arcsec resolution produces the best results in terms of the compromise between accuracy and implementation time.

7. Comparison and Results

The measurement study covered the area around the city of Hull, which was represented to measure the UHF TV band from 470 to 790 MHz with consideration of all radio and TV stations that feed the whole Hull area. Most of the channels transmitted into the area originate from the Belmont and Emley Moor transmitters (see Figure 3). The results of comparison of predicted path loss with measurements for two cases (excluding and including terrain modelling) are presented by using the previously defined criteria average error, standard deviation, RMSE and SCRMSE. Figure 5 shows our Graphical User Interface (GUI) of expected results including path loss curves for each propagation model and a table of calculated parameters.



Resolution	Av Error	STD. Dev	RMSE	SCRMSE
30 ARC Second	22.59	4.71	27.49	10.84
3 ARC Second	20.61	4.29	28.20	13.83
1 ARC Second	21.44	4.47	31.18	15.03

Figure 5. GUI showing propagation model comparison in the selected measurement locations

This analysis may be undertaken for all selected measurement points, by flexible selection of the transmitter name, terrain resolution, propagation model and transmitted channel. Results corresponding to all measurement locations and comparison of the three propagation models with real measurements along with parameter analysis, are discussed and classified in the following sections.

7.1. Influence of Terrain Resolution on Results from the Extended Hata Model.

In this and in the following sections, terrain profile databases with various spatial resolutions and equivalent single knife edge diffraction have been used to calculate the diffraction factors and then evaluate their impact on the performance of the propagation models. The results in Table 2 indicate that 30 arcsec resolution used with the Extended Hata model can be considered the best fit to the measured data with low error when applying the diffraction factor at different terrain resolutions. It can be clearly seen in Figure 6 that the behavior of path loss was influenced by diffraction, when compared with the path loss derived from measured data.

Thus, the propagation behaviour has been affected in most measurement locations when applying the terrain variation with 30 arcsec resolution. The impact of the propagation model is obvious after the third measurement point, where the first three points might be situated within the line of sight and the 1 km resolution results might have missed terrain features situated

along the path which might cause destructive or constructive diffraction. Thus, using 1 km resolution might not give accurate results. Also when using 1 arcsec resolution, the terrain variation becomes worse in this model, which additionally requires a significantly longer processing time. However, the 3 arcsec resolution gives only slightly different results and does not need a long time for calculation of terrain profile.

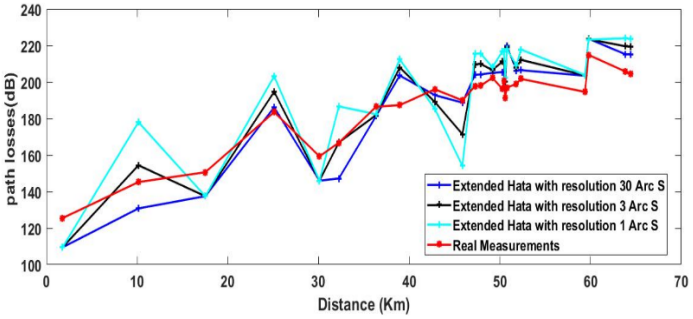


Figure 6. Impact of terrain resolution on results from the Extended Hata model

In Table 2 we can observe how the error statistics have been impacted by the diffraction factor and how the SCRMSE value are decreased in the selected terrain resolution. The results indicate that the 30 arcsec and extended Hata model has the best results of the SCRMSE, at 10.84 dB. On the other hand, the 3 arcsec result is seen to have less error compared with 1 arcsec by about 1.2 dB.

Table 2: Fitted Extended Hata model in different terrain resolution

Resolution, arcsec	Av Error, dB	STD. Dev, dB	RMSE, dB	SCRMSE, dB
30	23.34	4.87	23.52	11.07
3	21.43	4.47	22.82	9.95
1	22.23	4.64	25.02	9.81

7.2. Influence of Terrain Resolution on Egli Model Results.

Due to the nature of the terrain profile near the transmitter sites, which includes rough terrain and hills, the use of 1 arcsec and 3 arcsec resolutions will clearly affect the Egli propagation predictions, as illustrated in Figure 7. Here it may be seen clearly that there are large changes in the path loss at the distance of 48 km and that at other locations such around 52 km there is less variation where the receiver might be in line of sight of the transmitter.

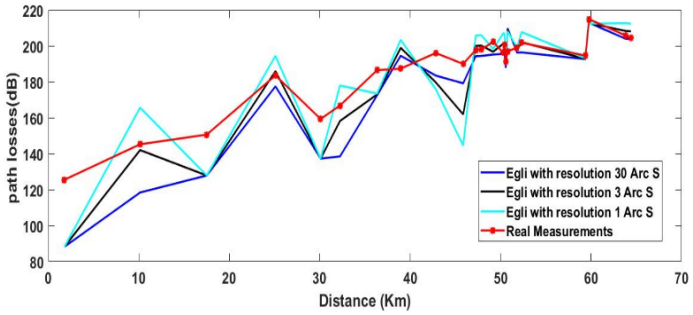


Figure 7. Impact of terrain resolution on the Egli model results

The error of the SCRMSE for the both 1 arcsec and 3 arcsec are slightly different by about 0.14 dB as can be seen in Table 3, whereas the SCRMSE values for 30 arcsec resolution have increased by 1.12 dB.

Table 3: Fitted Egli model in different terrain Resolutions

Resolution, arcsec	Av. Error, dB	STD. Dev, dB	RMSE, dB	SCRMSE, dB
30	22.59	4.71	27.49	10.84
3	20.61	4.29	28.20	13.83
1	21.44	4.47	31.18	15.03

7.3. Influence of Terrain Resolution on Davidson Model Results.

The Davidson model is clearly affected by terrain resolution in a similar manner to the Extended-Hata model at all selected resolution values, as shown in Figure 8.

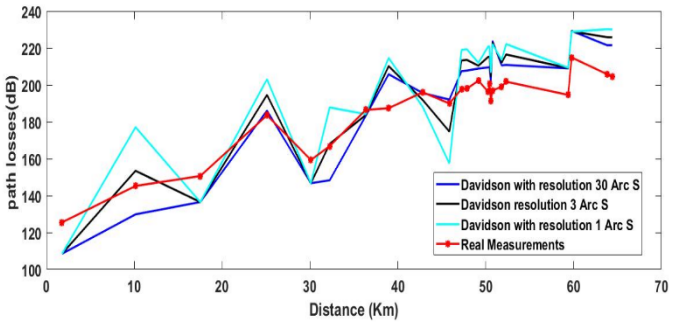


Figure 8. Impact of different terrain resolution on the Davidson model

It can be observed in the statistics of Table 4 that the SCRMSE is increased while the resolution value is decreased as Table 2 of the Extended-Hata behavior.

Table 4: Fitted Davidson model in different terrain Resolutions

Resolution, arcsec	Av Error, dB	STD. Dev, dB	RMSE, dB	SCRMSE, dB
30	24.46	5.10	30.70	12.48
3	22.44	4.68	31.50	15.60
1	22.99	4.79	34.38	16.94

According to the advantage and disadvantage of both previous results in terms of accuracy and processing time, we observed that 30 arcsec has short time and less accuracy, while 1 arcsec has high accuracy and long processing time. Therefore, the 3 arcsec resolution in the Egli model gives the best result when taking these chosen criteria into account when comparing SCRMSE results between different terrain resolutions.

8. Design of Flexible System for Creating TVWS Database by Using Different Propagation Models

Based on the previous results, which indicate that the Egli model is the best among the models that have been chosen for

comparison with the real measurements, a flexible system has been built that performs many functions related to propagation modelling and calculation of signal strength in each pixel. Among these tasks, it is possible to determine any geographic area based on latitude and longitude between two concentric points. In addition, it can be determined that the size of each pixel will affect the implementation time and propagation accuracy. Also, the system can perform three major operations at the same time to create a database for a selected geographic region that can be easily used when connected to the white space devices (WSD). In this work, for illustration, we considered only six transmitters, but more can be added using the "Add Transmitter Detail" button, as shown in Figure 9.

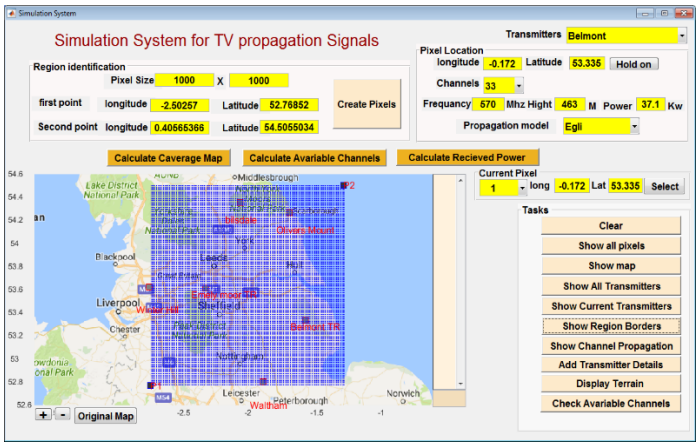


Figure 9. Display of all pixels in the selected region in the flexible simulation system for creating TVWS database.

8.1. Methodology for Calculation of Received Power

The second methodology to be implemented after creating the pixel file is to calculate the receiver power in each pixel in the frequency range 470 to 790 MHz, by considering the selected transmitters. The processing time depends on the number of pixels, propagation model and also the terrain resolution level. The process will be conducted only once to create a complete database of all the predicted TV signals in each pixel, as shown in Figure 10, which can be used for the next stages.

8.2. Methodology for Calculation of Available Channels

The main goal of the system is to calculate available channels with high accuracy and then store all available channels of each pixel in the database, in a way which makes it easy to retrieve the data from WSDs. All of the transmitter information, such as height, channels and transmitted power, is stored previously in the database. The process takes into account all channels of the selected transmitters that might be received in a specific pixel, considering the weak signals as well.

8.3. Methodology for Calculation of Coverage Map

This methodology must be used to translate the database that has been stored to show as a visual map of different levels of signal strength in the selected region for each transmitter, which are then stored in different files in the database as shown in Figure 11.

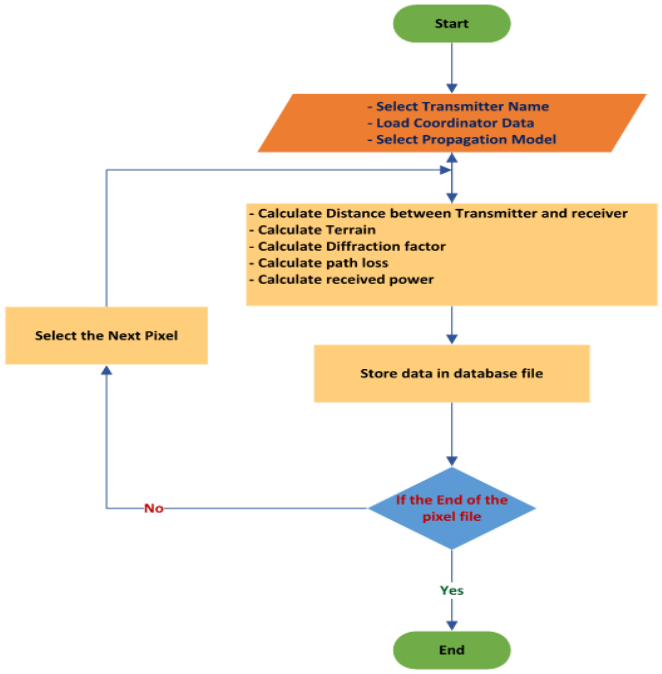


Figure 10. Algorithm of the received power calculation

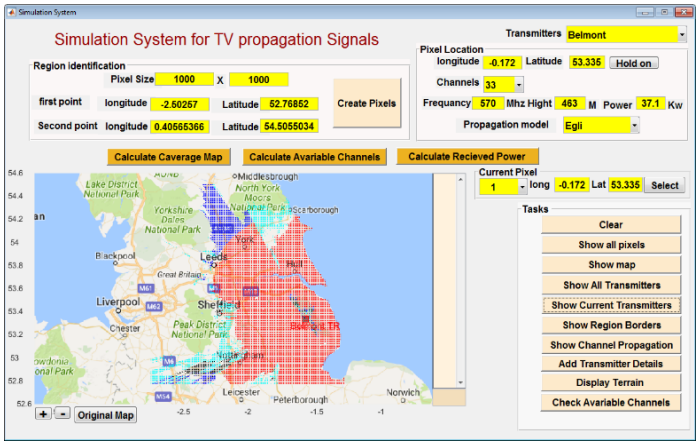


Figure.11 Display of the propagation signals for channel 33 using Egli model.

9. Conclusion

In this paper, TV signal strengths are calculated using various propagation models and then compared with real measurements that have been conducted in various locations. Using a single knife edge model to calculate the diffraction factor with consideration of terrain profile data at different resolutions, we investigate and prove how the terrain data resolution impacts the accuracy and implementation time of the propagation models. We have improved and extended results that have been published in our conference paper in 2016 [11]. RMSE and SCRMSE are the main criteria taken into account to assess the propagation models in different terrain resolutions. The results show that the Egli model still gives the best results when account is taken of the terrain profile data at a resolution of 3 arcsec (100m), providing SCRMSE of 0.14 dB, with shorter computation time and similar accuracy as compared with 1 arcsec. On the other hand, RMSE

and SCRMSE values of the extended Hata and Davidson models are still increasing and have poor performance when terrain data resolution is decreased. Therefore, 3 arcsec is considered the best resolution that can be used to calculate diffraction factor. The results also show that the Egli model is the best model giving a consistently good fit to measured data among other selected models and, with appropriate terrain data, will provide useful input to a system for facilitation of the cognitive radio decision process. In addition, the main benefits for designing the flexible system is to create a TVWS database for a specific area, by selecting the optimum pixel size, adding appropriate transmitter information and choosing a suitable propagation model. In future work, the system will be developed to use additional propagation models at various terrain data resolutions, providing a clear understanding of the differing results between propagation models taking into account the terrain resolution.

Acknowledgment

The lead author is grateful to his supervisor for his unlimited supporting to develop this study, by providing the suitable solutions and suggestions.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Gurney, D., Buchwald, G., Ecklund, L., Kuffner, S.L. and Grosspietsch, J., 2008, October. Geolocation database techniques for incumbent protection in the TV white space. In New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on (pp. 1-9). IEEE.

[2] National Imagery and Mapping Agency, "Digital Terrain Elevation Data," 30 May, 1990. 'www.globalsecurity.org/ intell/systems/dted.htm.

[3] Nekovee, M., 2009, October. A survey of cognitive radio access to TV white spaces. In 2009 International Conference on Ultra Modern Telecommunications & Workshops (pp. 1-8). IEEE.

[4] Iskander, M.F. and Yun, Z., 2002. Propagation prediction models for wireless communication systems. Microwave Theory and Techniques, IEEE Transactions on, 50(3), pp.662-673.

[5] P. Pardeep, P. Kumar and B. S. Rana, "Performance evaluation of different path loss models for broadcasting applications," *American Journal of Engineering Research (AJER)* , 3(4), pp.335-342, 2014.

[6] K. Stylianos, P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zettas, and J. Cosmas, "Comparison of Longley-Rice, ITU-R P. 1546 and Hata-Davidson propagation models for DVB-T coverage prediction, " In BMSB, pp. 1-4, 2014.

[7] J. J. Egli, "Radio propagation above 40 MC over irregular terrain. Proceedings of the IRE, " , 45(10), pp.1383-1391,1957.

[8] M. Hope, A. B. Bagula, M. Zennaro, and G. Lusilao-Zodi, "On the impact of propagation models on TV white spaces measurements in Africa," *International Conference on. IEEE*, pp.148-154, May 2015.

[9] P. Prajesh, and R. K. Singh, "Investigation of outdoor path loss models for wireless communication in Bhuj," *International Journal of Electronics and Communication Engineering & Technology (IJECET),* Volume 3, Issue 2, pp. 171-178, September 2012.

[10] J. You-Cheol, "Diffraction analysis and tactical applications of signal propagation over rough Terrain," *Air Force INST of Tech Wright-Patterson AFB OH School of Engineering*, no. 97J-01, Dec 1997.

[11] A. M. Fanan, N. G. Riley, M. Mehdawi, and M. Ammar, 2016, November. Comparison of propagation models with real measurement around Hull, UK. In Telecommunications Forum (TELFOR), 2016 24th (pp. 1-4). IEEE.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 20-27 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

Smart Grid Operational functions and Control Challenges by Implementing SSSC Tailored to Optimize performance in between Qatar and KSA on the GCC Electrical-power grid

Tariq Masood[*1], Muhammad Tajammal[2], Samer Karim Shah[3], Ghulam Hashmi[4], Suhail Aftab Qureshi[5], D. P Kothari[6]

[1]*Asset Integrity, Department Qatar Petroleum, Dukhan Operations, Qatar*

[2]*University of Hail, Hail, Kingdom of Saudi Arabia*

[3]*Maintenance Department, QAFAC, Doha Qatar*

[4]*Power System Engineering Department, Saudi Aramco, Dhahran, KSA*

[5]*University of Engineering and Technology, Lahore Pakistan*

[6]*J. D. College of Engineering & Management, Nagpur India*

ARTICLE INFO

Article history:
Received: 19 August, 2017
Accepted: 06 September, 2017
Online: 30 September, 2017

Keywords:
Reactive voltage
Voltage regulator
Power flow

ABSTRACT

This research work is novel technique to control and optimize SSSC (Subsynchronous Series Controller) functions with degree of precision in between Qatar and Kingdom of Saudi Arabia. The SSSC model developed and simulated in order to identify and determine its control and functioning parameters by introducing new tuning parameters based on that the SSSC can be adjusted stringently to witness desired results lead to address outstanding reactive power management issue. The proposed new parameters are contributing significantly to control SSSC functions in multiple directions in a power system network in between QATAR and Kingdom of Saudi Arabia at different time-based transmission contingencies on the GCC Electrical-power grid. Strategically, the SSSC capacity and capability can be utilized fully in between Qatar and Kingdom of Saudi Arabia by introducing and optimizing its control and tuning parameters more tangibly under both steady and dynamic states

1. Introduction

The SSSC has operational controllability results are clearly indicating that introduction of SSSC in between Qatar and Kingdom of Saudi Arabia power network will equitably improve the power system loadability, curtailing the losses and value-added sustainability of the power system enactment by addressing control and operational issue throughout the GCC Electrical-power grid.

Hereafter, new SSSC optimization technique can thus be magnificently expended for this type of power system process optimization. This work published in 13th the IET International AC/DC Conference held in Manchester, February 14-16, 2017. Whereas a SSSC determines and validates the three control and effective limits which have been made-to-order at (minimum (+/-) medium (+/-), and maximum (+/-) compensation. These rheostats functioning limits are regulated by constituting their consequent PI control-values by exercising D.J. Cooper PID regulator Performance to achieve the following [1].

(1) Augmented Power flow
(2) Developed consistency & controllability
(3) Augmented angle and voltage constancy.

Therefore, FACTS Controllers are convincing candidate technology options centered on that following benefits can be perceived.

[*]Corresponding Author: Tariq Masood, Asset Integrity, Department Qatar Petroleum, Dukhan Operations, Qatar | Email: t.masood.dr@bath.edu

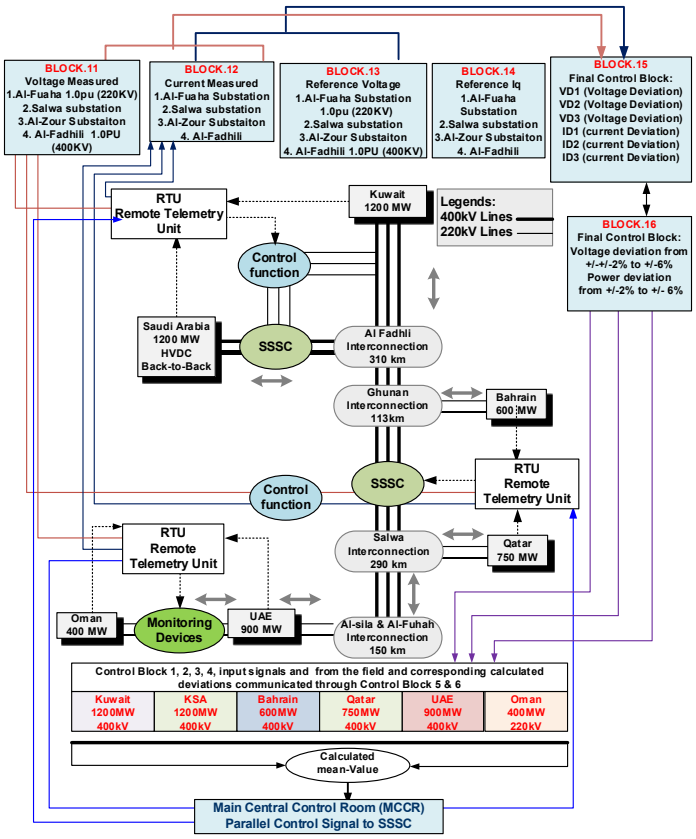


Figure 1. SSSC structure in between Qatar and Kingdom of Saudi Arabia

Figure 1 shows the GCC power network by retaining SSSC at augmented whereabouts by exercising Wideband-Delphi-Technique.

- Model reveals how much notional concentrated power can be dispensed in between QATAR and Kingdom of Saudi Arabia.
- At maximum conjectural power, what is borderline voltage
- Model reveals how much reactance must be retained by the orthodox series capacitance to double the power conveyance in between Qatar and Kingdom of Saudi Arabia if vital and prerequisite.
- Compute the introduced voltage by the series capacitor in between QATAR and Kingdom of Saudi Arabia
- By SSSC how much introduced and instilled/injected reactive voltage Vr is upheld at diverse compensation of SSSC in between Qatar and Kingdom of Saudi Arabia and Compute the determined power supplied at persistent introduced voltage Vr
- If the load angle decreases from 71 to 0 how much power will be produced by SSSC in between Qatar and Kingdom of Saudi Arabia.

2. GCC Electrical-power grid background

The GCC power-grid is constituted and employed into eight premeditated power system operational directions in order to meet national and industrial customers necessities in a consistent and viable demeanors at the GCC Electrical-power grid-network as shown in Figure 1 as considered and described with facts and figures, this is also published in [1] .

3. SSSC Operational Analyses

Principally, the SSSC-series counterbalance device location is not very much perilous it can be employed anyplace on the GCC Electrical-power grid. In this case the SSSC has been retained at the borderline to decrease the power-line transmission impedance unnaturally by SSSC which is controlling to upsurge the power flow in the power transmission system in (5) and (6) determine how much line impedance (X_C) has been Controlled and Managed after employing series but stable capacitor's K_{series} factor.

SSSC encompasses of capacitors and reactors to diverge the power-line impedance on domineering need basis vigorously. Predominantly, the SSSC injects the voltage in a series into the power transmission network at the midpoint of Kingdom of Saudi Arabia and Qatar whereas instilled/injected the voltage at midpoint Vr has been computed in (7). The SSSC stipulates or expends reactive-power in capacitive/inductive control-mode of functions on the GCC power-grid, wherein the SSSC instilled/injected voltage in phase quadrature with the line current of the power network.

Figure 2 denotes maroon trend in the graphical presentation of the SSSC processes in capacitive and inductive mode of functions. Practically, blue trends denote how much real power transferred can transferred when electrical-power transmission network compensation carried-out capacitive, inductive mode of operations or neutral without any compensation factor. Figure 3 signifies that a SSSC has a main and great influence on power steadiness and load flow, but ithas limited impact on voltage profile enhancement as exhibited in (28) that V_m, V_s, V_r has negligible or no effect with SSSC [2].

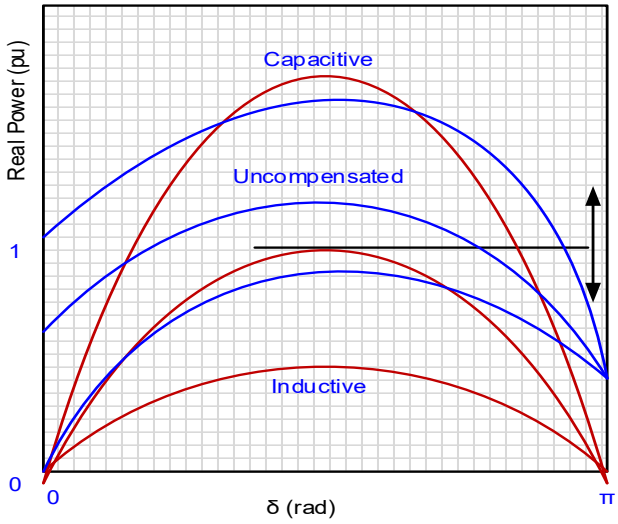


Figure 2. Signify the compensation factor by SSSC [2]

Principle of Operating parameters	SSSC Controller	Impact in between Qatar and KSA		
		Load Flow	Stability	Voltage Quality
Series Compensation: To Vary the power transmission line impedance	FSC: Fixed Series Compensation	o	ooo	o
	TPSC: Thryrister protected Series Compensation	o	ooo	o
	TCSC: Thryrister controlled Series Compensation	oo	ooo	o
Benchmark	o: Low or no Impact	o o: Medium		o o o: Strong

Figure 3. Denotes the SSSC impact.

3.1. By employing Series Fixed Capacitor to upsurge the Power on the GCC Electrical-power grid

Herein determined conjectural power can be dispensed and distributed without conservative or non-conventional series compensation, the voltage structured transitorily but not vigorously and unvaryingly which is outfitted at 800 km long electrical-power transmission network from Kingdom of Saudi Arabia to Qatar or vice versa along with its attendant substations. Now if series capacitor is connected at midpoint voltage as stated in Figure 3 to resource and expend sufficient reactive power to counterbalance as prerequisite. Equation (5) expended to compute series compensation factor to convalesce the power system steadiness and double the power at power transmission lines reactance X_C= 136.626 Ohms as intended in (6) and (7) [3].

$$K_{series} = \frac{X_c}{2 \times Z_a} cos\frac{\theta}{2} \tag{5}$$

$$X_c = 2 \times K_{series} \times Z_a tan\frac{\theta}{2} = 136.626\ Ohms \tag{6}$$

$$P_{max} = \frac{V^2}{Z_a - X_C \times Sin(43.26^{\circ})} = 1223\ MW \tag{7}$$

Equation (7) reveals the concentrated power by adapting power transmission line reactance by series capacitor Units.

3.2. By implementing Series-Capacitor to upsurge the Power on the GCC Electrical-power grid

The SSSC has a tremendous competence to upturn the electrical Power flow and improve dynamic steadiness to supplant usual and uneconomical fixed series capacitors are expended to govern the prerequisite reactive current at midpoint in between Qatar and Kingdom of Saudi Arabia for recompense [4].

Instill/Injected current with series capacitor is computed by exercising in (8), (9).

$$I_m = \frac{Vsin\frac{\delta}{2}}{(1 - K_{series})Z_a sin\frac{\theta}{2}} = 2.1159pu \tag{8}$$

Instilled/injected voltage as given below

$$\hat{V_r} = 1.5237 \times tan\frac{\theta}{2} = 0.8881pu \tag{9}$$

Equation (9) reveals the in a Phase injected/instill voltage

$$V_r = \hat{V_r} \times \frac{V}{\sqrt{3}} = 205.03\ kV \tag{10}$$

Equation (10) reveals the instilled/injected voltage in between QATAR and Kingdom of Saudi Arabia. Herein determined power flow can be sustained with unremitting voltage inserted by the SSSC at midpoint. Firstly, the SIL computed and determined by using Equation (11)

$$\hat{P} = \frac{P}{P_n} = Acos\delta + Bsin\frac{\delta}{2} \tag{11}$$

amount of real power can be dispensed in line with P_n, this is identified as a SIL in contradiction of the real electrical-power delivery.

Where

$$A = \frac{1}{sin\theta} = 1.40115\ \ ;\ \ B = \frac{\hat{V_r}}{2sin\frac{\theta}{2}} = 1.14735 \tag{12}$$

Equation (13) reveals electrical Power flow with unremitting reactive voltage instilled/injected whereas the $\hat{P}$ is determined

when $\frac{d\hat{P}}{d\delta} = 0$

$$\frac{d\hat{P}}{d\delta} = Acos\delta - Bsin\frac{\delta}{2} \tag{13}$$

Equation (14) resultant after substituting $\frac{d\hat{P}}{d\delta} = 0$ value

$$0 = Acos\delta - Bsin\frac{\delta}{2} \tag{14}$$

Unassumingly $cos\delta$ and $sin\delta$ are factorized as resultant in (15)

$$cos\delta = cos\left(\frac{\delta}{2} + \frac{\delta}{2}\right)\ \ ;\ \ sin\delta = sin\left(\frac{\delta}{2} + \frac{\delta}{2}\right) \tag{15}$$

Where

$$\cos(A + B) = CosA.CosB - SinA.SinB \tag{16}$$

Equation (17) resultant after replacing $cos\delta$ and $sin\delta$ values in (16)

$$cos\left(\frac{\delta}{2} + \frac{\delta}{2}\right) = cos\frac{\delta}{2}.cos\frac{\delta}{2} - sin\frac{\delta}{2}.sin\frac{\delta}{2} \tag{17}$$

Equation (17) is the streamlined and resultant in (18)

$$cos\delta = cos^2\frac{\delta}{2} - sin^2\frac{\delta}{2} \tag{18}$$

Equation (18) $cos\delta$ value subtract in (14) and resultant from (19)

$$0 = A\left(cos^2\frac{\delta}{2} - sin^2\frac{\delta}{2}\right) - Bsin\frac{\delta}{2} \tag{19}$$

Whereas

$$cos^2\frac{\delta}{2} = \left(1 - sin^2\frac{\delta}{2}\right) \tag{20}$$

Equation (20) $cos^2\frac{\delta}{2}$ subtracted in (19) and developed new (21)

$$0 = A\left(\left(1 - sin^2\frac{\delta}{2}\right) - sin^2\frac{\delta}{2}\right) - Bsin\frac{\delta}{2} \tag{21}$$

By adding the $x = sin\frac{\delta}{2}$

$$A(1 - x^2 - x^2) - Bx = 0 \quad (22)$$

$$-2Ax^2 + A - Bx = 0 \quad (23)$$

$$-(2Ax^2 - A + Bx) = 0 \quad (24)$$

$$2Ax^2 + Bx - A = 0 \quad (25)$$

Equation resultant from (25) to convert into Quadratic form.

$$\int(x) = \left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a} \quad (26)$$

Equation (26) expressed the Quadratic function as detailed below with factors.

$$\left(x + \frac{B}{4A}\right)^2 + \frac{(4 \times 2A \times -A) - B^2}{4A} = 0 \quad (27)$$

$$\left(x + \frac{B}{4A}\right)^2 - \frac{8A^2 + B^2}{4A} = 0 \quad (28)$$

Equation (27) and (28) Quadratic roles are defined and extracted and resultant in (29)

$$x = -\frac{B}{4A} \pm \frac{\sqrt{8A^2 + B^2}}{4A} \quad (29)$$

Equation (29) subtracted A and B factors values as defined and calculated in (12) and resultant in new (30)

$$x = -\frac{1.401}{4 \times 1.4735} \pm \frac{\sqrt{8 \times 1.4735^2 + 1.401^2}}{4 \times 1.4735} \quad (30)$$

Finally, the "x" values are signified as given below:

$$x = 0.5314 = sin\frac{\delta}{2} =\rightarrow \delta = 64.2° \quad (31)$$

Equation (31) reveals the $x = 0.5314$ and electrical power angle resultant

Where

$$\hat{P}_{max} = Asin\delta + Bcos\theta \quad (32)$$

$$\hat{P}_{max} = Asin(64.2°) + Bcos(45.5°)\hat{P}_{max} = 2.06519 \quad (33)$$

Equation (32) subtract the A, B, $\emptyset$ and δ values and resultant the $\hat{P}_{max}$ value.

$$P_{max} = \hat{P}_{max} \times \frac{400^2}{316.228} = 1044.9\,MW \quad (34)$$

Equation (34) signifies the concentrated electrical power dispensed and distributed in the GCC Electrical-power grid by retaining SSSC to uphold persistent Vr vigorously in the Midpoint of QATAR and Kingdom of Saudi Arabia as illustrated in Figure 4 [5].

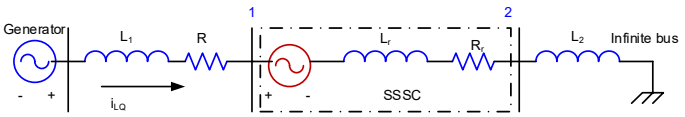


Figure 4: SSSC connected in between Qatar and Kingdom of Saudi Arabia

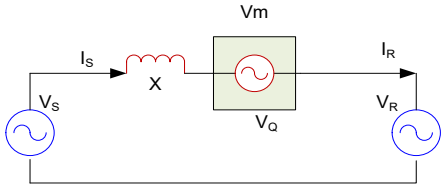


Figure 5: SSSC connected in between Qatar and Kingdom of Saudi Arabia

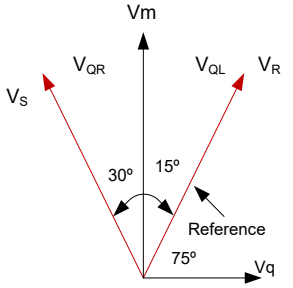


Figure 6. Phasor diagram of voltage profile

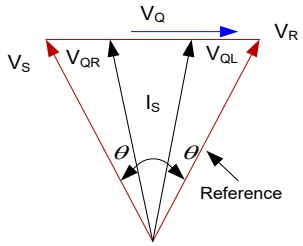


Figure 7. SSSC connected at the midpoint and instilled/injected V_Q toward V_R

Figure 5: signifies that counterbalance device employed in the middle of the of the power transmission line, this is voltage source only dispense reactive electrical power not a real power. Figure 6 displays the influence of location if counterbalance device employment on voltage profile, the voltage on left side of the counterbalance device would be V_{QL} and right side V_{QR}. Figure 7 illustrates that the series counterbalance device employed in the middle the worst voltage-profile occurred at each side of the counterbalance device, whereas the voltage-profile vector aligns with the current vector [5]. Therefore, the midpoint voltage-profile is the same with/without SSSC functions as exhibited in (37) and (38) SSSC power flow with $\delta = 0$[6].

$$P_{max} = 1.147 \times \frac{400^2}{316.2} = 580MW$$

$$\hat{V}_m = \frac{cos\frac{\delta}{2}}{cos\frac{\theta}{2}} \qquad \delta = 0 \quad (35)$$

Equation (35) reveals the midpoint voltage-profile if $\delta = 0$ and resultant midpoint voltage denotes from (36)

$$\hat{V}_m = \frac{1}{cos \times 22.76°} = 1.1\ pu \quad (36)$$

Equation (35) reveals the midpoint voltage if $\delta = 0$ therefore, $V_s and\ V_R$ voltage-profile are computed in (37) and (38). Sending end from the Kingdom of Saudi Arabia and receiving end at QATAR

$$V_s = V_m \angle 0 - j\frac{V_r}{2} = 1.171\angle 0 - j\frac{0.8881}{2} \quad (37)$$

$$V_R = V_m \angle 0 + j\frac{V_r}{2} = 1.171\angle 0 - j\frac{0.8881}{2} \quad (38)$$

3.3. SSSC input Data

Table 1: operating parameters and associated countries

Countries	Power Exchange
To/from Qatar to main transmission line	400kV, 50Hz, 750MW
To/from Kingdom of Saudi Arabia to main power-transmission line	400kV, 50Hz, 600 MW

Table 2: Conventional Reactive Power

Existing Power System	Reactive Power	Distance
VAR Demand at Al Jasra substation Kuwait	850MVAR	Ghunan Substation to Slawa Substation total distance 288Km
Reactive Power demand at Al-Al-Zour Substation	500MVAR	Slawa Substation total distance = 97Km

Table 3: Multivariable Controller's Configuration

Al-Zour Substation	Iq regular: Kp 14dB; Ki: 0.014 dB	Counterbalanced and regulated voltage at 0-2%
Rated SSC: +/- 500 MVAR	Ref V: 1.0 pu (400kV)	(VC voltage change from 2% to 4%)
P (proportional): 0.33, 0.38, 0.40, I (integral) 0.7, 0.9, 1.2	Droop: 0.033pu/100MVA; Kp:14dB; Ki: 3500dB	(VC voltage change from 4-6%)

4. Results and Discussion

4.1. Maximum level Vr voltage-injection

It has been simulated and authenticated that further real power can be up-surged in the power transmission line to touch its design specification by up-surging the SSSC compensation factor from 50% to 65% reactive voltage instilled/injected voltage would be 0.8864pu which is equivalent to 204kV and current Im = 2.1pu and $\boldsymbol{\delta = 63.87°}$ will activate consequently at different functioning condition total 1200MW power will be produced and distributed on the GCC electrical-Electrical-power grid. This is known as a determined compensation by implementing a SSSC and its compensation factor from 50% to 65% to obtained projected power produced and delivered in the electrical power transmission network on the GCC electrical-Electrical-power grid.

4.2. Medium level Vr voltage-injection

As simulated by Matlab/SIMULINK and scientific/ mathematical model V_r instilled/injected voltage 0.5936pu which is uniform and match to 137kV at 50% SSSC reactive voltage compensation whereas the midpoint current I_m = 1.4459pu, and $\boldsymbol{\delta = 71.32°}$ will activate consequently in this functioning circumstance to a total 1079 MW power which can be produced and distributed on the GCC Electrical-power grid as validated by the results. Infact, QATAR and Kingdom of Saudi Arabia system has been devised to produce and distribute concentrated 1200MW electrical-power in the GCC Electrical-power grid during a peak load, therefore Margin is obtainable to upsurge the power up to its projected specification. This is also identified as a medium compensation by a SSSC compensation factor from 20% to 50%. Therefore, the Control function block expended to readjust PI controller response as exhibited in Figure 10 to diverge the firing angle of the Thyrister control [7].

4.3. Minimum level Vr voltage-Injection

In the third challenge of operations, further this model was simulated and authenticated by exercising SSSC Controller. At minimum compensation of the SSSC compensation real power reduce considerably in the power transmission line to touch its projected specification by decreasing the SSSC compensation factor from 5% to 20% reactive voltage instilled/injected voltage would be 0.3710pu which is equal to 86.67kV and current Im = 0.974pu and $\boldsymbol{\delta = 77.57°}$ will activate consequently in this functioning condition total 986MW power will be produced and distributed on the GCC Electrical-power grid. This is identified as a minutest compensation by using a SSSC Controller and its compensation factor from 5 to 20% to achieve projected electrical power delivery of transmission network on the GCC Electrical-power grid. As instilled/injected voltage up-surged based on compensation from 5% to 20% there is considerable power flow and steadiness enhancement but very insignificant influence on voltage-profile as verified simulated consequences in Figure 8 (a) reactive power injection-waveform 8(b) indicates the total power delivered in between Qatar and Kingdom of Saudi Arabia, this is also computed in equation (34). The PI controller configured and demonstrated its operational response in Figure 9. Detailed configuration discussed as given in the appendices [8].

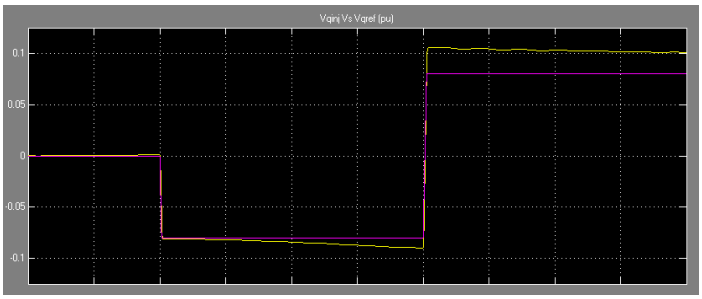


Figure 8 (a). Reactive power injection on the GCC Power Grid

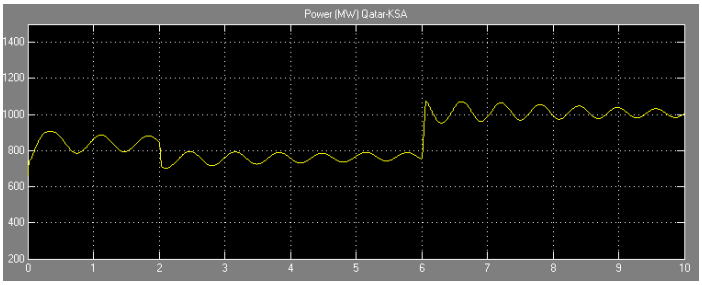


Figure 9 (b). Total active power delivered on the GCC Electrical-power grid

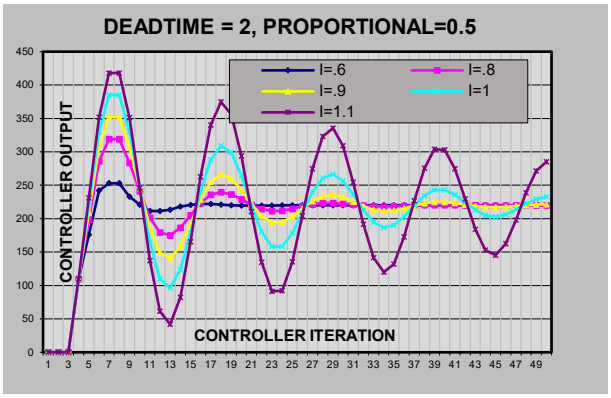


Figure 10. demonstrates the PI controller parameters

5. Conclusion

Fundamental creativity of the SSSC application in the GCC Electrical-power grid transmission line compensation is the subject of extensive significance. In this study, it has been computed and simulated how much maximum power can be produced and distributed in between QATAR and Kingdom of Saudi Arabia without any process vagueness. It has been reasonable how much reactance must be provided by the series capacitance to double the power supply in between Qatar and Kingdom of Saudi Arabia. It also has been also identified and unwavering amount of reactive voltage to be instilled/injected in between Qatar and Kingdom of Saudi Arabia in order to provide concentrated power delivery on the GCC Electrical-Power grid.

If the load angle as decreased below from 65 to 0 amount of electrical-power will be produced by SSSC in between Qatar and Kingdom of Saudi Arabia. Herein, the GCC electrical-power network will exclusively upsurge the power system loadability, decrease the losses and improve sustainability of the electrical-power system functioning. The results also illustrates that SSSC is multipurpose equipment with stupendous active capability to improve power system stability margin on the GCC Electrical-power grid. Based on these results, SSSC is a very strong-candidate to be instigated at GCC Electrical-power grid. It also demonstrates optimistic influence on neighboring countries' power system functions at the GCC Electrical-power grid.

Table 4 Terms and Abbreviations

Name	Abbreviations
GCC	Gulf Cooperative Council
Ir	Reactive current
Za	Line Impedance
Vm	Midpoint voltage
Kshunt	series compensation factor in the network
Pn	Surge Impedance
KSA	Kingdom of Saudi Arabia

Conflict of Interest

No conflict of interest in this paper has been identified.

Acknowledgment

I extend my special gratitude to all my coauthors in this paper in particularly late Prof. Dr Abdel-ETY Edris (Icon of FACTS Technology) at Santa Clara University for their unprecedented support and exceptional contribution to make this study a great success. I am also very grateful to my little Doll Princesses Fatima Masood, who kept me busy full time not giving me space during my research on smart control

References.

[1] Tariq Masood, R.K. Aggarwal, S.A. Qureshi "Novel Control of a SSSC Connecting the Oman and United Arab Emirates in the GCC power Grid Interconnection" published in IJAPE International Journal of Automation and Power Engineering, Volume 2, issue 1 January, 2013.

[2] Tariq Masood, Abdel Aty Edris, Suahil Aftab Qureshi, Muhammad Tajammal, Murtaza Hashmi DP Kothari, Samer Kareem "SSSC tailored to optimization performance in Between United Arab Emirates and Oman on the GCC Power Grid" published on 14-16 February, 2017 in the 13th IET International Conference on AC and DC power Transmission Conference.

[3] Mohammed Rasetgar, Mehdi Saradarzeda "An Improved D-SSSC Voltage and Current Load Balancing Control Strategy Under Unbalanced Load published in 25th Iranian IEEE ICEE 2017.

[4] Mohammed Abdul, Mohd Akram, "Power System Stability Enhancement using Static Synchronous Series Compensator" Published in IEEE (SCOPES)-2016

[5] S. Jamali, A. Kazemi, H. Shateri "Voltage Inversion due to Presence of SSSC on Adjacent Lines and Distance Relay Mal-Operation" published in 2008 IEEE Conference.

[6] D. Rai, S.O. Faried, "Impact of Imbalanced Phase Operation of SSSC on Damping Subsynchronous Resonance published in 2011 IEEE Conference,

[7] Tariq Masood "Improvement of Voltage and power flow control in the GCC Electrical-power grid by using coordinated FACTS Controllers" Ph.D Thesis, published in 2013 at University of Bath, UK.

[8] Tariq Masood, R.K. Aggarwal, Nasser Al-Eamdi "Enhanced Power Flow/Voltage Control in the GCC Grid by SSSC (SUBSYNCHRONOUS SERIES CONTROLLER) Robustness" published by Lambert in 2013 ISBN: 978-3-659-40995-0.

Appendix No.1
PI Controller response computed

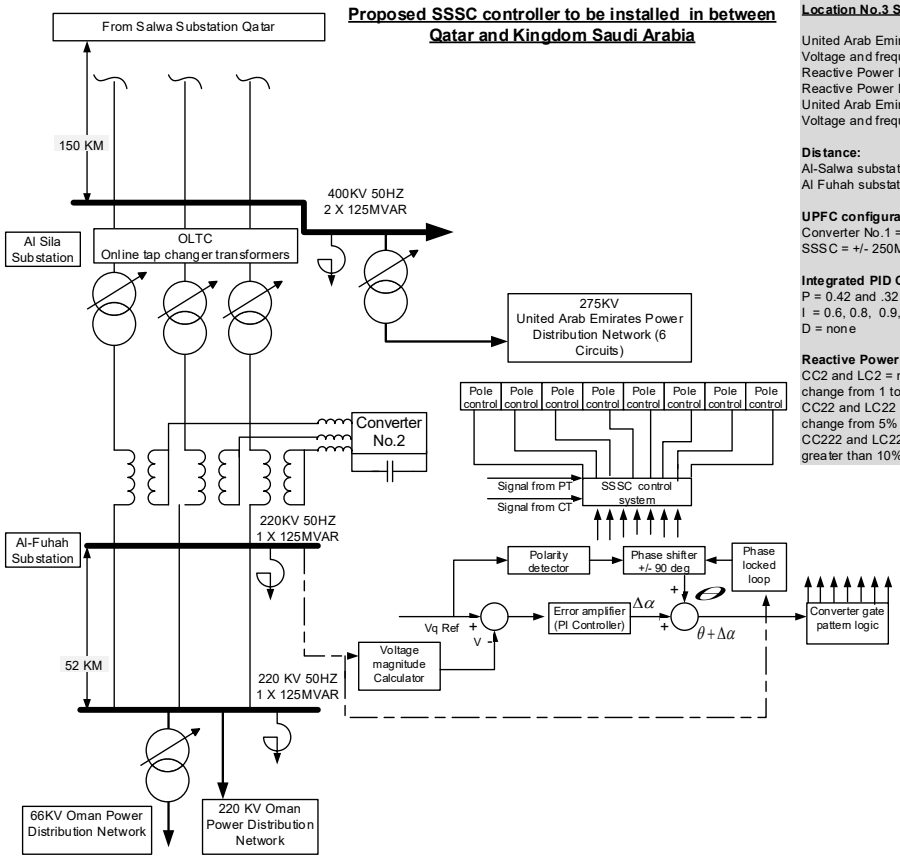
Proportional	0.5
Integral	0.8
SP	220.0

Iteration	PV	ERROR	ABSOLUTE ERROR	OUTPUT	INTEGRAL
				0.00	0.00
				0.00	0.00
1	0.0	220.00	220.00	110.00	88.00
2	0.0	220.00	220.00	198.00	176.00
3	0.0	220.00	220.00	286.00	264.00
4	110.0	110.00	110.00	319.00	308.00
5	198.0	22.00	22.00	319.00	316.80
6	286.0	-66.00	66.00	283.80	290.40
7	319.0	-99.00	99.00	240.90	250.80
8	319.0	-99.00	99.00	201.30	211.20
9	283.8	-63.80	63.80	179.30	185.68
10	240.9	-20.90	20.90	175.23	177.32
11	201.3	18.70	18.70	186.67	184.80
12	179.3	40.70	40.70	205.15	201.08

13	175.2	44.77	44.77	223.47	218.99
14	186.7	33.33	33.33	235.65	232.32
15	205.2	14.85	14.85	239.75	238.26
16	223.5	-3.46	3.46	236.53	236.87
17	235.7	-15.65	15.65	229.05	230.61
18	239.7	-19.75	19.75	220.74	222.71
19	236.5	-16.53	16.53	214.45	216.10
20	229.0	-9.05	9.05	211.58	212.48
21	220.7	-0.74	0.74	212.11	212.19
22	214.5	5.55	5.55	214.96	214.41
23	211.6	8.42	8.42	218.62	217.78
24	212.1	7.89	7.89	221.72	220.93
25	215.0	5.04	5.04	223.45	222.95
26	218.6	1.38	1.38	223.64	223.50
27	221.7	-1.72	1.72	222.64	222.81
28	223.4	-3.45	3.45	221.09	221.43
29	223.6	-3.64	3.64	219.61	219.98
30	222.6	-2.64	2.64	218.66	218.92
31	221.1	-1.09	1.09	218.38	218.49
32	219.6	0.39	0.39	218.68	218.64
33	218.7	1.34	1.34	219.31	219.18
34	218.4	1.62	1.62	219.99	219.83
35	218.7	1.32	1.32	220.49	220.36
36	219.3	0.69	0.69	220.70	220.63
37	220.0	0.01	0.01	220.64	220.63
38	220.5	-0.49	0.49	220.39	220.44
39	220.7	-0.70	0.70	220.09	220.16
40	220.6	-0.64	0.64	219.84	219.91
41	220.4	-0.39	0.39	219.71	219.75
42	220.1	-0.09	0.09	219.70	219.71
43	219.8	0.16	0.16	219.79	219.78
44	219.7	0.29	0.29	219.92	219.89
45	219.7	0.30	0.30	220.04	220.01
46	219.8	0.21	0.21	220.11	220.09
47	219.9	0.08	0.08	220.13	220.13
48	220.0	-0.04	0.04	220.11	220.11
49	220.1	-0.11	0.11	220.05	220.06
50	220.1	-0.13	0.13	220.00	220.01

Appendix No. 2

SSSC Controller configuration



Location No.3 SSSC input data:

United Arab Emirates to Oman = 400 MW
Voltage and frequency = 220KV and 50HZ
Reactive Power Demand at UAE Side = 125 MVAR
Reactive Power Demand at Oman Side = 125 MVAR
United Arab Emirates Power demand= 900 MW
Voltage and frequency = 400KV, 50 HZ

Distance:
Al-Salwa substation (Qatar) to Al-Sila substation (UAE) =150 KM
Al Fuhah substation (UAE) to Al-Wasset Substation (Oman) = 50 KM

UPFC configuration:
Converter No.1 = SSSC (Series connection Al-Fuhah Substation UAE)
SSSC = +/- 250MVAR

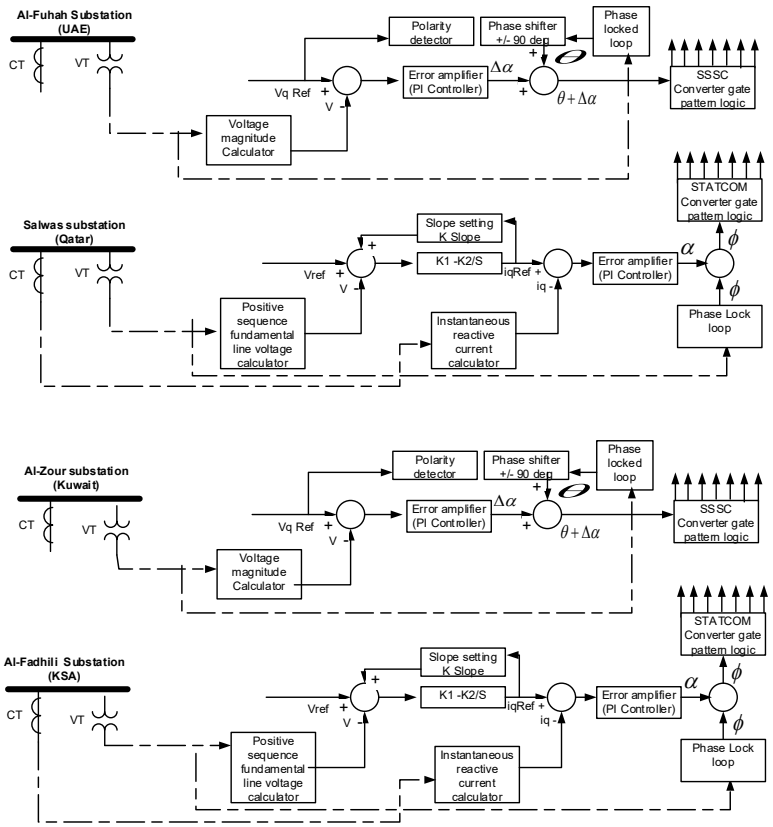
Integrated PID Controller's input data:
P = 0.42 and .32
I = 0.6, 0.8, 0.9, 1.0, 1.1
D = none

Reactive Power Compensation codes are derived in both Mode operations:
CC2 and LC2 = minimum compensation capacitive or inductive (VC voltage change from 1 to 5%)
CC22 and LC22 = medium compensation capacitive or inductive(VC voltage change from 5% to 10%)
CC222 and LC222 = full compensation capacitive or inductive(VC voltage change greater than 10%)

Appendix No. 3

SSSC Centralized and Decentralized Control

Decentralized controller configuration and implementation to control and monitor individual FACTS device operations at GCC Power Grid



Location Controller's input data

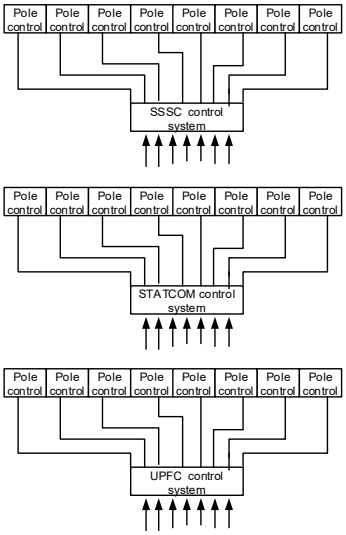
UPFC configuration:
Converter No.1 = STATCOM (shunt connection at Kingdom Saudi Arabia side)
Converter No.2 = SSSC (Series connection Kuwait side)
STATCOM = 250 MVAR
SSSC = 250 MVAR
RATED UPFC = +/- 500 MVAR
Location in between Al-Zour (Kuwait) and Al-Fadhil substation KSA
Droop =0.45/100VA

STATCOM configuration:
STATCOM (shunt connected at Salwa substation in between Bahrain and Doha)
RATED STATCOM = +/- 500 MVAR
Droop =0.37/100VA

SSSC configuration:
Converter No.1 = SSSC (Series connection Al-Fuhah Substation UAE)
SSSC = +/- 250MVAR

Integrated PID Controller's input data:

P = 0.42 and .32
I = 0.6, 0.8, 0.9, 1.0, 1.1



ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 28-39 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

A novel beamforming based model of coverage and transmission costing in IEEE 802.11 WLAN networks

Mehdi Guessous*, Lahbib Zenkouar

Mohammadia School of Engineers, Electrical Engineering Departement, Mohammed V University In Rabat, 10000, Morocco

ARTICLEINFO

Article history:
Received: 21 August, 2017
Accepted: 19 September, 2017
Online: 30 Spetember, 2017

Keywords:
Beamforming
Co-channel Interferences
Radio Resource Management
Transmit Power Control
Wireless Coverage
Wireless Local Area Network

ABSTRACT

IEEE 802.11 WLAN indoor networks face major inherent and environmental issues such as interference, noise, and obstacles. At the same time, they must provide a maximal service performance in highly changing radio environments and conformance to various applications' requirements. For this purpose, they require a solid design approach that considers both inputs from the radio interface and the upper-layer services at every design step. The modelization of radio area coverage is a key component in this process and must build on feasible work hypotheses. It should be able also to interpret highly varying characteristics of dense indoor environments, technology advances, service design best practices, end-to-end integration with other network parts: Local Area Network (LAN), Wide Area Network (WAN) or Data Center Network (DCN). This work focuses on Radio Resource Management (RRM) as a key tool to achieve a solid design in WLAN indoor environments by planning frequency channel assignment, transmit directions and corresponding power levels. Its scope is limited to tackle co-channel interference but can be easily extended to address cross-channel ones. In this paper, we consider beamforming and costing techniques to augment conventional RRM's Transmit Power Control (TPC) procedures that market-leading vendors has implemented and related research has worked on. We present a novel approach of radio coverage modelization and prove its additions to the cited related-work's models. Our solution model runs three algorithms to evaluate transmission opportunities of Wireless Devices (WD) under the coverage area. It builds on realistic hypotheses and a thorough system operation's understanding to evaluate such an opportunity to transmit, overcomes limitations from compared related-work's models, and integrates a hierarchical costing system to match Service Level Agreement (SLA) expectations. The term "opportunity" in this context relates also to the new transmission's possibilities that related-work misses often or overestimates.

1 Introduction

Designers face problems of two natures when building WLAN networks: issues that are inherent to radio resources' planning and those concerning their integration with other network parts and upper-layer services in response to developing customers' application needs such as mobility, real-time, interactive applications, business intelligence, and on-presence analytics. An accurate design should consider thoroughly both aspects and base its decision on a solid theoretical approach, on-field feedbacks, and best practices.

Design best practices dictate running two site surveys: one at the beginning of the project, before network's deployment, and a second, after network's deployment, to confirm resources' planning results and to calibrate the system's parameters in accordance to the real network's condition. An optional survey may be required if radio environment or its utilization has

*Mehdi Guessous, 7 avenue du General de Gaulle 91120 Palaiseau France, +33(0)614945485 & mehdiguessous@research.emi.ac.ma

https://dx.doi.org/10.25046/aj020604

changed and might affect the overall network capacity. The first survey qualifies mainly the exposure to on-reach foreign autonomous radio systems, and radio characteristics of the surface to cover (walls, floors, and other obstacles). The second site survey confirms the conformance of the actual deployment to design constraints in two ways:

1. by running a passive in-data-path survey that is based on passively reported statistics such as RSSI, SNR, PHY, MAC errors, packet count, etc. at radio interface level,

2. and an active one where in-data-path patterns simulate wireless client's traffic classes to gather actively statistics on Jitter, RTT, or any other measure concerning the rendered service, application or experience.

This approach is not very accurate as it lies on a set of preconfigured unchangeable parameters that result from special case deployments and testing contexts, which may be completely different from the on-field reality. In the market-leading implementations such as Cisco TPC or Aruba ARM, we observe that at a certain point, a reference is made to an "ideal" preconfigured parameter. The third neighbor's RSSI and a predefined hysteresis threshold [1] paremeters are used in the first case, and the "coverage index" [2] parameter in the second case, to decide on transmit power level planning network wide.

Inaccuracies come also from studies when they model the system as a pre-shaped radio coverage area in the form of a hierarchy of disks that delimit transmission ranges from interference or no-talk ones. Some limitations arise also from no-site-survey techniques that base their radio resource management solely on WLAN reported measures. Other limitations come from simplifications such as to consider that a client association and the corresponding CSMA/CA mechanisms rely entirely on the strongest access point's RSSI, because most of the interferences do not occur at MAC level, and radio measurements may differ from a vendor to another. Then it is necessary to ensure that model analysis inputs and outputs do not result in a contradiction.

Motivated by advances in beamforming techniques, especially from an array signal processing perspective, which make it possible to radiate energy in any direction and to easily estimate signal's direction of arrival, this work solution model aims at fixing many of previous limitations by considering the opportunity for an access point to transmit in one specific direction. This opportunity processing accepts data from upper-layer services such as to achieve a complete end-to-end integration with other network parts.

As it will be demonstrated in this work, the number of transmit directions at access point level, could be optimized in conjunction with planning transmit power levels and channel assignment to maximize the overall network capacity. A cost-based approach prices in a timely manner each direction for an eventual transmission. Access points in this scheme discover each other over-the-air and report useful information on transmit directions, power levels and channels, to a central intelligence that coordinates the overall network operation. The ultimate result is:

1. Cancellation of co-channel interferences in a majority of deployment cases.

2. Processing of new transmission opportunities that are missing in conventional work.

3. Enhancement of end-to-end network performance.

In the upcoming section, we present a foundation on unified WIFI architectures and beamforming as they relate to this work. In Sect. 4, we discuss more formally the problem and in Sect. 5, we present our solution model. Sect. 6 discusses our solution results. In the end, we conclude and further our work. This paper is an extension of work originally presented in 2017 International Conference on Information Networking (ICOIN) [3].

2 Related Work

Co-channel interference in WLAN networks is a major problem that has been widely studied. To tackle this issue RRM and related techniques such as Transmit Power Control (TPC), Dynamic Channel Assignment (DCA), constitute a very accurate solution. If we consider TPC that is the focus of this work, the related-work's approaches differ from each other in terms of the adopted interference's model, inputs TPC may work on, and their processing (predictive, heuristic, etc.). We further focus on related-work's adopted interference models.

To our knowledge, all models and derived TPC processing have considered only the transmission power level as a degree of freedom to limit co-channel interference. This work claims that the transmission direction could be an additional degree of freedom to overcome other models' limitations when it comes to modeling irregular coverage areas in indoor WLAN environments.

Authors in this work [4] concentrate, from a lower layer perspective, on the co-channel interferences as a function of the estimated distance between the Access Points (AP) and the Wireless Devices (WD). In this scheme, the WDs estimate the distance based on the RSSI information sent by the APs in RTS control packets. The transmission power level adjusts accordingly to achieve a better network performance and interference cancellation. The issue with this model is that the power level adjustment affects other on reach WDs and APs communications. In addition, this model does not include cases where the WD does not simply receive the RTS packets.

In [5], a WD measures the amount of interference as a function of the local AP's load, its signal's

strength and loads from the distant APs. Here interference is a linear function of the AP's load. This is interesting but as opposed to our work, this model misses the interpretation of interference from an energy radiation perspective. In our work, we still include AP's load consideration in the SLA applied to the processing of transmit opportunities.

Depending on the coverage area pattern: disks or Voronoi zones, the amount of interference in [6], is a function of different AP's areas overlap in the first case and to their borders in the second one. This work model is very accurate to describe how an AP's energy radiation could affect other neighboring APs. However, it is still insufficient to interpret many cases specific to indoor WLAN deployments that relate to on-reach and unreachable AP's. Subsection C of Sect. 4 discusses these cases in detail.

The authors in [7] consider both MAC and PHY layer inputs: data payload length, path loss condition, and frame retry counts, to find out the suitable PHY mode and power level combination. This procedure requires a control packet exchange that is similar to conventional RTC/CTS for this adaptation. This model is easy to implement but it is insufficient to interpret many cases specific to indoor WLAN similarly to the previous work. In addition and similarly to the first cited work, any PHY mode and corresponding power level change apply to all neighboring APs and WDs indistinguishably.

In [8], the authors concentrate on upper layer application performance such as FTP and HTTP, to find out an optimal power scheme or more exactly, an optimal combination between transmit power and supported load for a given distance between nodes. This model considers SLA as in the second and fourth cited works and augments its processing to include "behavioral" protocol aspects. However, the corresponding model misses the lower-level interpretation of AP's radiated energy.

The authors in [9], approach the problem from an inter-protocol coordination point of view: WIFI, BT, etc. which may be very helpful to base an integrated-system wide transmit decision or "etiquette" on the utilization of shared radio resources. In this scheme, all wireless devices broadcast the information on spectrum usage. This information includes transmit power level and frequency channel of use. In addition and based on costing policies, the system allocates resources: frequency, power and time, to wireless devices. This work is a generalization of the first and fourth previously cited works and may incur the same limitations.

In this work, we review the conventional co-channel wireless coverage models and show that the system could process more transmission opportunities and still conform to the SLA. In these works [4] [5], some of the transmission opportunities were missed in areas that were considered falsely as interference ranges or worst as no-talk ones. Our solution model considers inputs from upper-layer services to conform to the SLA on the end-to-end wireless client's experience as in works [7] [8] [9]. In addition, it takes advantage of the technological advances in beamforming and related techniques, to introduce, at a conceptual level in indoor WLAN networks, the direction of transmission as a new radio resource added to frequency channel and power level, to enhance work's like [9] results.

3 Theoretical Background

The next two subsections describe WIFI's unified architecture in addition to the beamforming techniques as they relate to our study. The main objective of this foundation is to justify the feasibility of the hypotheses we base our study on. These hypotheses relate to the possibility to correlate reported radio interface's information from different APs, to coordinate their operation network-wide, to concentrate the radiated energy in any direction, in a timely manner, and to estimate accurately the direction of known and unknown sources' signals.

3.1 WIFI's Unified Architecture

In dense indoor networks, the transmission is difficult to model and depends on radio environment's characteristics such as obstacles, interference, background noise, density of wireless clients, mobility applications, etc.

In this context, standalone APs' architectures do not scale with high number of APs and WDs that require high class QoS treatment. Such architectures are replaced gradually by controller-based or "unified" ones. They are marked as unified for two reasons:

1. a central decision-making and intelligence, that is a Wireless LAN Controller (WLC), manages the overall network's operation.
2. the WLAN integrates with the overall network parts: LAN, WAN, DCN, etc. from especially a QoS perspective.

The market-leading vendors implement such a central decision-making processor mainly in three ways:

1. appliance WLC-based,
2. virtual system WLC-based,
3. distributed AP-based.

In the latter implementation, the APs take over the WLC's role. The first two implementations require a WLC, a virtual or physical appliance, which is reachable by all network APs over the wire. Examples of such WLCs are Cisco 8540 Wireless Controller and Aruba 7280 Mobility Controller.

Cisco unified architecture [10] defines two protocols:

1. CAPWAP, used by APs to build protocol associations to the active WLC.

2. NDP, to send messages over the air (OTA) for APs to exchange proprietary and standard management or control information.

In addition to these protocols, Cisco APs integrate a set of on-chip RRM techniques: CLIENTLINK and CLEANAIR that monitor and measure radio environment characteristics among other features. The WLC gathers this information via the already established CAPWAP tunnels to APs. Further, platforms such as Cisco Prime Infrastructure (CPI) and Mobility Services Engine (MSE) extend the capability of CLEANAIR to perform analytics, locate clients, interferers, and build heat maps. Based on the APs' reported information, the WLC computes the channel assignment and power levels map network wide. Its decision still conforms to the pre-configured policy sets. Each of these sets defines a number of configurable variables and unchangeable settings such as acceptable signal levels, tolerable noise levels, usable power levels, frequencies and channels, etc.

In this work context, the WIFI unified architecture guarantees a coordinated operation between all network elements at a very low level, an accessible distributed intelligence, and a centralized decision-making. It focuses on how to enhance the centralized decision-making at WLC level and especially RRM processing. RRM is the set of tools, algorithms and services that processes inputs from AP's radio and LAN interfaces and outputs a wireless area coverage plan or map that conforms to a predefined utility function.

3.2 Beamforming

The majority of WLAN communicating systems are equipped with omnidirectional antennas or a set of limited number of unidirectional antennas to obtain an equivalent energy radiation's pattern. Ease and cost of such hardware configurations are some of the benefits of such designs but they lack many aspects regarding interference and noise cancellation, signal sources' localization, and power consumption.

Beamforming techniques by definition relate to both the direction of energy radiation and the estimation of signal's arrival direction [11] and from an array signal processing perspective, they offer the possibility to tackle previous issues and design limitations to the next level in three ways:

1. the radiation of energy toward actual receivers allows better performance and less impact on the other system's elements.

2. the estimation of direction of known and unknown sources allows better costing of a given directed transmission.

3. the on-chip beamforming techniques allows for a hardware real-time adaptation to extremely changing radio environments like in our context.

With reference to [11], beamforming concerns both sides of a communicating system: the emitter and the receiver. It allows the radiated energy to be concentrated in one direction and to estimate its corresponding source's location for both legitimate and rogue interferers at receiver level. Basically, beamforming could be achieved mechanically by manually or electronically gearing a directional antenna. From a signal processing perspective, beamforming is achieved by changing certain transmitter's characteristics such as array elements' phase shifting, source signals' amplitude increasing or addition weighting. Additionally, the desired array gain and beam's direction depend on the number and geometry of array elements.

To ease our study we consider that array beamformers achieve equivalent performance to mechanical ones concerning both the resultant beam radiation's angle and signal's gain.

Let us consider an antenna with a beamwidth of $2 * \theta_n$ such as the radiated energy in the zone delimited by this angle is equal to the maximum energy radiated by the antenna minus $n\ dB$. n is calculated in such manner to allow an AP to transmit in all corresponding directions at the same time without interfering with each other. A rough estimation of this number is:

$$N \approx (2 * \pi)/(2 * \theta_n) \tag{1}$$

In the contexts of [12] [13] studies on antenna arrays, the beamwidth is expressed as a function of the distance between array elements d, the number of elements M, and the wave length λ:

$$\Delta\theta_{main} \approx (2 * \lambda)/(d * M) \tag{2}$$

In general, the evolving antenna array's beamformers could be classified as data independent, statistically optimum, adaptative or partially adaptative [11]. In this work, data dependent data beamformers (statistically optimum and adaptative) are of interest. Some of these beamformers are Multiple Sidelobe Canceller (MSC), Reference Signal (RS), Max SNR, Linearly Constrained Minimum Variance (LCMV), Lean Mean Square (LMS), Recursive Least Squares (RLS).

The other part of beamforming relates to the estimation of the Direction of Arrival (DOA). DOA estimation methods may include Bartlett, Capon, MUSIC, Min-Norm, DML, SML, WSF, Root-Music, ESPRIT, IQML, and Root-WSF. SML, WSF and Root-WSF are statistically the most efficient depending on the array geometry [14]. They could be yield into two main categories: spectral-based techniques, which are computationally optimal, and parametric methods, which are more accurate than the previous ones.

Recent algorithms are more accurate in the estimation of signal's arrival of both known and unknown sources. They depend on the array size, antenna's radiation pattern, source signals' correlation,

data statistics, corresponding data generation framework, among others. However, for the rest of this study we could consider that beamforming techniques allow:

1. sufficient number of directions to mimic an omnidirectional operation,
2. geographically any direction is achievable,
3. patterns from different directions at the same AP level do not interfere and are uncorrelated,
4. on-chip real-time transmits and synchronization.

4 Problem Statement

In this section, we formalize mathematically the problem of determining the opportunity to transmit at any point under the wireless coverage area. We first focus on related-work's models. Then in the next sections, we present our solution model and compare its results to preceding studies.

In this problem statement, we focus mainly on how related-work estimates the interference at a given point under the wireless coverage area, applies radio and upper-layer constraints, and processes the corresponding transmit opportunity.

To ease this work we consider the wireless coverage area as a two-dimensional Euclidean plane. The APs and WDs are points of this plane. They transmit in the same channel and cause the majority of co-channel interference in this area. This study focuses on co-channel operation but is easily extendible to tackle cross-channel interference or noise issues in the context described before.

Let us define:

$\{AP_1, AP_2, \ldots, AP_n\}$— set of n access points.

P_j — a point under wireless coverage area.

(x_i, y_i)— coordinates of AP_i or P_i points in an Euclidean two-dimensional plane.

$w_i, w_{min}, w_{max}, w_{opt}$— AP_i's corresponding current, minimum, maximum and optimum transmit power levels.

$d_{i,j}$ — distance between AP_i and P_j points.

$I(), O()$— interference and transmit opportunity functions.

We categorize related-work's models into two categories: Range-based and Zone-based. In the upcoming subsections, we describe each of them and discuss their limitations.

4.1 Range-based Model

In works similar to [4], an AP's wireless coverage corresponds to one of these ranges: a transmission, interference or no-talk range. These ranges correspond to the estimation of the distance between the AP and a receiving point P (AP or WD). Further, they consider that an AP's wireless coverage pattern is omnidirectional in the form of a circle or a disk. In this configuration, the three circles that correspond to each range type and centered at the AP define the whole AP's wireless coverage area as in Figure 1.

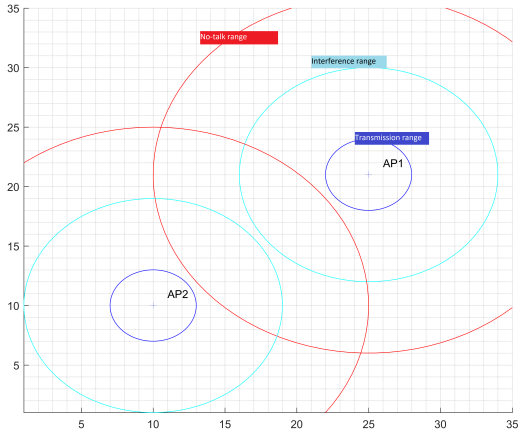


Figure 1: Range-based model's representation of an AP's wireless coverage

Let us further define:

$C_{i,tx}, C_{i,if}, C_{i,nt}$— AP_i attached circles that delimit corresponding transmission, interference and no-talk ranges.

$R_{i,tx}, R_{i,if}, R_{i,nt}$— AP_i corresponding ranges.

If we suppose that:

$$R_{i,tx} \leq R_{i,if} \leq R_{i,nt} \tag{3}$$

Then the interference and transmit opportunity are given as follows:

$$I(P_i) \approx \sum_{k=1}^{3} \alpha_k \sum_{j \neq i} \beta_j * intersection(C_{j,k}, C_{i,1}) \tag{4}$$

$$O(P_i) \approx \frac{1}{\sum_{k=1}^{3} \alpha_k \sum_{j \neq i} \beta_j * intersection(C_j, P_i)} \tag{5}$$

A weighted version of these functions is given further to account for inaccuracies and other constraints. The α and β weights are meant to reflect the non-linearity of interference addition, the power level's effect on interference strength, and the type of transmission's range. Here we consider that the interference at a given point P_i is a function of other AP_j's

signals except from the AP_i to which the point is associated. The opportunity processing is slightly different as it considers all AP's interference to P_i that is not associated to any AP_i yet.

4.2 Zone-based Model

Works such as [5] [6] that are based on this model are different from the previous ones as ranges here are not function only of the corresponding AP transmission's characteristics: channel, power level, etc., but depend also on the neighboring APs. The result of this is that the transmission range shape is no more a solid circle but a convex polygon with straight sides. Each straight side defines a borderline that separate two neighboring APs' transmission ranges. Consequently, it is important to note that a point in a transmission range of one AP could not be in another AP's transmission range. A Zone-based AP's wireless coverage is represented in Figure 2.

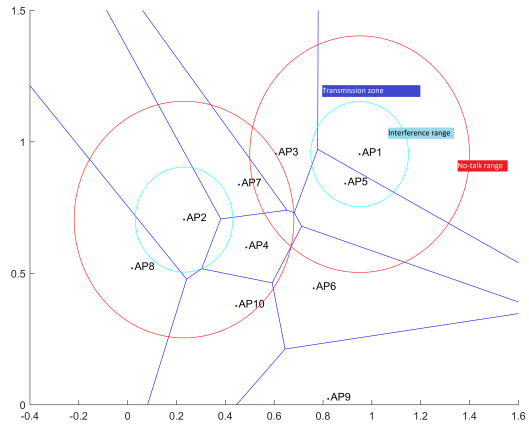


Figure 2: Zone-based model's representation of an AP's wireless coverage

Let us further define:

G_i — AP_i corresponding transmission range polygon or zone.

Then the interference and transmit opportunity are given as follows:

$$I(P_i) \approx \alpha_1 \sum_{j \neq i} \beta_j * intersection(C_{j,1}, C_{i,1})^{-\lambda} + \sum_{k=2}^{3} \alpha_k \sum_{j \neq i} \beta_j * intersection(C_{j,k}, C_{i,1}) \quad (6)$$

$$O(P_i) \approx \frac{1}{\sum_{k=2}^{3} \alpha_k \sum_{j \neq i} \beta_j * intersection(C_{j,k}, G_i)} \quad (7)$$

Similarly to formulas 4 and 5, α and β weights are meant to represent the impact of AP's transmission characteristics. λ is representative of the neighboring APs impact on AP_i borderlines. In this configuration, only transmission range is affected. The processing of other ranges are the same as in the previous model. It is also to note that opportunity in this model correspond to P_i in AP_i transmission's range or zone G_i rather than intersection of the neighboring APs like in the previous model.

4.3 First Observations

The presented models' functions of interference and opportunity do not reflect the accuracy of interference and opportunity measurements but rather the concept behind them and their possibilities to account for different WLAN design cases and aspects. At this stage, we see clearly that both models are limited in these ways:

1. both models are limited to consider that the strength of interference is only inversely proportional to the distance of an AP from interfering neighbors.
2. both models would interpret an increase in a transmission power level as an expanded reach in all directions: uniformly in case of Range-based models but depending on neighboring APs in case of Zone-based ones.
3. a point could not be in two transmission ranges of two different APs at the same time in Zone-based models.
4. transmission range does not depend on the strength of interference in Range-based models.
5. both models would interpret falsely obstacles to the signal propagation, as a weaker signal from an AP in the context of indoor WLAN networks does not mean necessarily that this AP is out of reach.
6. similarly to the previous point a stronger signal does not mean necessarily that an AP is in reach: it signal may be guided under some conditions.

The consequence of these limitations in indoor WLAN networks, is to reduce or overestimate the transmit opportunities that the network may offer. The upcoming section presents our solution to fix these modelization limitations.

5 Solution model

Our solution model is suitable for WLC-controlled indoor networks in the already explained conditions. For the purpose of this study, we suppose that APs are equivalent, support MIMO technologies such as beamforming and DOA, and operate over the same channel. The wireless devices (WD) are equipped with omnidirectional antennas and transmit at a lowered power level in comparison with APs.

The model runs three algorithms:

1. Transmit Direction Discovery (TDD) algorithm that builds a per direction neighbor discovery map, computes the optimal number of transmit directions and the corresponding per direction power levels.
2. Transmit Direction Mapping (TDM) algorithm estimates the impact that may have an AP on the network when it transmits in a specific direction and builds consequently, a per-direction transmit cost-based map that represents this information.
3. Transmit Direction Opportunity (TDO) algorithm processes data from both TDM and the upper-layer network services to optimize the transmission opportunities and build a per-direction transmit opportunity-based map to represent this information.

The next workflow in Figure **??** is an overview of our solution processing.

At initialization, the number of transmission directions is set to the minimum and the power level to the maximum. TDD processing stops after all APs are discovered and exactly one or no neighbor AP is discovered at any AP direction.

Based on results from TDD, TDM categorizes the coverage area points and costs a potential transmission at each of them. The coverage area points belong to one of these three categories: Not-On-Any-Transmit-Direction (NOATD), On-Discovered-Neighbor-Direction (ODND), and Not-ODND-On-Transmit-Direction (NOOTD). After the coverage area point's categorization, the system processes the transmission costs.

TDO algorithm evaluates the opportunity to transmit at a specific coverage area point. It returns the AP that may handle the transmission and the corresponding direction. This processing accepts inputs from TDM and the upper-layer SLA. The SLA processing is tight to the AP's wired network interface as opposed to the radio interface.

5.1 TDD algorithm

TDD algorithm processes the optimal number of transmission directions that APs may support and the corresponding per-direction transmit power levels. To ease this study, we consider these hypotheses and simplifications:

1. only a 2D plane operation is considered.
2. a uniform beamwidth in all directions.
3. APs are able to achieve an emulated omnidirectional operation.
4. two signals from the same AP in two different directions could not interfere.
5. all APs have the same optimal and maximal numbers of transmission directions and corresponding power levels.
6. neighbor discovery is bidirectional, occurs at the opposite direction, and at the same transmit power level.
7. an AP could not discover the same neighbor AP in two separate directions.
8. an AP could not discover itself.
9. at least one neighbor AP is reachable in any AP direction.

Points 2, 3, 4 and 5 are idealistic physical characteristics of AP's beamforming implementation. The 9^{th} point indicates that a co-channel interference condition is detected from a neighboring AP. The 6^{th}, 7^{th} and 8^{th} assumptions are further developed in separate work to cover and interpret cases where:

1. the optimal numbers of transmission directions and power levels are not the same among APs.
2. neighbor discovery is unidirectional or asymmetric.

Let us define:

(d_i, w_i)— the direction and power level associated with AP_i.

$d_{-j}(-x,-y)$— the opposite direction of $d_j(x,y)$ in an euclidean 2D plane.

$L(AP_i, d_j, w_k)$— the segment that represents the transmission range of AP_i in d_j direction at w_k transmit power level.

$a(i,j)$, $b(i,j)$, $c(i,j)$— respectively the segment $L()$ slope, intercept and end point.

$X(AP_i, d_j, w_k)$— the set of discovered AP neighbors by AP_i in d_j direction at w_k power level.

$O(AP_i, AP_j)$— a segment that represents an obstacle between AP_i and AP_j.

We formalize the problem as:

$$\forall i \in \mathbb{N}, \quad \exists j,k \in \mathbb{N} \quad \text{s. t.} \quad X(AP_i, d_j, w_k) \neq \emptyset \tag{8}$$

$$\forall i,j,k \in \mathbb{N}, \quad AP_i \notin X(AP_i, d_j, w_k) \tag{9}$$

$$\forall i,j,j',k,k' \in \mathbb{N}, \quad \text{and} \quad j \neq j', \quad X(AP_i, d_j, w_k) \cap X(AP_i, d'_j, w'_k) = \emptyset \tag{10}$$

$$\forall i,i',j,j',k,k' \in \mathbb{N}, \quad \text{and} \quad i \neq i', \quad AP_{i'} \in X(AP_i, d_j, w_k) \equiv AP_i \in X(AP_{i'}, d_{j'}, w_{k'}) \quad \text{and} \quad d_{j'} = d_{-j}, \quad w_k = w_{k'} \tag{11}$$

$$\forall i,j \in \mathbb{N}, \quad L(AP_i, d_j) = \left\{ a_{i,j} * x + b_{i,j} \quad | \quad x \in \left[x_i, x_i + c_{i,j} \right] \right\} \tag{12}$$

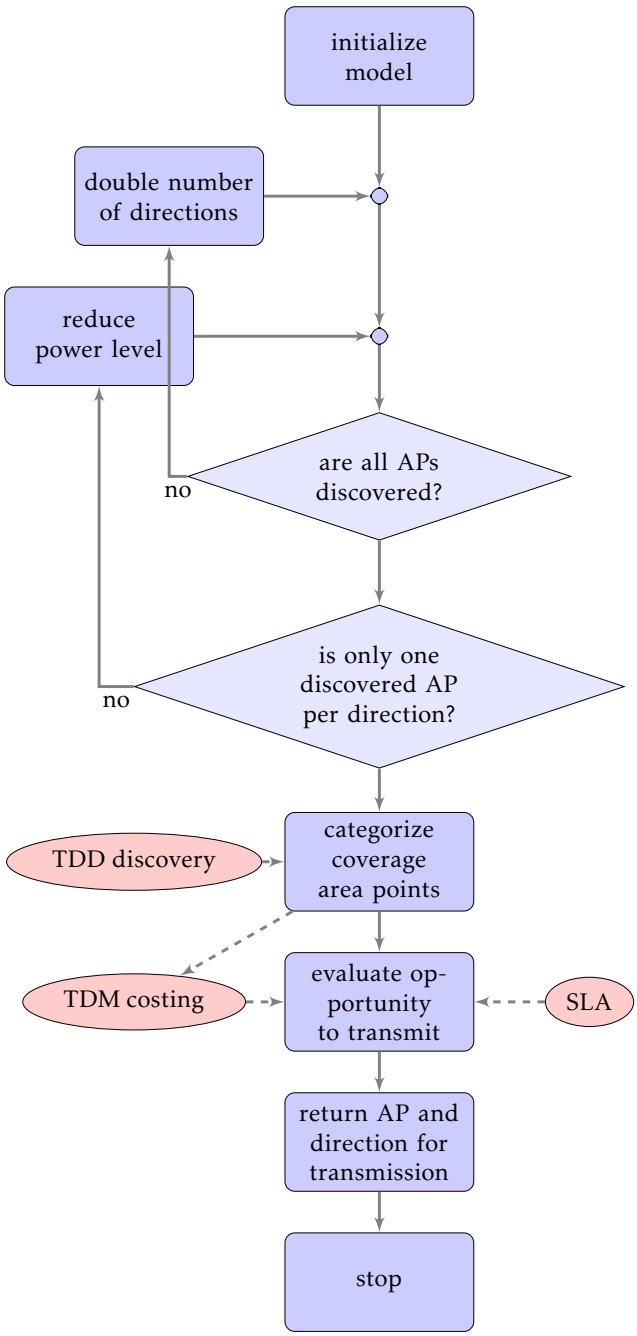


Figure 3: Flowchart of our solution model processing

The APs' distribution on the plan could be any random function or resulting from on-site surveys. However, we consider that the minimum and maximum distances between any of these APs are linear functions respectively of the corresponding minimum and maximum supported power levels. Figure 4 represents a simple network of seven APs with eight transmission directions each. Eight beams in red color are formed roughly around AP_i corresponding transmission directions.

The AP neighbor's per-direction discovery is based on the distance between APs that are on each other theoretical transmission range, on-site surveys, and radio interface sensing. Additionally the solution model allows the discovery of these two design special-case neighbors:

1. "isolated" neighbor AP that is at reach but separated by an obstacle.
2. "guided" neighbor AP that is not at reach but received with an acceptable signal level.

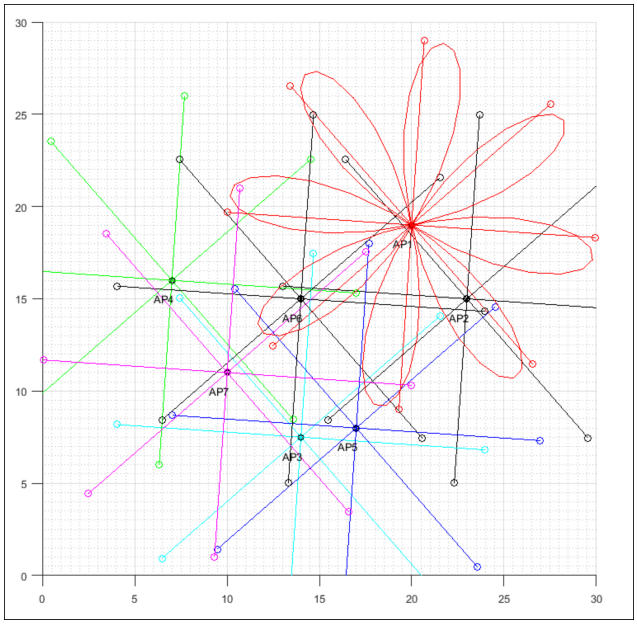


Figure 4: TDD Discovery Map example of 7 APs and 8 transmit directions

With reference to Figure 4 network's representation, the results of neighbor AP's discovery are given in Table. 1.

Table 1: Neighbor AP's discovery results

AP	d_1	d_2	d_3,	d_4	d_5	d_6	d_7	d_8
AP_1	AP_6	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	AP_2	$\emptyset$
AP_2	AP_5	AP_6	AP_1	$\emptyset$	$\emptyset$	$\emptyset$	AP_6	$\emptyset$
AP_3	$\emptyset$	$\emptyset$	AP_7	AP_6	$\emptyset$	AP_5	$\emptyset$	$\emptyset$
AP_4	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	AP_6	$\emptyset$	AP_7	$\emptyset$
AP_5	$\emptyset$	AP_3	$\emptyset$	$\emptyset$	AP_2	$\emptyset$	$\emptyset$	$\emptyset$
AP_6	AP_7	AP_4	$\emptyset$	$\emptyset$	AP_1	AP_2	$\emptyset$	AP_3
AP_7	$\emptyset$	$\emptyset$	AP_4	$\emptyset$	AP_6	$\emptyset$	AP_7	$\emptyset$

The optimization of APs transmission directions' number is a three-step process. For directions with more than one discovered neighbor, we calculate a new number such as to have the least possible number of neighbors per any AP's direction. Otherwise, we adjust the power levels to get some of these APs out of reach from the local AP. If no success, we revert to Range-based-like interference scheme to cost this transmission zone.

A simpler form of the proposed algorithm is given next:

1: $(d^{opt}, w^{opt}) \leftarrow (d^{min}, w^{max})$
2: $X \leftarrow \bigcup_{j=1}^{opt} X(AP_i, d_j^{max}, w_{opt})$
3: **for** $j = 1$ to opt **do**
4: **if** $opt \leq max$ **then**
5: **if card**$(X(AP_i, d_j^{opt}, w_{opt})) > 1$ **then**

```
        d_opt ← d_{2*opt}
      end if
    else if card(X(AP_i, d_j^opt, w_opt)) > 1 then
      repeat
        if ⋃_{j=1}^{opt} X(AP_i, d_j^opt, w_{opt-1}) = X then
          w_opt ← w_{opt-1}
        end if
      until ∀j,  card(X(AP_i, d_j^opt, w_opt)) ≤ 1
    end if
  end for
```

The optimum numbers of directions and power levels are respectively initialized to the minimum and the maximum numbers supported by AP_i. For every AP's direction, we check the number of discovered neighbors. If many neighbors are detected, we multiply the initial number of directions by two until we obtain only one neighbor per direction. If the maximum number of direction is reached then we reduce by one level the power and try to get the same result without altering X. X is the set of all neighboring APs that were detected previously in the maximum supported directions and at the maximum supported power level.

5.2 TDM algorithm

In the previous section, TDD returned the optimal number of transmission directions, the corresponding optimal number of power levels and the sets of AP's per-direction discovered neighbors. In addition to these elements, TDD interprets the presence of two kinds of obstacles: those that are preventing isolated neighbors from seeing each other, and those that are creating new transmission opportunities among guided neighbors. One way for TDD to interpret such kind of obstacles is by comparing the on-the-air discovery's results with on-site surveys' information. Based on the previous elements, TDM algorithm builds a transmission-costing map for APs to reach any point under the overall wireless coverage area.

A simpler form of the proposed algorithm is given next:

```
for j = 1 to opt do
  if P_i ∈ L(AP_i, d_j) and X(AP_i, d_j) ≠ ∅ then
    Case_1
  else if P_i ∈ L(AP_i, d_j) and X(AP_i, d_j) = ∅ then
    Case_2
  else
    Case_3
  end if
end for
```

TDM costing system categorizes WDs as NOATD, ODND, or NOOTD. These WDs could be represented roughly by circles with sufficiently small radii in comparison with APs transmit ranges, to not false processing results. We process carefully NOOTD points, which are not on discovered-neighbor APs' directions, as they are hard to detect.

Let us define:

$C_i^W D$— WD attached circle.

ρ_i — weights associated with each category.

$Cost(P_i)$— function of transmissions cost at P_i

Then for this preliminary work, the cost of a transmission could be expressed as follows:

$$Cost(P_i) \approx \begin{cases} \rho_1 I(P_i), & \text{if } NOATD \\ \rho_2 I(P_i), & \text{if } ODND \\ \rho_3 I(P_i), & \text{if } NOOTD \end{cases} \quad (13)$$

Where,

$$I(P_i) \approx \alpha_1 \sum_j \sum_k \beta_{j,k} * intersection(L_{j,k}, C_i^{WD}) + \sum_{k=2}^{3} \alpha_k \sum_{j \neq i} \beta_j * intersection(C_{j,k}, C_{i,1}) \quad (14)$$

Cost calculation results were plotted on Figure 5, for the same Figure 4 network example set of seven APs and eight transmission directions. On this map, transmission costs range from white color, lowest cost points, to red color that correspond to the highest costly points.

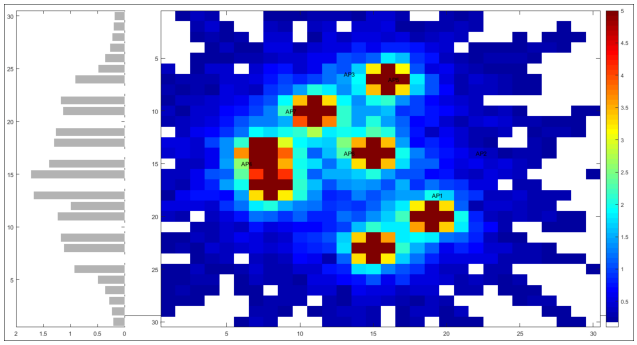


Figure 5: TDM Cost Map example of 7 APs and 8 transmit directions

5.3 TDO algorithm

TDO takes into consideration the worst-case analysis done by TDM when all the APs transmit at the same time in all supported directions at the maximum power level. Then for a given transmission, it prioritizes the directions with the least cost in duality with TDM but in addition, it integrates data from the upper service layers and returned SLA tests both on the passive and active communication's paths.

A more complex form of this opportunity processing relies on a system-wide APs transmit operation's synchronization to cancel costly directions and on-the-air tests to optimize NOOTD costs.

Let us define:

$s_{1,i}, s_{2,i}$— AP_i weights associated with defined SLA passive and active test results.

$O(P_i)$— opportunity of transmission function at P_i.

Then for this preliminary work, the transmission opportunity calculation at any point P_i is given by:

$$O(P_i) = s_{1,i} s_{2,i} \overline{I(P_i)} \tag{15}$$

The transmission opportunity processing results were plotted on Figure 6 for a set of seven APs and eight transmit directions.

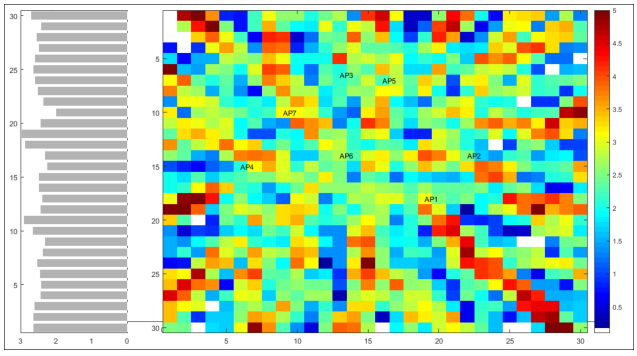


Figure 6: TDO Opportunity Map example of 7 APs and 8 transmit directions

$s_{1,i}$ and $s_{2,i}$ weights are calculated based on the simulated SLA active test results that may include: Packet Loss Rate, Round-Trip Time (RTT), Jitter, and on SLA passive measurements at PHY and MAC levels like: error rates, transmitted packets, etc. that are directly taken from the radio and LAN interfaces. In addition, these weights take into consideration the nature of the end-user application: real-time applications do not require the same service from the network as bulk ones.

5.4 Layered Transmit Opportunity Processing

With regard to this work opportunity processing, a transmission is a four-step procedure: discover, cost, evaluate, and transmit. Discovery is primarily done by APs and could be extended to any WD. The interpretation of discovery results, costing, and opportunity processing, are done at WLC level. Costing in this work has been solely based on interference estimation but could integrate any other radio interface's measurements: Signal to Noise Ratio (SNR), Received Channel Power Indicator (RCPI), and so on. Opportunity evaluation is based both on previous radio interface costing and on measures from upper layer services. These two measures have to be fully uncorrelated to not false results.

5.5 Transmit Synchronization

In this configuration, a transmission is not tight to the access point WD is associated with, but rather to the AP that offers the best transmit opportunity among a set of APs. In fact, the WD is associated with a virtual or pseudo AP that is representative of a set of physical APs with common characteristics with regard to this association. Any physical AP attached to the pseudo-AP could source an effective transmission to the associated WD. In the context of this work, we match the WD's corresponding pseudo-AP to a single physical AP.

6 Evaluation and Conclusion

At this stage, the evaluation is based on Matlab simulations of both our model and models from related-work [4] [5] [8] Range-based, and [6] Zone-based solutions. In the next step, we consider implementing this solution on a Linux-based APs and WDs and compare its performance with simulated related-work and vendor's implementations.

We compare our model performance to both Range-based and Zone-based models in these three transmission conditions:

1. a free space transmission condition,
2. the presence of obstacles, "isolated" neighbor's condition,
3. "guided" neighbor's condition.

We first evaluate the compared transmission opportunity for any set of 30 APs and ten times a randomly selected 100 WDs on a 2D grid. To ease this work we limit the maximum number of APs to 30 and the number of WDs test points to 100.

Figure 7 shows that for 30 APs and 10 random sets of 100 WDs on the grid, the average performance calculations is almost: 77% for our solution, 68% for Zone-based and only 57% for Range-based models. Figure 8 shows the average transmit opportunity for each model per WDs sets.

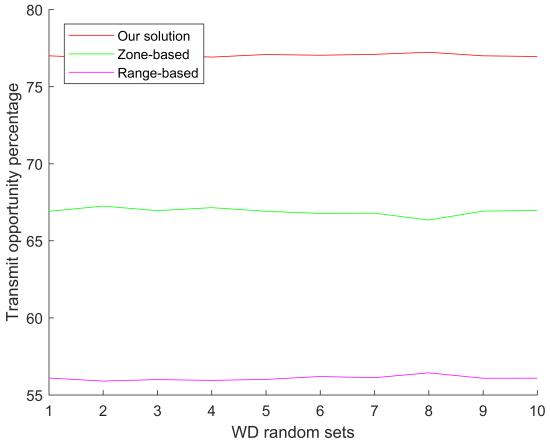


Figure 7: Performance comparison for 30 AP and 10 random sets of 100 WDs

Second, as in Figure 9, we evaluate the APs number impact on the compared performance for any set of 100 WDs. APs number ranges from 1 to 30. In Figure 10, we compare Zone-based to Range-based models.

We notice a general tendency of all models to perform similarly when APs are of limited number. However, for greater APs' numbers, the Zone-based models perform better than Range-based ones as in Figure 10. In Figure 9, our model returns a better transmit opportunity estimation when APs' number increases in comparison to other models.

Third, in Figure 11, we vary the number of the supported directions that is applicable only to our model and observe the compared average performance. For this test purpose, this number is tight to a variable α that takes these values: 12.5%, 25%, 50% and 75%. We only compare our model to Zone-based one.

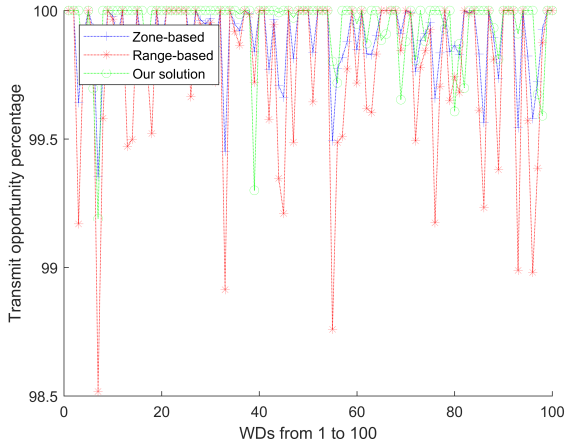


Figure 8: Average performance comparison per WDs in a network of 30 APs and 100 WDs

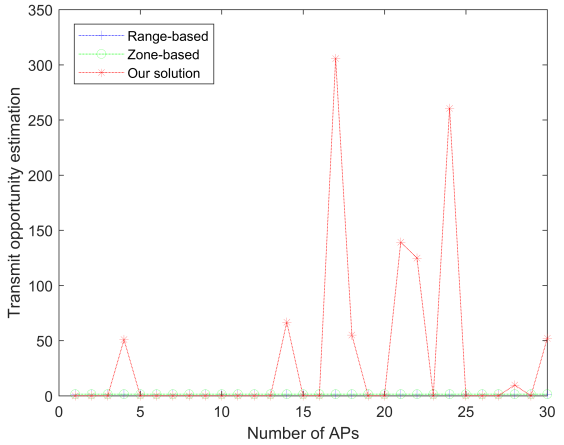


Figure 9: Comparison of our solution, Zone and Range-based models for any set of 100 WDs by APs number

We see clearly that our solution is comparable to Zone-based one when α is almost 75%. Our solution is near optimal when α is 12,5% but this is idealistic. However, achieving a 50% α is realistic with regards to these-days vendors APs system characteristics. A 50% α coefficient may correspond roughly to four transmission directions or beams.

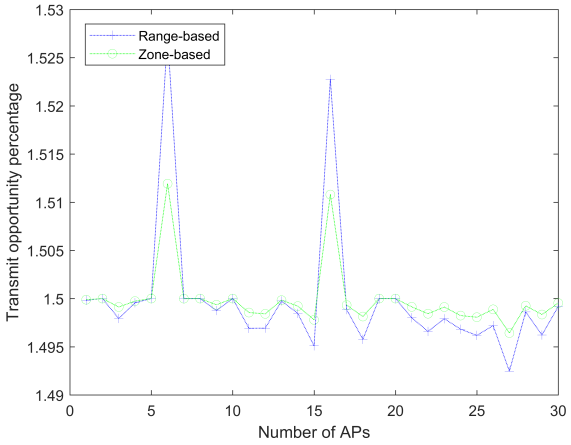


Figure 10: Comparison of Zone and Range-based models for any set of 100 WDs by APs number

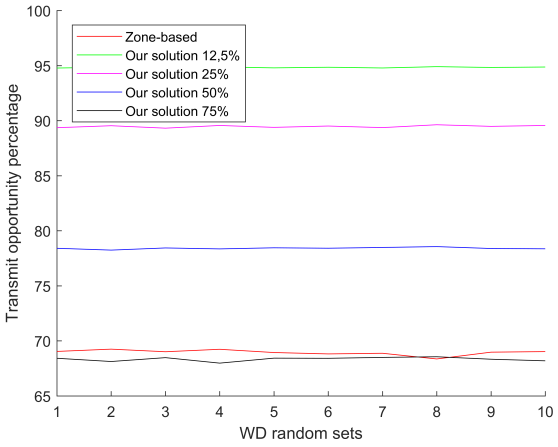


Figure 11: AP directions number impact on performance

As a conclusion, this work transmission processing is built on a new costing and opportunity evaluation model that leverages advances in antenna array processing such as beamforming techniques and WLAN integrated design best practices. In addition to the transmission channels and corresponding transmission power levels, it considers the direction of transmission or beam as an additional degree of freedom of dynamic RRM purposes. A comparison to other models indicates that our model is stable and performs better in very dense WLAN indoor and WLC-based network deployments.

Further work may focus on:

1. more complex neighbor AP discovery scenarios,
2. angle of incidence in TDM costing,
3. APs transmit operation coordination,
4. more complex costing layered approach for end-to-end to communications.

Conflict of Interest The authors declare no conflict of interest.

Acknowledgment We would thank colleagues : researchers, engineers, and anonymous reviewers for sharing their precious comments and on-field experience that improved the quality of this paper.

References

[1] "Radio Resource Management White Paper." [Online]. Available: https://www.cisco.com/c/en/us/td/docs/wireless/controller/technotes/8-3/b_RRM_White_Paper.html

[2] "ARM Coverage and Interference Metrics." [Online]. Available: http://www.arubanetworks.com/techdocs/ArubaOS_64x_WebHelp/Content/ArubaFrameStyles/ARM/ARM_Metrics.htm

[3] M. Guessous and L. Zenkouar, "Cognitive directional cost-based transmit power control in ieee 802.11 wlan," in *Information Networking (ICOIN), 2017 International Conference on*. IEEE, 2017, pp. 281–287.

[4] R. Ruslan and T.-C. Wan, "Cognitive radio-based power adjustment for wi-fi," in *TENCON 2009-2009 IEEE Region 10 Conference*. IEEE, 2009, pp. 1–5.

[5] N. Ahmed and S. Keshav, "A successive refinement approach to wireless infrastructure network deployment," in *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, vol. 1. IEEE, 2006, pp. 511–519.

[6] R. Prateek and P. Om, "Interference-constrained wireless coverage in a protocol model," in *Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems Terromolinos, Spain: ACM*, 2006.

[7] A. Akella, G. Judd, S. Seshan, and P. Steenkiste, "Self-management in chaotic wireless deployments," *Wireless Networks*, vol. 13, no. 6, pp. 737–755, 2007.

[8] D. Qiao, S. Choi, A. Jain, and K. G. Shin, "Adaptive transmit power control in ieee 802.11 a wireless lans," in *Vehicular Technology Conference, 2003. VTC 2003-Spring. The 57th IEEE Semiannual*, vol. 1. IEEE, 2003, pp. 433–437.

[9] D. Raychaudhuri and X. Jing, "A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands," in *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*, vol. 1. IEEE, 2003, pp. 172–176.

[10] "Enterprise Mobility 8.1 Design Guide." [Online]. Available: https://www.cisco.com/c/en/us/td/docs/wireless/controller/8-1/Enterprise-Mobility-8-1-Design-Guide/Enterprise_Mobility_8-1_Deployment_Guide.html

[11] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[12] V. Rabinovich and N. Alexandrov, "Typical array geometries and basic beam steering methods," in *Antenna Arrays and Automotive Applications*. Springer, 2013, pp. 23–54.

[13] T. C. Cheston and J. Frank, "Phased array radar antennas," *Radar Handbook*, pp. 7–1, 1990.

[14] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 40-48 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

A New profiling and pipelining approach for HEVC Decoder on ZedBoard Platform

Habib Smei[*1, 2], Kamel Smiri [1, 3], Abderrazak Jemai [1, 4]

[1]*Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire LIP2, 2092, Tunis, Tunisie*

[2]*Direction Générale des Etudes Technologiques, Institut Supérieur des Etudes Technologiques de Rades, Rades, Tunisie*

[3]*Université de Manouba, Institut Supérieur des Arts Multimédias Manouba, Campus Universitaire Manouba, 2010, Tunisie*

[4]*Université de Carthage, INSAT, B.P. 676, 1080 Tunis, Cedex, Tunisie*

ARTICLE INFO

Article history:
Received: 01 August, 2017
Accepted: 22 September, 2017
Online: 02 October, 2017

Keywords:
Co-Design flow
Profiling
Performance estimation
Embedded Systems
Pipeline
HEVC decoder
Zynq
Zedboard

ABSTRACT

New multimedia applications such as mobile video, high-quality Internet video or digital television requires high-performance encoding of video signals to meet technical constraints such as runtime, bandwidth or latency. Video coding standard h.265 HEVC (High Efficiency Video Coding) was developed by JCT-VC to replace the MPEG-2, MPEG-4 and h.264 codecs and to respond to these new functional constraints. Currently, there are several implementations of this standard. Some implementations are based on software acceleration techniques; Others, on techniques of purely hardware acceleration and some others combine the two techniques. In software implementations, several techniques are used in order to decrease the video coding and decoding time. We quote data parallelism, tasks parallelism and combined solutions. In the other hand, In order to fulfill the computational demands of the new standard, HEVC includes several coding tools that allow dividing each picture into several partitions that can be processed in parallel, without degrading neither the quality nor the bitrate.

In this paper, we adapt one of these approaches, the Tile coding tool to propose a pipeline execution approach of the HEVC / h265 decoder application in its version HM Test model. This approach is based on a fine profiling by using code injection techniques supported by standard profiling tools such as Gprof and Valgrind. Profiling allowed us to divide functions into four groups according to three criteria: the first criterion is based on the minimization of communication between the different functions groups in order to have minimal intergroup communication and maximum intragroup communication. The second criterion is the load balancing between processors. The third criterion is the parallelism between functions. Experiments carried out in this paper are based on the Zedboard platform, which integrates a chip Zynq xilinx with a dual core ARM A9. We start with a purely sequential version to reach a version that use the pipeline techniques applied to the functional blocks that can run in parallel on the two processors of the experimental Platform. Results show that a gain of 30% is achieved compared to the sequential implementation.

1. Introduction

This paper is an extension of the work originally presented in [1].

In recent years, the number of applications processing digital video is steadily growing. Streaming videos, videoconferencing, web cameras, mobile video conversations are examples of digital video that require good video quality.

In addition, statistics show that by 2020 [2], Internet video streaming and downloads will increase to over 80% of all consumer Internet traffic. This growing demand for digital video processing encourages digital video coding market operators to design and develop new solutions that can meet this growing need.

[*]Corresponding Author: Habib Smei, Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire LIP2, 2092, Tunis, Tunisie
Email: habibsmei@gmail.com

Indeed, we are witnessing an increase in companies and research groups working to develop standards of video codec with better quality.

The JCT-VC (Joint Collaborative Team on Video Coding) group, formed by VCEG and MPEG, was created at the end of 2009 to create a new standard that meets the new requirements of video processing [3]. Following the work of this group, a new standard called H.265 / HEVC (High Efficiency Video Coding) was created with the aim of compressing a half-rate video from the previous H.264 / AVC standard to one Quality. This improvement in terms of efficiency is necessary to manage beyond high definition resolutions like HD, Quad-HD and Ultra-HD.

Many research groups and companies participate in standardization meetings and contribute to the growth of the standard with new algorithms. As a result, many implementations of this standard have been developed. One of these implementations is the test model software [4], which is a complete encoder and a decoder for the new algorithms.

To design video processing devices that implement standard video encoders, designers investigate several solutions. All solutions are based on improvement techniques based on software or hardware optimization approaches or a combination of both approaches.

The work undertaken in this paper is part of the LIP2 co-design flow.

In a co-design process, designers typically begin to run embedded application code on a host machine. Then, a thin profiling step is done to describe each function. This description can affect the execution time, memory size, number of calls, relationship between functions and other parameters that may be useful for the designer to make the best design decisions.

In this paper, we present a new approach to implement a pipeline solution of the HEVC H.265 application decoder based on the Zedboard platform [5]. The results show that a gain of 30% is achieved compared to the sequential implementation.

The rest of this article is organized as follows: Section 2 presents an overview of the HEVC Standard and HEVC h.265 decoders. In section 3, we present some works in relation with HEVC parallelization.

Section 4 gives a detailed description and profiling of the version of the test model of the decoder and analyses the results obtained. Section 4 shows the execution of the decoder pipeline on the Zedboard platform. Finally, Some conclusion remarks and future directions are given in Section 6.

2. Overview of the HEVC Standard

Due to the complexity of multiprocessor systems, the probability of failure is all the more important that it requires special consideration. Indeed, during a failure, a part of the application state disappears and the application may pass in an inconsistent state that prevents it from continuing normal execution.

The HEVC [5,6] is the acronym of "High Efficiency Video Coding". This is the latest video-coding format used by JCT-VC. This standard is an improvement of the H.264 / AVC standard. It was created on January 2013 [7], when a first version was finalized. It was developed jointly by the ISO / IEC Moving Picture Experts Group (MPEG) & ITU-T Video Coding Experts Group (VCEG).

The main objective of this standard is to significantly improve video compression compared to its predecessor MPEG-4 AVC / H.264 by reducing bitrate requirements by as much as 50% compared to H.264/ AVC, with equivalent quality.

The HEVC supports all common image definitions. It also provides support for higher frame rates, up to 100, 120 or 150 frames per second.

Video coding standards have evolved mainly through the development of the well known ITU-T and ISO / IEC standards. ITU-T produced H.261 [8] and H.263 [9], ISO / IEC produced MPEG-1 [10] and MPEG-4 Visual [11], and the two organizations together produce H. 262 / MPEG -2 Video [12] and H.264 / MPEG-4 Advanced Video Coding (AVC) [13,14] standards.

The two standards that have been produced jointly have had a particularly strong impact and have found their way into a wide variety of products that become increasingly common these days. In the course of this evolution, efforts have been multiplied to increase the compression capacity and to improve other characteristics such as the robustness against data loss. These efforts take into account the capability of practical IT resources to be used in products at the time of the expected deployment of each standard.

The HEVC standard is intended to complement various objectives and to meet even stronger needs to encode videos with an efficiency more important than H.264 / MPEG -4. Indeed, there is a growing diversity of services such as ultra high definition television (UHDTV) and video with a higher dynamic range.

On the other hand, the traffic generated by video applications targeting peripherals and mobile tablets and transmission requirements for video-on-demand services impose serious challenges in current networks. An increased desire for quality and superior resolutions also occurs in mobile applications.

The HEVC standard defines the process of encoding and decoding the video. As an input, the encoder will process an uncompressed video. It performs the prediction, transformation, quantification and entropy coding processes to produce a bitstream conforming to the H.265 standard.

The decoding process is divided into four stages. The first stage is the entropy decoding for which relevant data such as reference frame indices; intra-prediction mode and coding mode are extracted. These data will be used in the following stages. The second is called reconstruction step, which contains the inverse quantization (IQ), inverse transform (IT) and a prediction process, which can be either intra-prediction or motion compensation (inter-prediction). In the third stage, a de-blocking filter DF is applied to the reconstructed frame. Finally a new filter called Sample Adaptive Offset (SAO) is applied in the fourth stage. This filter adds offset values, which are obtained by indexing a lookup table to certain sample values [14].

In HEVC, the coding structure is based on a quaternary tree representation allowing partitioning into multiple block sizes that

easily adapt to the content of the frames. Three units are defined to treat each frame. First, the Coding Unit (CU) that defined the frame‘s based size on pixels. It allows sizes ranging from 8x8 to 64x64 pixels to be adapted according to the application. Second, the Prediction Unit (PU) that defines the partitioning size according to predicts type (inter and intra). Finally, the Transform Unit (TU) defines the quantification and transforms size to be applied to a prediction unit. Four levels of decomposition are possible for this TU, which take sizes ranging from 4x4 to 32x32 pixels.

Three types of coding structures are defined; ALL intra or intra-only (AI), Low Delay (LD) and Random Access (RA). In the first structure, all pictures are encoded as intra, which yields very high quality, and no delay; this mode is mainly aimed for studio usage.

In the low delay one, the first is an intra frame while the others are encoded as generalized P or B pictures (GPB). This structure is conceived for interactive real-time communication such as video conferencing or real-time uses where no delay for waiting for future frames is permitted.

Finally, the random access structure is similar to hierarchical structure and intra pictures are inserted periodically, at the rate of about one per second. It is designed to enable relatively frequent random access points in the coded video data. This coding order has an impact on latency, since it requires frame reordering: for this reason the decoder might have to wait to have decoded several frames before sending them to output. This is the most efficient mode for compression but also requires the most computational power.

Each of these three modes has a low complexity variant where some of the tools are disabled or switched into a faster version. As an example low complexity uses CAVLC instead of CABAC

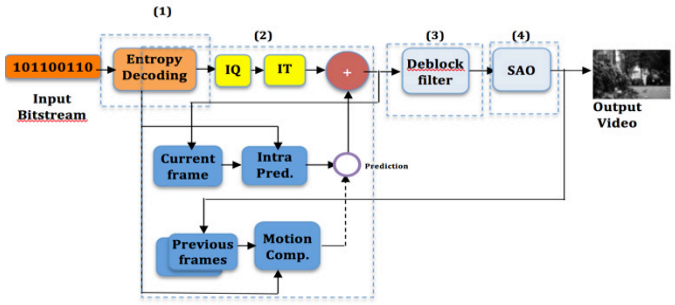


Figure 1. Functional structure of the HEVC decoder

3. Related works

HEVC includes several parallel specifications for many distributed multi-core systems. This allows division of each picture into several partitions that can be processed in parallel.

For these specifications, coarse grain communications levels are used. We cite the Group Of Pictures (GOP) level, the Slice Level, the Tile level and the wavefront parallel processing (WPP) level.

All these communication granularity has been justified by the small data dependency between processes acting on separate partition blocks.

For an embedded multiprocessor System on Chip (SoC) implementation, the communication granularity specified is not appropriate in all cases given the limited on-chip resources (memory, CPU frequency, bandwidth, …). For this, more fine grain communications granularities (CTU, TU, PU) are specified in HEVC standard based on a fine partition of a frame.

In [15], Authors carried out the analysis of parallel processing tools to understand the effectiveness of achieving the purpose for which they are targeted. In [16], an optimized method for MV-HEVC is proposed. It uses multi-threading and SIMD instructions implementation on ARM processors. The proposed method of MV-HEVC showed improvement in terms of processing speed on advanced RISC multi-core processors mobile platforms (ARM). In [17], Authors used the Wavefront Parallel Processing (WPP) coding and implemented it on multi- and many-core processors. The implemented approach is named Overlapped Wavefront (OWF), an extension of WPP tool that allows processing multiple picture partitions as well as multiple pictures in parallel with minimal compression losses. Results of her experimentations show that exploiting more parallelism by increasing the number of cores can improve the energy efficiency measured in terms of Joules per frame substantially. [18] discussed various methods in which the throughput of the video codec has been improved including at the low level in the CABAC entropy coder, at the high level with tiles, and at the encoder with parallel merge/skip tool. In [19], many optimizations are proposed to achieve HEVC real-time decoding on a mobile processor. These code optimizations include the adoption of single-instruction multiple-data (SIMD). An important speedup is accomplished with multiple threads assigned each one to a picture to be decoded. However, in the proposed solution, the SAO filter has not been implemented and optimized. Further, for the AI configuration, no additional speedup is achieved since no CPU resources are available anymore to decode a picture. In [20], a hybrid parallel decoding strategy for HEVC is presented. It combines task and data parallelism without any constraint or coding tools. The proposed approach aims to balance execution time of different stages, especially with SSE (Streaming SIMD Extensions) optimization. Another parallelism approach based on entropy slices is presented in [21]. This approach is not based on many slices because it can reduce the coding efficiency. It assigns one thread per LCU block to parallelize the HEVC decoder. Moreover, these threads are synchronized using the Ring-Line Strategy to maintain the wave front dependencies. This solution is great for high resolutions; however, it’s not the same for others.

4. Profiling of Test Model HEVC Decoder

In the specification and modelling phase of a co-design flow, the complete embedded system is written in a high level language such as C, C++, Java, Matlab and then the software is functionally verified. It will then be simulated on a host machine in order to understand its behaviour and measure the runtime performance of the program and eventually return feedback and performance statistics to the designer

This specification is an input to the profiling step, which consists in understanding the behaviour of the system in the various execution cases in order to deduce a clear representation of the various functions that represent it.

The behaviour of the system can be understood and mastered by following a parametric analysis of its execution in the different configurations and situations that the system is subject to.

This analysis can only be undertaken after having completed a detailed description of different functions of the system. The description can include the execution time of each function, its sizes, call graphs, dependencies between them and other information that can help designers to make the best decisions and technical choices. This process of identification and description is called profiling.

Once the profiling is done, designers can test the system with several configurations and estimate its performance parameters.

Several works dealing with performance estimation of embedded systems have been realized. In the case of the HEVC codec, performance estimation aims to test the ability of codecs to process video for real-time applications and to support new resolutions (HD, UHD, Full HD, 4k, 8k) and to measure footprint and the space occupied by the hardware components that make up the device.

In the work carried out by [22], a complete profiling of the encoder was carried out. The aim of this work is to present the different functions of the encoder, their execution times and the types of operations carried out in order to deduce the functions that are candidates for a hardware migration. The results are presented in terms of types of assembly level instructions in each encoder function. In [23], the authors propose a hybrid parallel decoding strategy for HEVC, which combines task level parallelism and data level parallelism. In [24], the authors use a performance estimation analysis to prove a power model based on bit derivations that estimates the energy required to decode a given HEVC coded bitstream. In [25], authors proposed a method to improve H.265/HEVC encoding performance for 8K UHDTV moving pictures by detecting amount or complexity of object motions.

4.1. Functions of the Test Model

HEVC HM Test Model is an open source project under BSD license. It is intended for the implementation of an efficient C++ HEVC decoder. The version used in this work was released from [4] and it is compliant with most of the HEVC standard. The code of the application is composed of C++ classes, and the main classes and functions are listed in the table 1.

The graph below (Figure 2) shows the most important classes in the decoding process and their associated functions. We note that the most significant classes are TComDataCU, TComTrQuant and TComSlice. The TComDataCU class represents the declaration of the CU data structure. The TComTrQuant class includes inverse transformation and inverse quantization, and the TComSlice class includes the decoder decompression process.

The optimization of the codes of the functions making up these classes can have a positive impact on the execution time and the footprint of the decoder.

As DecodeCTU, DecompressCtu called recursive function xDecompressCU which, in turn, decompresses each CU with the adequate prediction mode. When all CTUs in a frame are processed, the loop ends and the decoder performs both DF and SAO filters to correct artefacts.

Table 1: Main classes and principal functions of test model decoder

Main Classes	Principal functions	Size of code (Ko)
TDecGop	FilterPicture decompressSlice	8
TAppDecCfg	parseCfg	9
TAppDecTop	Decode x WriteOutput	21
TDecEntropy	decodePredInfo decodePUWise decodeInterDirPU	28
TdecTop	Decode Executeloopfilter	32
TComLoopFilter	xDeblockCU LoopFilterPic	35
TDecCU	decompressCtu decodeCtu xdecodeCU xcompressCU xReconInter xReconIntraQT	44
TComSlice	decompressSlice	126
TComTrQuant	partialBufferflyInverse32 partialBufferflyInverse16 partialBufferflyInverse8 xDeQuant	135
TComDataCU	initCu copySubCu	140

As shown in the diagram in Fig. 3, the process of decoding a frame is done by calling the DecompressCTU function of the TDecSlice class. This function executes a read loop of all CTUs in the current frame. Each CTU is decoded through the DecodeCtu function of the TDecCu class. The XdecodeCu function of the same class, recursively performs the decoding of all CUs in the same CTU. When decoding a CTU, the TDecSlice makes a second call to the DecompressCTU function of the TDecCu class. Like DecodeCTU, DecompressCtu call the recursive function xDecompressCU, which, in turn, decompresses each CU with the proper prediction mode. When all CTUs in a frame are processed, the loop ends and the decoder perform DF and SAO filters to correct the artefacts.

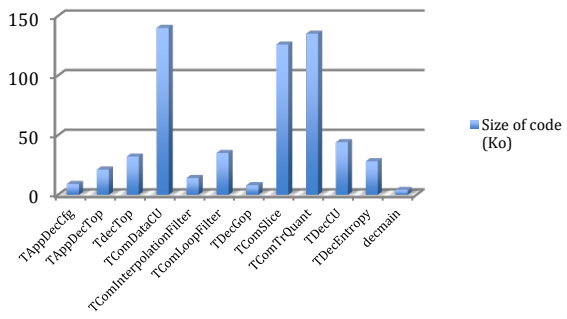


Figure 2. Main functions of Test Model Decoder

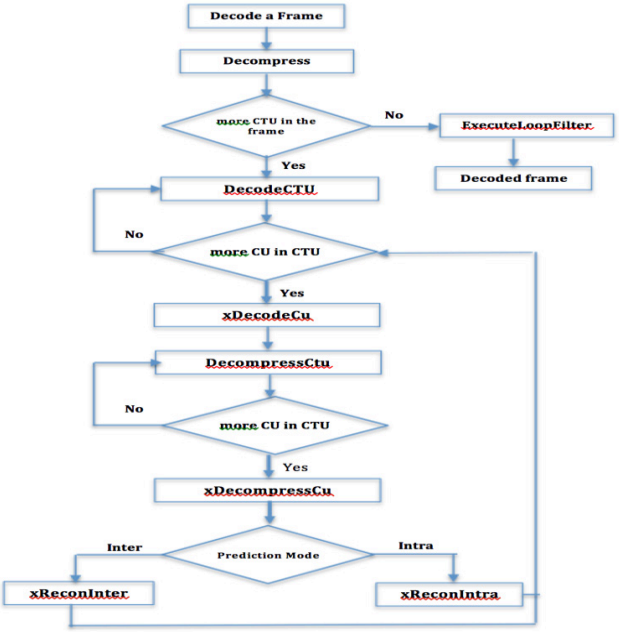


Figure 3. Frame decode Diagram of HM version

4.2. Call Graph functions of HM HEVC Decoder

Our goal in this step is to check the behaviour of the decoder described above, and to evaluate the performance of its various components in detail.

To achieve the profiling operation, we made use of profiling techniques based on code injection and also the use of tools such as gprof [26, 27] and Valgrind [28]. Gprof provides a flat profiling to give the execution time of each function of the decoder and also the number of times this function is called. Gprof can give us as a result, the call graph. This indicates for each function code, the number of times it was called, either by other functions or by itself. This information shows the relationships between the different functions and can be used to optimize certain code paths. Roughly, the results of profiling the performance generated by GProf revealed a runtime complexity of balance in terms of computing time in the functions Motion Compensation (36%) and ExecuteLoofilters of (16.7%).

Valgrind is a GPL licensed programming tool that can be used for profiling under the Linux operating system. It includes several tools, one of which is Callgrind. We used Callgrind to obtain the exact number of operations that were performed while decoding an entire video sequence. The results are separated in terms of classes, and in every class in terms of functions.

The resulting graph shows for each function, the execution time as a percentage of the total execution time (Valingring) and the functions to which they appealed. Normally, the graph shows that the most time-consuming functions are Motion Compensation and ExecuteLoofilters.

The aim of the profiling step is to understand the behaviour of the decoder described above and to evaluate the performance of its various components in detail.

To do so, we used profiling tools such as Gprof and Valgrind. These tools allow us to present the call graphs of the different functions and their execution times, and thus the state of the stack used. In order to have a complete profiling which includes the execution sequences in the different execution scenarios and configurations, a manual injection of code has been realized into significant functions.

The Gprof tool can give as a result, the call graph. This indicates for each function, the number of times it was called. This information shows the relationships between the different functions and can be used to optimize some code paths.

The profiling results generated by GProf revealed a computational complexity in terms of execution time in Motion Compensation (36%) and ExecuteLoofilters (16.7%) functions.

Valgrind is a GPL licensed tool. It includes several tools, one of which is Callgrind. We used Callgrind to get the exact number of operations performed when decoding a complete video sequence. The results are separated in terms of classes and in each class in terms of functions.

The resulting graph shows, for each function, the execution time as a percentage of the total execution time and the functions to which they appealed.

The graph shows that the most time-consuming functions are Motion Compensation and ExecuteLoofilters.

4.3. Profiling and functions Execution trace

As explained above, in order to determine the sequence of execution of the various functions of the decoder and the number of calls, a combination of code injection (manual code injection) and sampling profiling (using tools such as valgrind , Gprof, ...) has been made. The code injection in each function of the decoder has allowed us to locate called and calling functions. This also helps us to know the execution sequential order of different functions.

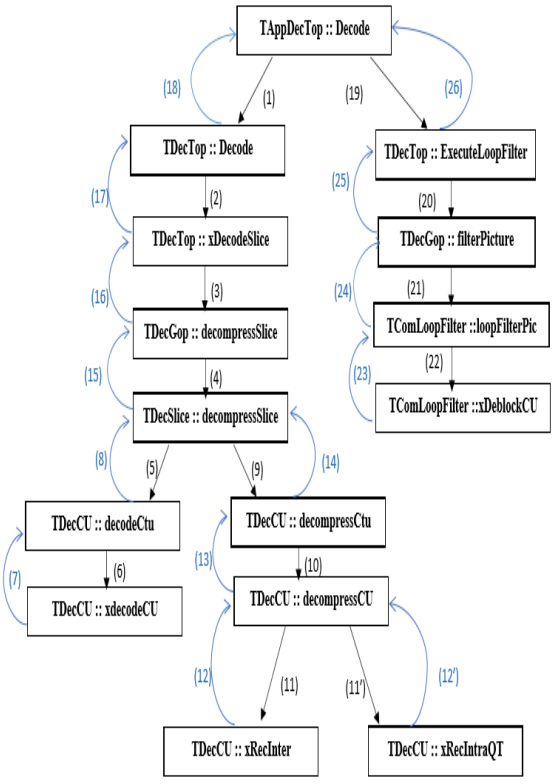


Figure 4. Functions execution Sequences

Values in parentheses indicate the execution order of decoder's functions.

The distribution of functions into groups is mainly based on three criteria: The first is to minimize communication between different groups of functions in order to have minimal inter-group communication and maximum intra-group communication. The second criterion is load balancing between the processors. Indeed, the target platform (Zedboard) is equipped with the two ARM V9 processors. To ensure optimal utilization of two processors, we must balance the charge of activity between the two cores. The third criterion is the execution parallelism. Indeed, we must ensure maximum parallel execution to minimize the average processing time of images.

Taking into account all the criteria, we came to the distribution of assigning the group functions 1, 2 and 4 to processor 0 (proc0) and functions of the group 3 to processor 1 (Proc1).

The calculations we have developed suggested the division mentioned above with the following performance data:

■ CPU.0:

Group 1: (1) (2) (3) (4) (15) (16) (17) (18) = 24.46%
Group 2: (5) (6) (7) (8) = 6.85%
Group 4: (19) (20) (21) (22) (23) (24) (25) (26) = 16.70%
Total activity time = 48.01%

■ CPU.1

Group 3: (9)(10)(11)(11') (12)(12')(13)(14)= 51.95%

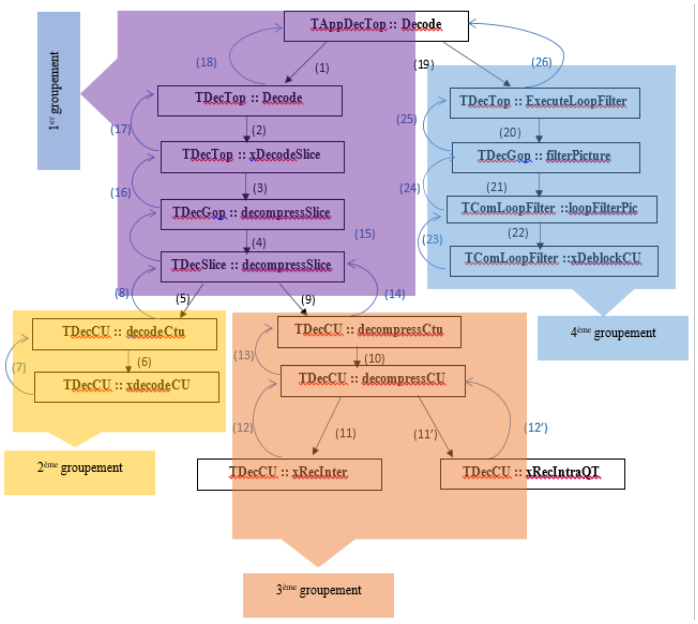


Figure 5. Grouping decoder functions

According to profiling procedure, we can divide the code on four main parts (Figure 4). The first contains the configurations functions. The second group represents the entropy decoding functions (decodeCTU and xdecodeCU). The third group is the reconstruction. Finally, the HEVC filters are grouped. We analyse their decoder's execution time in order to know its behaviour at runtime.

5. Experimental results

5.1. Test sequences

As an effort to carry out a good evaluation of the standard, the JCT-VC developed a document with some reference sequences and the codec configuration, which should be used with each one [29]. The sequences are divided into 6 groups (Classes) based on their temporal dynamics, frame rate, bit depth, resolution, and texture characteristics

A subset of six video sequences was selected from this list. These six video sequences were selected from three classes, such that two sequences from each class, namely Class A, Class C, Class F. Detailed descriptions of the sequences are given in Table II.

Table 2: Bitstreams used in the experimentations

Resolution (Sequence class)	Name	Frame rate	Frame number
2560x1600 (Class A)	*PeopleOnStreet*	30	150
	Traffic	30	300
832x480 (Class C)	*BQMall*	60	600
	PartyScene	30	500
1024x768 (Class F)	ChinaSpeed	50	500
	Slideediting	30	300

5.2. Zedboard Platform

Zedboard platform [30] (table III) is an evaluation platform based on a Zynq-7000 family [31]. It contains on the same chip two components. A dual-core ARM Cortex MPCore based on a high-performance processing system (PS) that can be used under Linux operating system or in a standalone mode and an advanced programmable logic (PL) from the Xilinx 7th family that can be used to hold hardware accelerators in multiple areas.

The two parts (PS and PL) interact between them by using different interfaces and other signals through over 3,000 connections [32]. Available four 32/64-bit high-performance (HP) Advanced eXtensible Interfaces (AXI) and a 64-bit AXI Accelerator Coherency.

Table 3: Zedboard technical specifications

Component	Characteristics
Processeur ZYNQ-7020 AP SOC XC7Z020-7CLG484CES	2 ARM Cortex A9 cores at 667 MHz
Memory	512 MB DDR3, 256 MB Quad-SPI Flash et SD Card
Communication	10/100/1000 Ethernet, USB OTG et USB UART
Extension	FMC (Low Pin Count) et 5 Pmod headers (2*6)
Display	HDMI output, VGA output et 128*32 OLED
Input / Output	8 switches, 7 push butons et 8 leds
Current and Voltage	3.0 A (Max) et 12V DC input
Certification	CE and RoHS certifier

In our experimentations, we have configured and compiled a custom Linux kernel downloaded from the official Linux kernel from Xilinx [31].

5.3. Pipeline execution approach

Since ZedBoard is based on Zynq chip that has dual cores, it is possible to distribute functions between the two cores.

Because of strict sequentially that we have seen in the functions execution of the HEVC decoder (HM Test Model version), the two processors will necessarily be in mutual exclusion of activity. To create parallel activity, we can transform the sequential video processing by a "pipelined" treatment.

If we refer to the grouping of functions explained above, we can consider that video sequence processing goes through four stages:

Stage 0: Download to a frame buffer (FB) the frame to be processed. This step corresponds to transition (1) in the graph (Fig. 5)

Stage 1: Decoding of the frame, corresponding to transitions (2, 3, 4, 5, 6, 7, 8). This stage corresponds to the Entropy decode process.

Stage 2: Decompression, corresponding to transitions (9, 10, 11-11 ', 12-12', 13, 14). This stage corresponds to the reconstruction process (that contains Inverse Transform, Inverse Quantification, Inter- prediction and Intra-prediction).

Stage 3: Filtering, which corresponds to transitions (15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26). This stage corresponds to the Loop Filter process

A parametric analysis of the application has enabled to deduce that the overall execution time of Stages 0, 1 and 3 is approximately equivalent to the execution time of Stage 2. Therefore, we can affect stages 0, 1 and 3 in the first PROC0 processor and Stage 2 to the second processor PROC1.

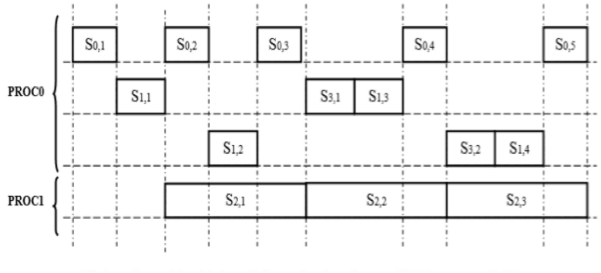


Figure 6. Pipeline execution based on dual cores Processor

Si,j represents the execution sequence of stage i applied to the frame number j. The first three rows reflect the sequencing of stage Si,j on the same processor PROC0:

Process P0, P1 and P2 support respectively the first, second and the third row.

The fourth row reflects the execution sequence of stage2 of the processor PROC1 applied to successive frames. A process P3 supports it.

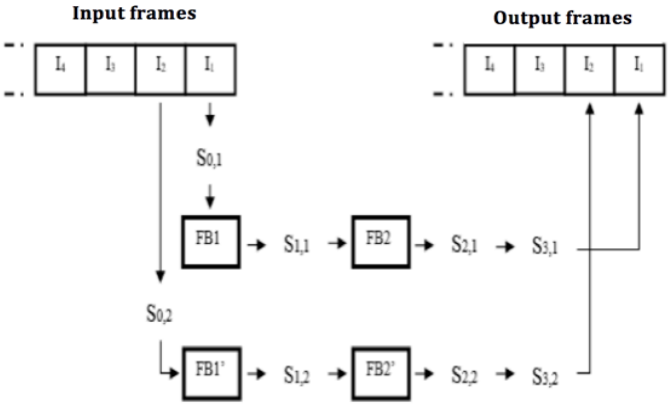


Figure 7. Sequence Process on two processors (cores)

Results obtained for the RA, AI and LD configurations are presentaed in tables IV, V and VI..

Table 4: Results for AI configuration

Class	Sequence	Frames Number	QP	Sequential Time (ms)	Pipelined Time (ms)
A	People On Street	150	22	594.28	386.28
			32	454.59	295.49
	Traffic	150	22	610.68	396.94
			32	432.75	281.29
Average execution time				523.075	340
Gain					35%
C	Party Scene	500	22	307.41	218.26
			32	219.16	155.61
	BQ Mall	600	22	259.30	184.10
			32	191.85	136.21
Average execution time				244.43	173.54
Gain					29%
F	China Speed	500	22	382.39	287.68
			32	295.93	207.15
	Slide Editing	300	22	279.75	199.83
			32	242.71	169.9
Average execution time				300.19	210.14
Gain					28%

Table 5: Results for RA configuration

Class	Sequence	Frames Number	QP	Sequential Time (ms)	Pipelined Time (ms)
A	People On Street	150	22	409.51	274.37
			32	256.97	172.17
	Traffic	150	22	293.26	196.48
			32	179.19	120.06
Average execution time				284.77	190.77
Gain					33%
C	Party Scene	500	22	153.41	107.38
			32	89.64	62.75
	BQ Mall	600	22	131.76	93.55
			32	84.08	59.70
Average execution time				114.72	80.84
Gain					29.5%
F	China Speed	500	22	216.84	156.12
			32	140.75	101.34
	Slide Editing	300	22	56.93	38.14
			32	50.25	33.67
Average execution time				116.19	82.31
Gain					29.15%

Table 6: Results for LD configuration

Class	Sequence	Frames Number	QP	Sequential Time (ms)	Pipelined Time (ms)
A	People On Street	150	22	446.65	294.79
			32	266.68	176.01
	Traffic	150	22	310.86	220.71
			32	164.25	116.62
Average execution time				297.11	202.03
Gain					32%
C	Party Scene	500	22	176.65	125.42
			32	91.99	60.71
	BQ Mall	600	22	142.03	99.42
			32	85.04	56.13
Average execution time				123.9275	85.42
Gain					31%
F	China Speed	500	22	217.94	155.84
			32	143.83	99.93
	Slide Editing	300	22	45.43	35.35
			32	41.39	27.85
Average execution time				112.14	74.01
Gain					28.89%

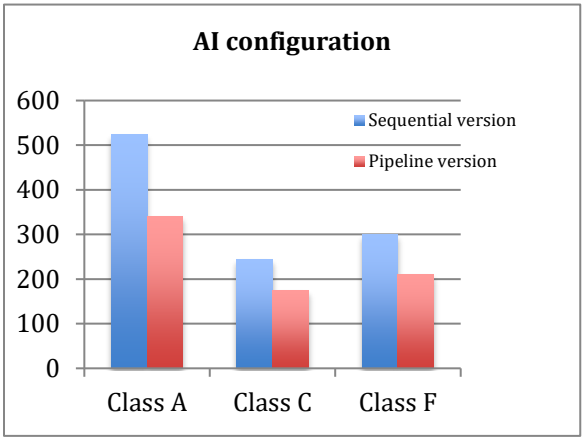


Figure 8. Execution time gain (AI configuration)

The experimental tests use reference sequences for the three configurations (AI, RA and LD) and with different quantification parameters (QPs). For the different configurations, results show a considerable gain compared to the sequential version (around 30%). We also note that class A frames have more processing time than the other classes (C and F) with a slightly more significant gain. This is logical since the resolution in this class is more important than for the other classes.

6. Conclusion

The objective of this article was to study the behaviour of the HEVC decoder represented by the reference application Test Model (HM). This study permitted us to discover the different functions of the decoder, their size, the call graphs, the percentage of CPU utilization and the number of instructions for each function.

Moreover, we have identified the most consuming functions in terms of CPU execution time and memory size, and then we have represented the execution trace of the functions by the use of an injection code technics. All these profiling results allowed us to divide the functions into groups. Once the regrouping is done, we proposed a parallelism approach based on the pipeline principle in order to run the application on the two cores of the Zedboard Platform. The experimental results found show an acceleration of about 30% compared to the sequential version.

Other technics can be added to improve this approach such as the increase of the number of processors or the implementation of hardware accelerators in the PL side.

References

[1] H. Smei, K. Smiri, A. Jemai, “Pipelining the HEVC Decoder on ZedBoard Plateform”, in 11th International Design & Test Symposium, IDT 2016, Hammamet, Tunisia, December 18-20, 2016. IEEE 2016, ISBN 978-1-5090-4900-4(IDT 2016).

[2] Cisco, Visual Networking Index (VNI): Forecast and Methodology, 2015-2020.

[3] JCT-VC - Joint Collaborative Team on Video Coding - [Online]:http://www.itu.int/en/ITU-T/studygroups/2013-2016/16/Pages/video/jctvc.aspx

[4] BSD Licence HEVC decoder (HM), Reference web site : [Online]:https://hevc.hhi.fraunhofer.de, code available Online at https://github.com/bbc/vc2-reference. Accessed, September 2016

[5] Bingjie Han , Ronggang Wang, Zhenyu Wang, Shengfu Dong, Wenmin Wang, Wen Gao, “HEVC decoder acceleration on multi-core x86 platform” IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2014.

[6] D. Marpe and T. Wiegand, J. Sullivan, Microsoft Corporation, IEEE Communications Magazine. August 2006.

[7] Article about HEVC - [Online]:http://en.wikipedia.org/wiki/High_Efficiency_Video_Coding

[8] Video Codec for Audiovisual Services at px64 kbit/s, ITU-T Rec. H.261, version 1: Nov. 1990, version 2: Mar. 1993.

[9] Video Coding for Low Bit Rate Communication, ITU-T Rec. H.263, Nov. 1995 (and subsequent editions).

[10] Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mbit/s—Part 2: Video, ISO/IEC 11172-2 (MPEG-1), ISO/IEC JTC 1, 1993.

[11] Coding of Audio-Visual Objects—Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Visual version 1), ISO/IEC JTC 1, Apr. 1999 (and subsequent editions).

[12] Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video, ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG 2 Video), ITU-T and ISO/IEC JTC 1, Nov. 1994.

[13] Advanced Video Coding for Generic Audio-Visual Services, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), ITU-T and ISO/IEC JTC 1, May 2003 (and subsequent editions).

[14] H.Krichene Zrida, A.Jemai, A.C.Ammari, M.Abid, “High Level H.264/AVC Video Encoder Parallelization for Multiprocessor Implementation”, Date’09, International Conference on Design, automation &Test. Nice, France.

[15] P. GB, P. NS and R. Adireddy. “Analysis of HEVC/H265 Parallel Coding Tools”. PathPartner Technology consulting Pvt. Ltd. White Paper. (2014).

[16] W. Liu, J. Li and Y. B. Cho. “A novel architecture for parallel multi-view HEVC decoder on mobile device”. Journal on Image and Video Processing (2017).

[17] C. Ching, M. Alvarez-Mesa, J. L. Ben Juurlink and T. Schierl. “Parallel HEVC Decoding on Multi- and Many-core Architectures. A Power and Performance Analysis”. J Sign Process Syst (2013).

[18] M. Zhou, V. Sze and M. Budagavi. “Parallel Tools in HEVC for High-Throughput Processing”. Applications of Digital Image Processing XXXV, Vol. 8499, 849910, (2014).

[19] B. Bross and al. “HEVC Real-time Decoding”, SPIE Proceedings Vol. 8856: Applications of Digital Image Processing XXXVI, (2013).

[20] B. Han and al. “HEVC Decoder acceleration multi-core X86 plateform”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), (2014).

[21] M. Alvarez-Mesa and al. “Parallel video decoding in the emerging HEVC standard”, 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), (2012).

[22] F. Saab, I.H. Elhajj, A. Kayssi and A. Chehab, “ Profiling of HEVC encoder", Article in Electronics Letters, July 2014.

[23] B.Han, R.Wang, Z.Wang, S.Dong, W.Wang and W.Gao, "HEVC Decoder Acceleration On Multi-Core X86 Platform", IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2014.

[24] C.Herglotz, D.Springer, M.Reichenbach, B.Stabernack, and A.Kaup, "Modeling the Energy Consumption of the HEVC Decoding Process", DOI 10.1109/TCSVT.2016.2598705, IEEE, 2016.

[25] R.Harada, Y.Matsuo And J.Katto, "Improvement Of H.265/HEVC Encoding For 8K UHDTV By Detecting Motion Complexity", IEEE International Conference on Consumer Electronics (ICCE), 2016.

[26] Gprof web Site. https://sourceware.org/binutils/docs/gprof/

[27] Gprof tutorial. [Online]:https://www.gadgetdaily.xyz/apply-gradual-transition-effects-to-page-elements/

[28] Valgrind Web site. [Online]:http://valgrind.org/.

[29] F. Bossen, "Common HM test conditions and software reference configurations," in 12th Meeting: Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC 1/SC 29/WG 11, Geneva, 2013, document JCTVC-L1100.

[30] Zedboard Plateform web site. www.zedboard.org. Accessed, September 2016

[31] Xilinx, Inc. (2016). Zynq-7000 All Programmable SoC Technical Reference Manual. http:// www.xilinx.com/support/documentation/user_ guides/ug585-Zynq-7000-TRM.pdf.

[32] Linux kernel from Xilinx. [Online]:https://github.com/Xilinx/linux-xlnx

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 49-55 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Flexible Aperture Tuning Solution for Cellular Main Antenna in Metallic Back Cover Mobile Phone

Yew Choon Mark Tan*, Guan Hong Ng, Yew Siow Roger Tay

AAC Technologies Pte. Ltd., Singapore

ARTICLE INFO

Article history:
Received: 06 September, 2017
Accepted: 11 October, 2017
Online: 19 October, 2017

Keywords:
Metallic Back Cover
Tunable Antenna
Aperture Tuning

ABSTRACT

Metal housing has been used extensively on portable communication devices such as on mobile phones and tablets. The choice of metal housing ranges from metallic rim to metallic back cover. This metal housing tends to improve the outlook appearance of the mobile devices, and add mechanical strength towards the mobile devices. However, from the aspect of the communication antenna, the metal housing often posts great challenges towards the flexibility in antenna design and reduction in antenna performance. This paper presents an approach to overcome the challenges by integrating the metal housing of the mobile phone as part of the antenna, along with the introduction of tunable antenna concept to provide different forms of Aperture Tuning to the Cellular Main Antenna, to satisfy its wide frequency band coverages for the 2nd, 3rd and 4th Generation (2G, 3G and 4G) mobile network.

1. Introduction

Over the years, there had been a popular trend for mobile phones to be equipped with metal housing. The metal housing often comes in different forms, which examples are as shown in Figure 1. One form of metal housing is the Closed Metal Ring as shown in Figure 1(a), with a metal ring running around the edges of the mobile phone. Another type of metal housing, quite similar to the Closed Metal Ring, is the Broken Metal Ring (with open slot) as shown in Figure 1(b). In the Broken Metal Ring, there are slots cutting on the metal ring, dividing the metal ring into two or more sections. In recent years, there is another type of metal housing that has been becoming popular. This metal housing comes in the form of a Metallic Back Cover as shown in Figure 1(c). The use of metal housing on mobile phone tends to improve the outlook appearance of the mobile phone. Apart from that, the robustness of the metal housing tends to add mechanical strength to the mobile phone. However, the use of metal housing on mobile phone often post disadvantages towards the mobile phone antenna. These disadvantages include introducing unwanted coupling towards the antenna, creating a shielding effect over the antenna hence reducing the sensitivity of the antenna etc. Additionally, maintaining the aesthetics of the metal housing on the mobile phone comes with an antenna design flexibility cost.

(a) (b) (c)

Figure 1. Different forms of metal housing: (a) Closed Metal Ring; (b) Broken Metal Ring (with open slot); (c) Metallic Back Cover;

This paper serves as an extension of work originally presented in 2017 11th European Conference on Antennas and Propagation (EuCAP2017) [1]. This paper presents a flexible Aperture Tuning solution [2 – 4] for the cellular Main Antenna in a Metallic Back Cover mobile phone, to provide coverage for multi frequency bands in the frequency range of 700MHz to 960MHz (LTE700, GSM850, GSM900), 1710MHz to 2170MHz (DCS1800, PCS1900, WCDMA2100) and 2300MHz to 2690MHz (LTE2300, LTE2600). The flexible Aperture Tuning solution as proposed in this paper aims to overcome the disadvantages post by the metal housing as mentioned above. In the earlier paper [1], the flexible Aperture Tuning solution for the cellular Main Antenna has been briefly introduced, along with the Free-Space passive measurement data of the Main Antenna's Reflection Coefficient S11 and Total Efficiency. This paper will further elaborates on the concept of the different Aperture Tuning solutions that are deployed at different parts of the Main Antenna for tuning the antenna's Low (700 – 960MHz) Band, Mid (1710 – 2170MHz)

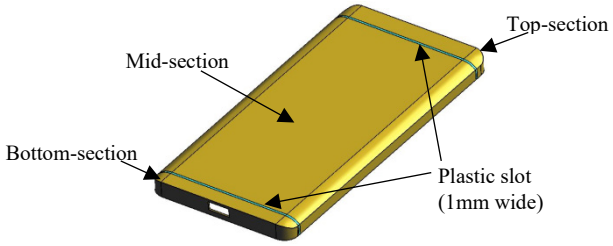
*Corresponding Author: Yew Choon Mark TAN, 85 Science Park Drive, #02-08 The Cavendish, Singapore 118259, (+65) 91767574 & marktan@aactechnologies.com

Band and High (2300 – 2690MHz) Band. Besides the Free-Space passive measurement data, this paper also presents the active measurement data (Total Radiation Power TRP) taken with the mobile phone mockup in the Free-Space and Phantom Head & Hands conditions.

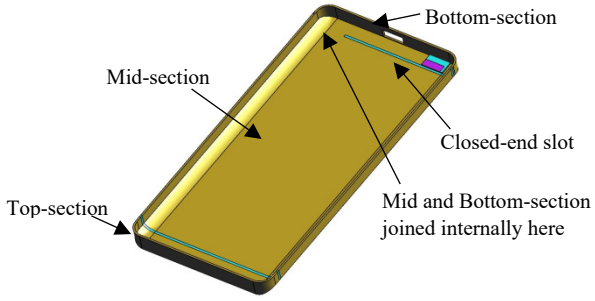
2. Metallic Back Cover Mobile Phone Mockup

2.1. Metallic Back Cover

The metal housing for the mobile phone presented in this paper takes the form of a Metallic Back Cover, as mentioned in Figure 1(c). Figure 2 shows pictures illustrating the exterior and interior appearance of the Metallic Back Cover. From the exterior appearance shown in Figure 2(a), the Metallic Back Cover seems to be divided into three sections, namely Top-section, Mid-section and Bottom-section. In between each section, there is a plastic slot (at 1mm wide) running across the Metallic Back Cover. However, from the interior appearance shown in Figure 2(b), it can be seen that the Mid-section and Bottom-section of the Metallic Back Cover are actually joined together internally. This makes the slot at the Bottom-section more like a Closed-end slot instead of a through slot as seen at the Top-section. The overall dimension of the Metallic Back Cover is 144mm by 70mm by 6.5mm, which is suitable for mobile phone with 5.5inch display.



(a) Exterior appearance



(b) Interior appearance

Figure 2. Metallic Back Cover

In this proposed solution, the Main Antenna is located at the bottom section of the mobile phone. The entire Bottom-section metallic cover and partial of the Mid-section metallic cover are integrated as part of the Main Antenna design. Figure 3 shows pictures mainly highlighting the interior appearance at the bottom section of the Metallic Back Cover. There is a Capacitive Feed Patch located near the bottom left corner of the Metallic Back Cover. This Capacitive Feed Patch is a thin layer of conductive rectangular patch, separated from the inner surface of the Metallic Back Cover by a thin layer of plastic substrate with thickness of 0.5mm. The Capacitive Feed Patch is positioned on the Bottom-section side of the Metallic Back Cover in a way such that the edge of the Capacitive Feed Patch will align well with the edge of the Closed-end slot, as can be seen in Figure 3. Figure 3 also highlights the installation of a Tuner Board on the inner surface of the Metallic Back Cover. The Tuner Board is a tiny piece of Printed Circuit Board (PCB) that is used to house a Tunable Capacitor TC1. The Tuner Board is fully grounded onto the Mid-section of the Metallic Back Cover, positioned just by the edge of the Closed-end slot. The output of the Tunable Capacitor TC1 is bridged across the slot and get into direct contact with the Bottom-section of the Metallic Back Cover.

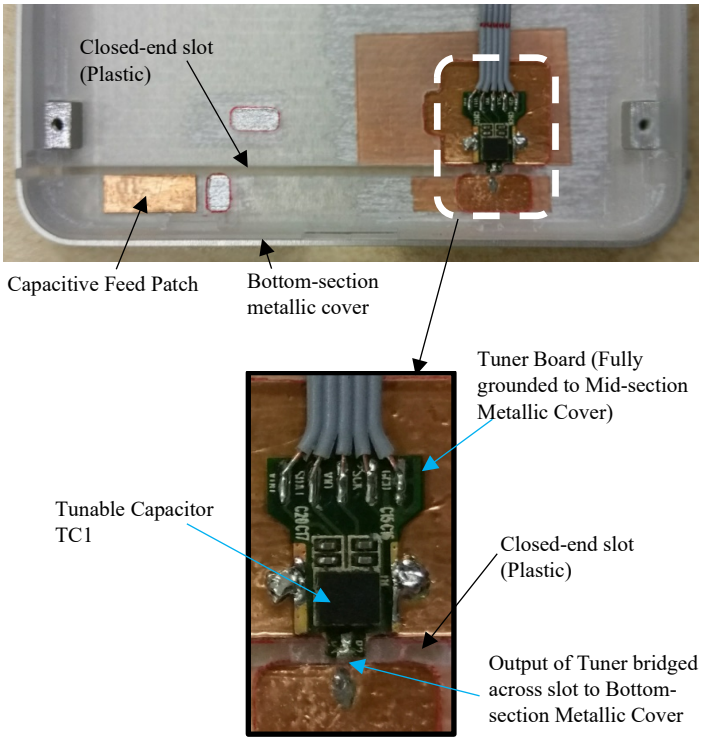


Figure 3. Capacitive Feed Patch and Tuner Board on Metallic Back Cover

2.2. Main Antenna Printed Circuit Board (PCB)

Figure 4 shows partial of the Main Antenna Printed Circuit Board (PCB) used in the mobile phone mockup, mainly highlighting the environment in the Main Antenna region at the bottom section of the mobile phone. The PCB is equipped with a Type-C Micro-USB and a speaker which are components that are commonly found at the bottom section of mobile phone. On the PCB, there is a Feed-Pin that is traced to the source of the Main Antenna. The Feed-Pin can be in the form of a pogo-pin or spring contact to establish a direct connection with the Capacitive Feed Patch at the location (red-dot) as highlighted in Figure 4. Located next to the Feed-Pin is Gnd-Pin1, which is traced to a Tunable LC Resonant Circuitry, with reference to the common ground. This Tunable LC Resonant Circuitry is made up of a shunt inductor in parallel with a shunt Tunable Capacitor TC2, to provide Aperture Tuning to the Main Antenna. The Gnd-Pin1 can be in the form of a pogo-pin or spring contact to make direct contact with the Bottom-section metallic cover at the location (blue-dot) as indicated in Figure 4. Located above Gnd-Pin1 on the PCB is Gnd-Pin2 which is traced to a standalone Tunable Capacitor TC3 connecting to the reference ground. This standalone Tunable Capacitor is meant for providing Aperture Tuning to the Main Antenna. The Gnd-Pin2 makes direct connection with the Mid-

section of metallic cover at the location (purple-dot) as indicated in Figure 4. To the left of Gnd-Pin2, located near the speaker is Gnd-Pin3 which is connected directly to the reference ground. Gnd-Pin3 on the PCB is meant to make direct connection with the Mid-section of metallic cover at the location (green-dot) next to the Tuner Board as shown in Figure 4. This arrangement of the grounding connections between the respective ground pins (Gnd-Pin1, Gnd-Pin2 and Gnd-Pin3) on PCB and the Metallic Back Cover has integrated the Metallic Back Cover as part of the Main Antenna design. The three Tunable Capacitors (TC1, TC2 and TC3) as mentioned above on the Metallic Back Cover and Main Antenna PCB are Micro-Electro-Mechanical Systems (MEMS) based variable capacitors. Each capacitor has 16 different states of capacitance value ranging from 0.3pF to 2.97pF programmed in its register. The proposed tunable solution for the Main Antenna in this paper is not limited to the technology or type of the Tunable Capacitor.

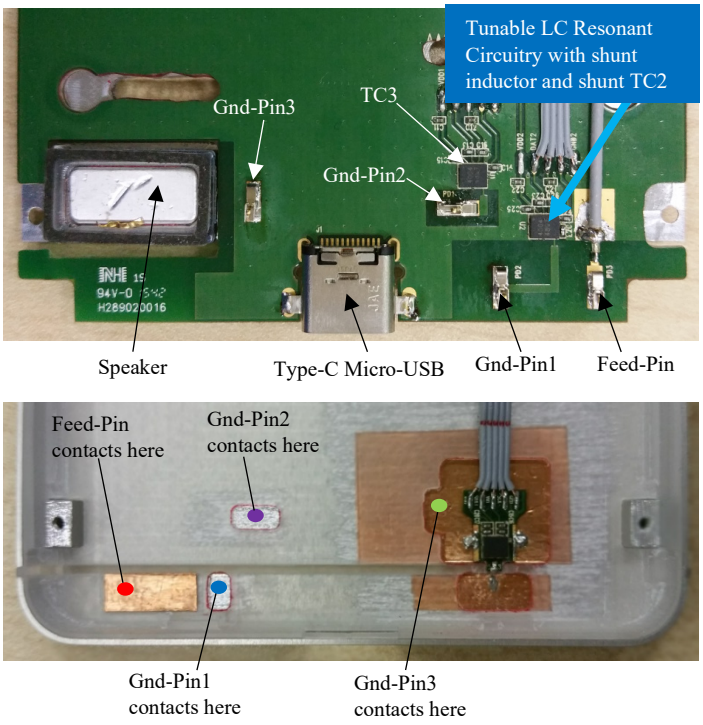


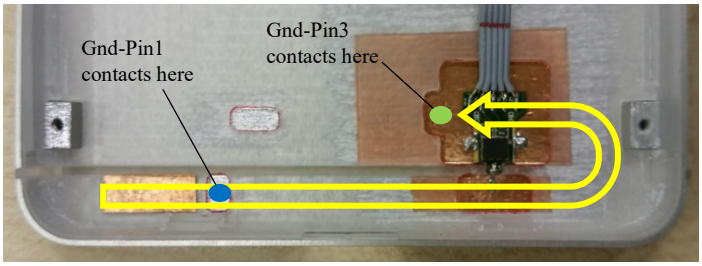
Figure 4. PCB layout for the Main Antenna and integrating the Metallic Back Cover as part of the Main Antenna design

3. Flexible Aperture Tuning Solution for Main Antenna

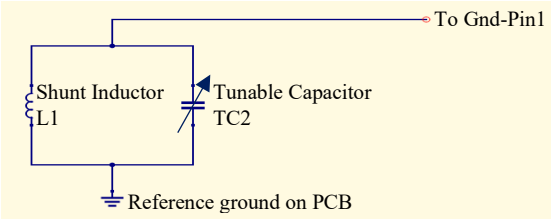
The Main Antenna of the mobile phone has operating frequency bands that are divided into the Low Band, the Mid Band and the High Band. The Low Band covers frequency range from 700MHz to 960MHz, the Mid Band covers frequency range from 1710MHz to 2170MHz and the High Band covers frequency range from 2300MHz to 2690MHz. These ranges of frequency bands are sufficient to provide coverage for the 2G, 3G and 4G network. The Aperture Tuning of the Main Antenna as proposed in this paper will be broken down to sections with respect to the following bands: Low Band, Mid Band and High Band.

3.1. Low Band Tuning (700MHz to 960MHz)

The proposed Main Antenna solution is fed using the Capacitive-Fed Coupling Technique [5 – 6] to the Bottom-section metallic cover via the Capacitive Feed Patch as mentioned above in Figure 3. Energy from the Capacitive Feed Patch is coupled onto the Bottom-section metallic cover. Current on the Bottom-section metallic cover will then flows in the direction as indicated by the yellow-coloured arrow (as shown in Figure 5(a)) to Gnd-Pin3 on the Main Antenna PCB. This forms a Loop Antenna for the Low Band of the Main Antenna. Aperture Tuning is applied to this Loop Antenna at Gnd-Pin1, via the Tunable LC Resonant Circuitry [2 – 4] (as mentioned in Figure 4) which equivalent circuit is as shown in Figure 5(b). The resonance at the Low Band of the Main Antenna can be made to shift within the frequency range of 700MHz to 960MHz by tuning the capacitance value of the Tunable Capacitor TC2 on the PCB. In tuning the Low Band of the Main Antenna, the capacitance value of the Tunable Capacitor TC1 on the Metallic Back Cover and the Tunable Capacitor TC3 at Gnd-Pin2 on the PCB are kept constant at 2.97pF and 0.66pF respectively.



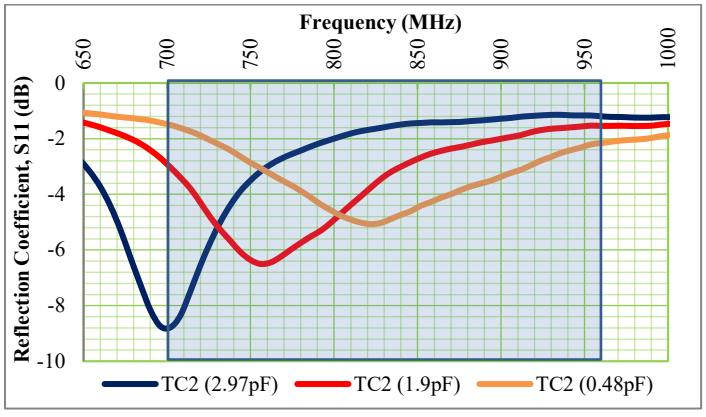
(a) Current flow for Low Band Tuning



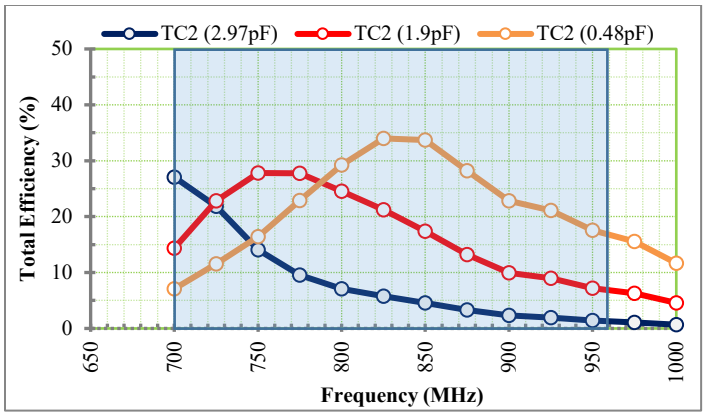
(b) Equivalent circuit of Tunable LC Resonant Circuit

Figure 5. Low Band Tuning of Main Antenna

Figure 6 shows plots of the passive measurement result for the Low Band Tuning of the Main Antenna, in Free-Space condition. Figure 6 (a) is the Reflection Coefficient S11 of the Main Antenna measured using a Network Analyzer. Figure 6(b) shows plots of the Total Efficiency of the Main Antenna in the Low Band region, measured in an Anechoic Over-the-Air (OTA) Chamber. From the two plots shown in Figure 6, it can be seen that three different



(a) Reflection Coefficient S11 in Low Band region



(b) Total Efficiency in Low Band region

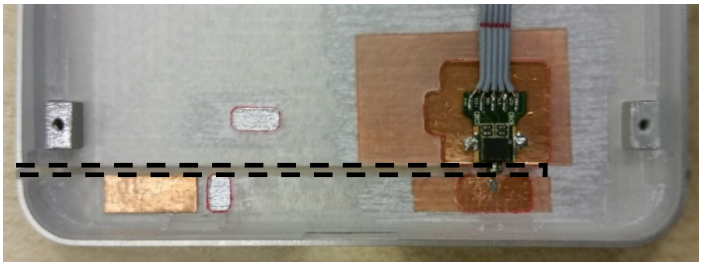
Figure 6. Passive measurement results in Low Band region

capacitance value of TC2 (at 0.48pF, 1.9pF and 2.97pF) are used to shift the Low Band resonance across the Low Band region of 700MHz to 960MHz. This shift in the Total Efficiency plot (as seen in Figure 6(b)) helps to maintain the Total Efficiency in the Low Band region at above 20%.

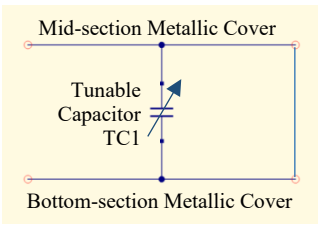
3.2. Mid Band Tuning (1710MHz to 2170MHz)

The Mid Band of the Main Antenna is resulting from a Slot Antenna form on the Metallic Back Cover by the Closed-end slot, as highlighted with the black-dotted line in Figure 7(a). The Tunable Capacitor TC1, which output is bridged across the Closed-end slot on the Metallic Back cover, is being used to provide Aperture Tuning to this Slot Antenna. The configuration of having the Tunable Capacitor TC1 across the Closed-end slot can be best described by the equivalent circuit as shown in Figure 7(b). Hence, by tuning the capacitance value of Tunable Capacitor TC1, the wavelength of the Slot Antenna will change accordingly, shifting the Mid Band resonance within the frequency range of 1710MHz to 2170MHz. In tuning the Mid Band of the Main Antenna, the capacitance value of both the Tunable Capacitor TC2 and TC3 on the Main Antenna PCB are kept constant at 2.97pF.

Figure 8 shows plots of the passive measurement result for the Mid Band Tuning of the Main Antenna, in Free-Space condition. Figure 8(a) is the measured Reflection Coefficient S11 of the Main Antenna and Figure 8(b) shows plots of the measured Total Efficiency of the Main Antenna in the Mid Band region. From the two plots shown in Figure 8, it can be seen that the Mid Band resonance can be made to shift across the Mid Band region of 1710MHz to 2170MHz by tuning the capacitance value of TC1 at 1.55pF, 1.9pF, 2.44pF and 2.97pF respectively. This shift in the Total Efficiency plot (as seen in Figure 8(b)) helps to maintain the Total Efficiency in the Mid Band region at above 20%.

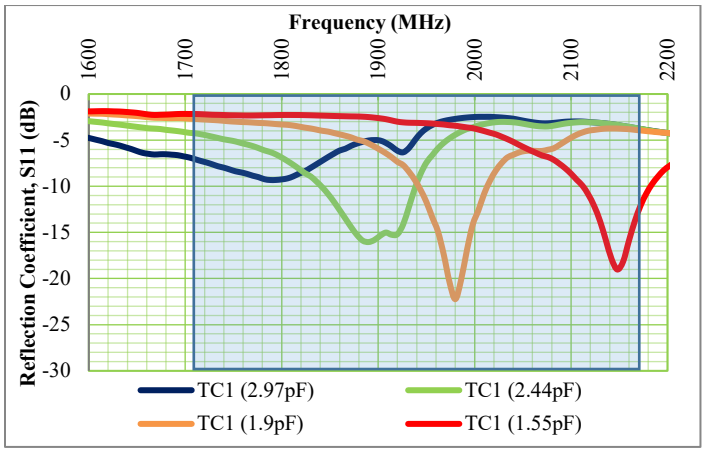


(a) Slot Antenna for Mid Band Tuning

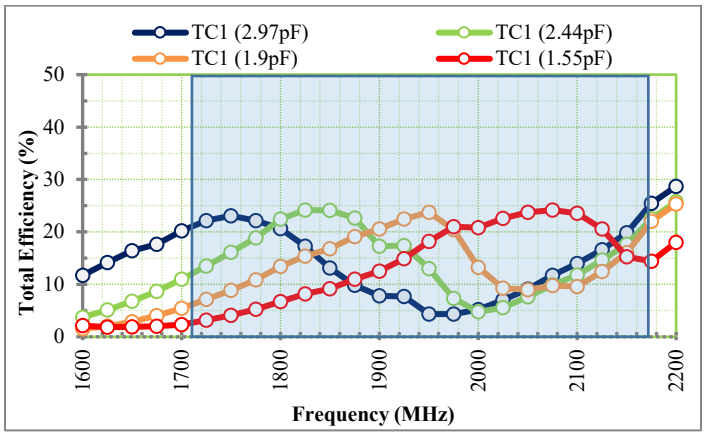


(b) Equivalent circuit of Tuner Board on Metallic Back Cover

Figure 7. Mid Band Tuning of Main Antenna



(a) Reflection Coefficient S11 in Mid Band region



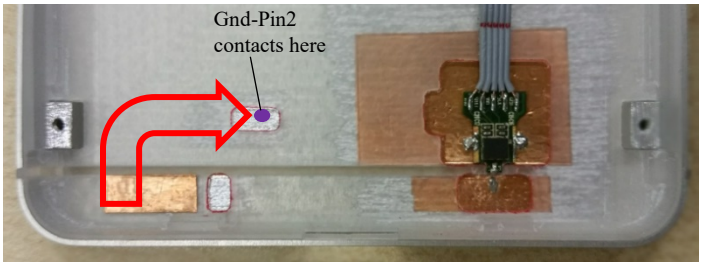
(b) Total Efficiency in Mid Band region

Figure 8. Passive measurement results in Mid Band region

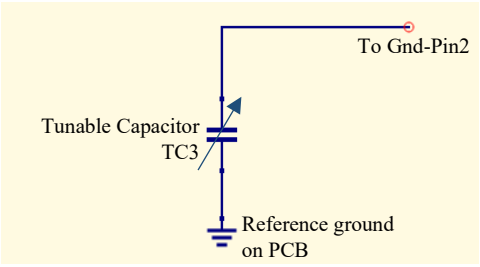
3.3. High Band Tuning (2300MHz to 2690MHz)

In High Band Tuning of the Main Antenna, the energy from the Capacitive Feed Patch on the Bottom-section metallic cover is coupled across the Closed-end slot to the Mid-section metallic cover. Current on the Mid-section metallic cover will then flows in the direction as indicated by the red-coloured arrow to Gnd-Pin2 on the Main Antenna PCB, as shown in Figure 9 (a). This actually forms a small Loop Antenna for the High Band of the Main Antenna. Aperture Tuning is applied to this small Loop Antenna at Gnd-Pin2, via the Tunable Capacitor TC3, which equivalent circuit is as shown in Figure 9 (b). The resonance at the High Band of the Main Antenna can be made to shift within the frequency range of 2300MHz to 2690MHz by tuning the capacitance value of the Tunable Capacitor TC3 on the PCB. In tuning the High Band of the Main Antenna, the capacitance value of the Tunable

Capacitor TC1 on the Metallic Back Cover and the Tunable Capacitor TC2 in the Tunable LC Resonant Circuit on the PCB are both kept constant at 2.97pF.

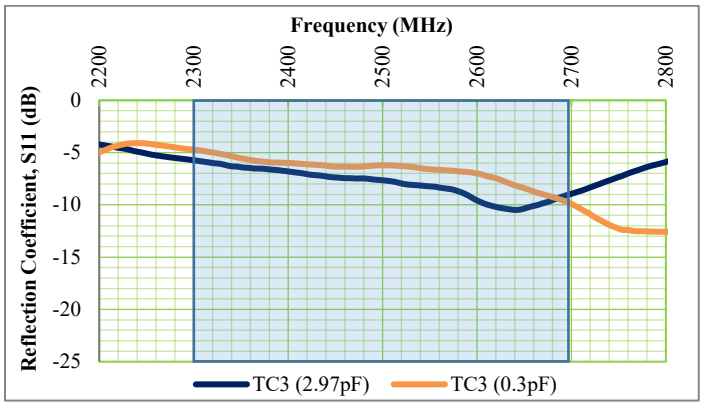


(a) Current flow for High Band Tuning

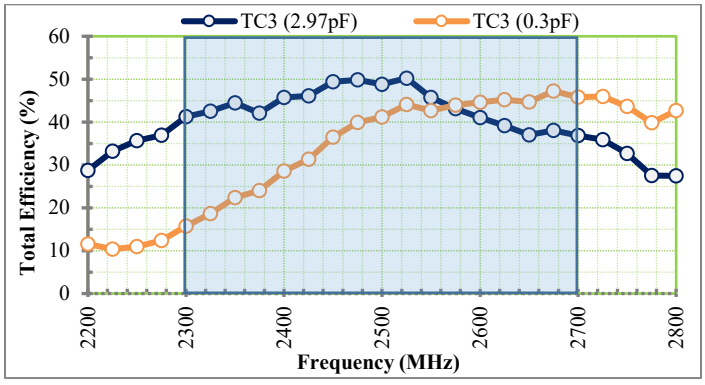


(b) Equivalent circuit of Aperture Tuning at Gnd-Pin2

Figure 9. High Band Tuning of Main Antenna



(a) Reflection Coefficient S11 in High Band region



(b) Total Efficiency in High Band region

Figure 10. Passive measurement results in High Band region

Figure 10 shows plots of the passive measurement result for the High Band Tuning of the Main Antenna, in Free-Space condition. Figure 10 (a) is the measured Reflection Coefficient S11 of the Main Antenna and Figure 10 (b) shows plots of the measured Total Efficiency of the Main Antenna in the High Band region. From the two plots shown in Figure 10, it can be seen that two-different capacitance value of TC3 (0.3pF and 2.97pF) are used to shift the High Band resonance across the High Band region of 2300MHz to 2690MHz. This shift in the Total Efficiency plot (as seen in Figure 10(b)) helps to maintain the Total Efficiency in the High Band region at above 40%.

4. Mobile Phone Mockup in Active Phone Condition

In this section, the Metallic Back Cover mobile phone mockup as reported above is configured to be an active phone. Figure 11 shows a picture illustrating the Metallic Back Cover mobile phone mockup in its active phone configuration. The mockup is equipped with an in-house designed Main PCB which is used to house a Long-Term Evolution (LTE) module for conducting active Total Radiated Power (TRP) measurement in an Anechoic OTA Chamber. The Main PCB is equipped with power contact points to draw power from a mobile phone battery, to power up the LTE Module as well as the Tunable Capacitors (TC1, TC2 and TC3 mentioned above in section 2 and 3) during the active measurement.

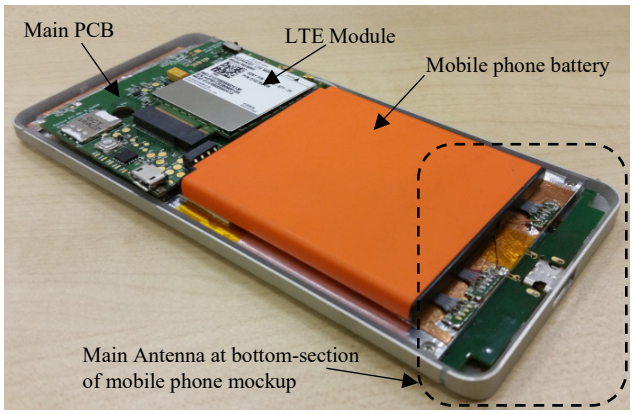


Figure 11. Mobile phone mockup in active condition

4.1. Measurement in Free-Space

The active TRP measurement in the Free-Space condition was conducted in an Anechoic OTA Chamber, for four different 3G bands and five different 4G LTE bands. These bands are WCDMA Band 1, 2, 5 and 8 in the 3G network, and FDD LTE Band 1, 4, 7, 8 and 13 in the 4G LTE network. Table 1 tabulates the active TRP measurement results of the Main Antenna, with the metallic cover mobile phone mockup in Free-Space condition. The first column in Table 1 lists the respective nine frequency bands in the 3G and 4G network. The second column displays the mid channel for each band respectively, and the third column lists the frequency (in MHz) corresponding to respective mid channel in the second column. As can be seen from the third column, this measurement covers the various frequencies which fall within the Low (700MHz to 960MHz), Mid (1710MHz to 2170MHz) and High (2300MHz to 2690MHz) Band regions of the Main Antenna. The fourth column in Table 1 lists the measured Conducted Power (in dBm) of the LTE Module, transmitting at the respective frequencies listed in the third column. And lastly, the measured Free-Space TRP (in dBm) of the Main Antenna on the Metallic Back Cover mobile phone mockup is as listed in the fifth column. From Table

1, the Conducted Power of the LTE Module is ranging from 20 to 23dBm, and the Free-Space TRP ranges from 14 to 20dBm.

Table 1. Active TRP measurement results in Free-Space

Band	Channel	Frequency (MHz)	Conducted Power (dBm)	Free-Space TRP (dBm)
WCDMA Band5	4182	836.6	22.3	17.8
WCDMA Band8	2787	897.6	22.2	16.7
WCDMA Band2	9400	1880.0	22.9	15.1
WCDMA Band1	9750	1950.0	22.4	14.1
FDD LTE Band13	23230	782.0	22.0	16.9
FDD LTE Band8	21625	897.5	21.2	16.6
FDD LTE Band4	20175	1732.5	21.8	15.7
FDD LTE Band1	18300	1950.0	21.7	13.8
FDD LTE Band7	21100	2535.0	20.2	19.7

4.2. Measurement with Phantom Head & Hands

To study the user effect towards the Main Antenna solution on the metallic cover mobile phone mockup, the active TRP measurement was further extended to the Phantom Head & Hands configuration. Figure 12 shows an example of the setup of the Phantom Head + Right Hand (BHHR) in the talk mode configuration. In this setup, the metallic cover mobile phone mockup will be gripped by the Hand, and placed beside the Phantom Head, as according to the CTIA Test Plan [7]. Two Phantom configurations had been considered in this measurement: Phantom Head + Left Hand (BHHL) and Phantom Head + Right Hand (BHHR). This active TRP measurement of the Main Antenna, with the mobile phone mockup on the Phantom Head & Hands was conducted in an Anechoic OTA Chamber, for four different 3G bands (WCDMA Band 1, 2, 5 and 8). Table 2 tabulates the active TRP measurement results of the Main Antenna, with the mobile phone mockup on the Phantom Head & Hands (BHHL and BHHR) talk mode configuration. The first three columns in Table 2 list the frequency bands, its mid channel and the corresponding frequency in the 3G network respectively. The fourth and fifth column in Table 2 list the measured TRP (in dBm) of the Main Antenna, in the BHHL and BHHR configuration respectively. From the BHHL TRP and BHHR TRP results listed in Table2, it can be clearly seen that the left hand has a greater influence on the Main Antenna performance as compared to the right hand. This is likely due to the hand is covering the opening of the Closed-end slot on the Metallic Back Cover in the BHHL configuration. The variation in TRP between BHHL and BHHR is between 1 to 3.3dBm, with the worst case scenarios taking place at WCDMA Band 2 (1880MHz). A comparison between the Free-Space TRP (in Table 1) and BHHL TRP (in Table 2) shows a reduction in the antenna performance in the latter case with variation of between 5.9 to 9.6dBm, and the worst case scenarios taking place at WCDMA Band 8 (897.6MHz). The comparison between the Free-Space TRP and BHHR TRP also shows a reduction in the antenna performance in the latter case, with variation of between 4.6 to 8.4dBm, and the worst case scenarios taking place at WCDMA Band 5 (836.6MHz).

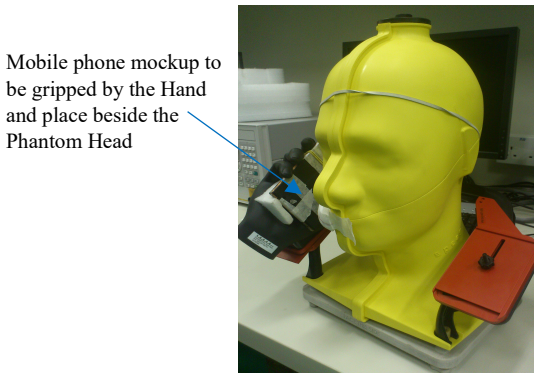


Figure 12. Phantom Head + Right Hand (BHHR) in talk mode setup

Table 2. Active TRP measurement results in BHHL and BHHR

Band	Channel	Frequency (MHz)	BHHL TRP (dBm)	BHHR TRP (dBm)
WCDMA Band5	4182	836.6	8.3	9.3
WCDMA Band8	2787	897.6	7.1	9.6
WCDMA Band2	9400	1880.0	7.0	10.3
WCDMA Band1	9750	1950.0	8.2	9.5

5. Conclusion

It has always been challenging in providing solution for antennas on mobile phones with metal housing. The metal housing often introduces unnecessary coupling, shielding effect towards the antennas as well as reducing the flexibility in designing the antenna. In this paper, a tunable solution has been proposed for the cellular Main Antenna of a mobile phone with metal housing in the form of metallic back cover. The proposed solution demonstrates the integration of the metal housing as part of the antenna design, along with the use of different Aperture Tuning circuitries at different parts of the Main Antenna, to allow flexible tuning of the antenna's Low, Mid and High Band. In an attempt to study the user effect towards the Main Antenna on the metallic cover mobile phone, the measurement results suggested that the Left Hand influence on the antenna is greater as compared to the Right Hand. In the comparison between Free-Space and BHHL/BHHR condition, the antenna performance sees a reduction of less than 10dBm.

References

[1] Mark Y.C. Tan, Guan Hong Ng, Roger Tay, "A tunable LTE main antenna solution for metallic cover mobile phone", 2017 11th European Conference on Antennas and Propagation (EuCAP), 19 – 24 March 2017. https://doi.org/10.23919/EuCAP.2017.7928219

[2] P. A. Tornatta, R. Gaddi, "Aperture tuned antennas for 3G-4G applications using MEMS digital variable capacitor", 2013 IEEE MTT-S International Microwave Symposium Digest (IMS), 2 – 7 June 2013. https://doi.org/10.1109/MWSYM.2013.6697602

[3] Sameer Akhtar Shah, Liu Qianyun, "A novel approach to design an antenna aperture tuning circuit", 2015 Asia-Pacific Microwave Conference (APMC), 6 – 9 Dec 2015. https://doi.org/10.1109/APMC.2015.7413123

[4] Mark Y.C. Tan, Guan Hong Ng, Roger Tay, "Tuning of LTE Main Antenna Using Flexible Tunable LC Resonant Circuitry", 2017 11th European Conference on Antennas and Propagation (EuCAP), 19 – 24 March 2017. https://doi.org/10.23919/EuCAP.2017.7928223

[5] Wei-Yu Li, Chun-Yih Wu, Kin-Lu Wong and Ming-Fang Tu, "Internal small-size PIFA for LTE/GSM/UMTS operation in the mobile phone", 2010 IEEE Antennas and Propagation Society International Symposium (APSURSI), July 2010. https://doi.org/10.1109/APS.2010.5562325

[6] Norakamon Wongsin, Thanakarn Suangun, Chatree Mahatthanajatuphat, Prayoot Akkaraekthalin, "A multiband fractal ring antenna fed by capacitive coupling", 2011 8th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 17 – 19 May 2011. https://doi.org/10.1109/ECTICON.2011.5947808

[7] "Test plan for wireless device Over-the-Air performance", CTIA, version 3.6.2, May 2017. https://www.ctia.org/docs/default-source/certification/ctia-test-plan-for-wireless-device-over-the-air-performance-ver-3-6-2.pdf?sfvrsn=2

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 56-62 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

A New Identification Approach of MIMO Hammerstein Model with Separate Nonlinearities

Chekib Ghorbel *, Zeineb Rayouf, Naceur Benhadj Baraiek

Advanced System Laboratory, Polytechnic School of Tunisia BP 743, 2078, La Marsa, Tunisie.

ARTICLEINFO

Article history:
Received: 20 September, 2017
Accepted: 09 October, 2017
Online: 23 October, 2017

Keywords:
Parametric identification
Hammerstein model
Separate nonlinearity
Quadruple-tank process

ABSTRACT

A new coupled structure identification of Multi-Input Multi-Output (MIMO) Hammerstein models with separate nonlinearities is proposed. It is based on the use of the Recursive Least Squares (RLS) algorithm. A comparative study between a decoupled and coupled structures identification of MIMO Hammerstein models is discussed. A quadruple-tank process is used to illustrate the e ectiveness of the new structure.

1 Introduction

This paper is an extension of work originally presented in Conference on Science of Electronics, Technologies of Information and Telecommunications (SETIT 2016) [1].
The latter work presented two methods of parametric identification of decoupled multivariable Hammerstein model.
It consists of one nonlinear static block and one linear dynamic block. Many process (chemical and biological process, signal processing, *etc.*) have this structure, for example: pH neutralization processes [2], polymerization reactor [3], distillation columns [4] and dryer process [5].
Many system identification methods have been used to identify the single-input single-output (SISO) model. They can be divided into stochastic methods [6,7], iterative methods [8], over-parameterization methods [9], separable least squares methods [10,11], blind identification methods [12] and frequency domain methods [13].
It is possible to transform the SISO Hammerstein model to MISO and MIMO model which is linear in the parameters [14]. Serval approches have been proposed to identify MIMO Hammerstein model. In [15,16,17], neuronal networks and fuzzy logic have been used to deal with more general nonlinearities. An approach based on multivariable cardinal cubic spline functions to model the static nonlinearities have been proposed in [18]. The Least Squares Support Vector Machines (LS-SVMs) have been presented in [19,20]. A generalized Hammerstein model consisting of a static polynomial function in series with time-varying linear model is developed in order to model the Hammerstein-like multivariable processes whose linear dynamics vary over the operating space in [21]. In this work, we propose a new coupled structure identification of MIMO model with separate nonlinearities. It is organized as follows: SISO Hammerstein system is presented in part 1 of section 2. A new coupled structure for MIMO Hammerstein system is developed in part 2 of section 2. Simulation results of a quadruple-tank process is given in section 3. Finally, a conclusion is made.

2 Parametric identification of Hammerstein model

2.1 Parametric identification of SISO Hammerstein model

Assume that the Hammerstein model of Figure 1 is composed of a nonlinear block $F(.)$ associated with a linear sub-system $\frac{B(q^{-1})}{A(q^{-1})}$. It is described by:

$$\begin{cases} y_k = \frac{B(q^{-1})}{A(q^{-1})} v_k + w_k \\ v_k = F(u_k) \end{cases} \quad (1)$$

*Corresponding Author: Chekib Ghorbel, Advanced System Laboratory, Polytechnic School of Tunisia, BP 743, 2078, La Marsa, Tunisie.
Contact No: +21698722239.
Email: chekib.ghorbel@enicarthage.rnu.tn

https://dx.doi.org/10.25046/aj020607

with:

$$A\left(q^{-1}\right) = 1 + a_1 q^{-1} + a_2 q^{-2} + \ldots + a_{n_A} q^{-n_A}$$

$$B\left(q^{-1}\right) = b_1 q^{-1} + b_2 q^{-2} + \ldots + b_{n_B} q^{-n_B}$$

$$v_k = \lambda_1\, u_k + \lambda_2\, u_k^2 + \ldots + \lambda_N\, u_k^N$$

q^{-1} delay operator, u_k input of the system, y_k output, v_k the unmeasurable internal signal and w_k represents the modeling error, external disturbances, *etc*.

In order to have a unique parameterizations of the Hammerstein model structure, the first coefficient of the nonlinear function $F(.)$ equals to 1, $\lambda_1 = 1$ [11,22].

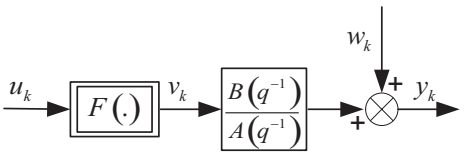


Figure 1: SISO Hammerstein model

The output y_k is given by:

$$y_k = -\sum_{i=1}^{n_A} a_i\, y_{k-i} + \sum_{i=1}^{n_B} b_i \left(u_{k-i} + \sum_{\rho=2}^{N} \lambda_\rho\, \mathrm{u}^\rho{}_{k-i} \right) + w_k \quad (2)$$

Equation (2) can be written in the following form:

$$y_k = \Phi_k^T\, \theta_k + w_k \quad (3)$$

with:

$$\Phi_k = \begin{pmatrix} Y_k \\ U_k \end{pmatrix},\ \theta = \begin{pmatrix} a \\ b \\ s \end{pmatrix},$$

Φ_k and $\theta_k \in R^{n_R}$ where $n_R = n_A + N\, n_B$,

$Y_k = \begin{pmatrix} -y_{k-1}, -y_{k-2}, \ldots, -y_{k-n_A} \end{pmatrix} \in R^{n_A}$,

$U_k = \begin{pmatrix} U_{1k}, U_{2k}, \ldots, U_{Nk}, \end{pmatrix} \in R^{N n_B}$,

$U_{j_k} = \begin{pmatrix} u^j{}_{k-1}, u^j{}_{k-2}, \ldots, u^j{}_{k-n_B} \end{pmatrix} \in R^{n_B}$,

for $j = 1, 2, \ldots, N$

$a = \begin{pmatrix} a_1, a_2, \ldots, a_{n_A} \end{pmatrix} \in R^{n_A}$,

$b = \begin{pmatrix} b_1, b_2, \ldots, b_{n_B} \end{pmatrix} \in R^{n_B}$,

$s = \begin{pmatrix} \lambda_2 b, \lambda_3 b, \ldots, \lambda_N b \end{pmatrix} \in R^{N n_B}$.

The parameter vector θ can be estimated using the RLS algorithm. It is described by the following equations:

$$\begin{cases} \hat{\theta}_k = \hat{\theta}_{k-1} + P_k\, \Phi_k\, \varepsilon_k \\ P_k = P_{k-1} - \dfrac{P_{k-1}\, \Phi_k\, \Phi_k^T\, P_{k-1}}{1 + \Phi_k^T\, P_{k-1}\, \Phi_k} \\ \varepsilon_k = y_k - \hat{\theta}_{k-1}^T\, \Phi_k \end{cases} \quad (4)$$

where P_k is the adaptation gain matrix, Φ_k is the observation vector and θ_k is the parameters vector.

2.2 Parametric identification of MIMO Hammerstein model

Two structures are used to identify MIMO Hammerstein models in the literature: with separate nonlinearities [21] or with combined nonlinearities [23]. The second case is the most general, but it can cause a very challenging parameter estimation problem because of the large number of parameters to be estimated.

In this paper, we developed a new coupled structure for MIMO Hammerstein model with separate nonlinearities. They are presented in Figure 2 and 4 where u_j, $v_{i,j}$, y_i for $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, m$ are the system input, internal signal and system output. $F_{i,j}(.)$ are nonlinear function.

2.2.1 A decoupled structure identification of MIMO Hammerstein model

The decoupled structure of MIMO Hammerstein model is given in Figure 2. Each output $y_{i,k}$, $i = 1, 2, \ldots p$, of the multivariable system corresponds to a linear model, which at its input are introduced a nonlinear functions, Figure 3.

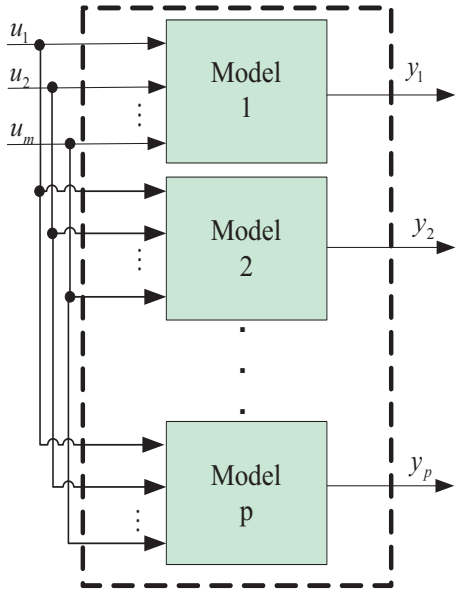


Figure 2: Decoupled structure of MIMO Hammerstein model

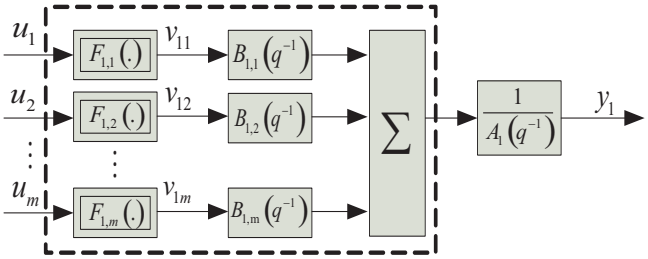


Figure 3: Structure of first model

Each output $y_{i,k}$, $i = 1, 2, \ldots p$, of the MIMO Hammerstein model is proposed as:

$$\begin{cases} A_i\left(q^{-1}\right)y_{i,k} = \sum_{j=1}^{m} B_{i,j}\left(q^{-1}\right)v_{j,k} \\ v_{i,j,k} = F_{i,j}\left(u_{j,k}\right) = u_{j,k} + \sum_{\rho=2}^{N} \lambda_{i,j,\rho}\, u_{j,k}^{\rho} \end{cases} \quad (5)$$

with:

$A_i\left(q^{-1}\right) = 1 + a_{i,1}q^{-1} + a_{i,2}q^{-2} + \ldots + a_{i,n_{A_i}}q^{-n_{A_i}}$

$B_{i,j}\left(q^{-1}\right) = b_{i,j,1}q^{-1} + b_{i,j,2}q^{-2} + \ldots + b_{i,j,n_{B_{i,j}}}q^{-n_{B_{i,j}}}$

System (5) can be rewritten as:

$$\begin{aligned} y_{i,k} = &-\sum_{\tau=1}^{n_{A_i}} a_{i,\tau}\, y_{i,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}\, u_{j,k-\tau} \\ &+ \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}}\sum_{\rho=2}^{N} b_{i,j,\tau}\,\lambda_{i,j,\rho}\, u_{j,k-\tau}^{\rho} \end{aligned} \quad (6)$$

then:

$$\begin{aligned} y_{i,k} = &-\sum_{\tau=1}^{n_{A_i}} a_{i,\tau}\, y_{i,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}\, u_{j,k-\tau} \\ &+ \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}}\sum_{\rho=2}^{N} s_{i,j,\tau,\rho}\, u_{j,k-\tau}^{\rho} \end{aligned} \quad (7)$$

Equation (7) can be written in the following form:

$$y_{i,k} = {\Phi_{i,k}}^{T}\ \theta_i \quad (8)$$

with:

$$\Phi_{i,k} = \begin{pmatrix} Y_{i,k} \\ U_k \\ \varphi_k \end{pmatrix},\ \theta_i = \begin{pmatrix} A_i \\ B_i \\ S_i \end{pmatrix}$$

$\Phi_{i,k}$ and $\theta_i \in R^{n_R}$; $n_R = n_{A_i} + \sum_{j=1}^{m} N n_{B_{i,j}}$

$$Y_{i,k} = \left(-y_{i,k-1},\ -y_{i,k-2},\ \ldots,\ -y_{i,k-n_{A_i}}\right) \in R^{n_{A_i}}$$

$$U_k = (U_{1,k},\ U_{2,k},\ \ldots,\ U_{m,k}) \in R^{m\times n_{B_{i,j}}}$$

$$U_{j,k} = \left(u_{j,k-1},\ u_{j,k-2},\ \ldots,\ u_{j,k-n_{B_{i,j}}}\right) \in R^{n_{B_{i,j}}}$$

$$\varphi_k = (\varphi_{1,k},\ \varphi_{2,k},\ \ldots,\ \varphi_{m,k}) \in R^{(N-1)\times m\times n_{B_{i,j}}}$$

$$\varphi_{j,k} = \left(\varphi_{j,1,k},\ \varphi_{j,2,k},\ \ldots,\ \varphi_{j,n_{B_{i,j}},k}\right) \in R^{(N-1)\times n_{B_{i,j}}}$$

$$\varphi_{j,\tau,k} = \left(u_{j,k-\tau}^{2},\ u_{j,k-\tau}^{3},\ \ldots,\ u_{j,k-\tau}^{N}\right) \in R^{(N-1)}$$

$$A_i = \left(a_{i,1},\ a_{i,2},\ \ldots,\ a_{i,n_{A_i}}\right) \in R^{n_{A_i}}$$

$$B_i = (B_{i,1},\ B_{i,2},\ \ldots,\ B_{i,m}) \in R^{m\times n_{B_{i,j}}}$$

$$B_{i,j} = \left(b_{i,j,1},\ b_{i,j,2},\ \ldots,\ b_{i,j,n_{B_{i,j}}}\right) \in R^{n_{B_{i,j}}}$$

$$S_i = (S_{i,1},\ S_{i,2},\ \ldots,\ S_{i,m}) \in R^{(N-1)\times m\times n_{B_{i,j}}}$$

$$S_{i,j} = \left(S_{i,j,1},\ S_{i,j,2},\ \ldots,\ S_{i,j,n_{B_{i,j}}}\right) \in R^{(N-1)\times n_{B_{i,j}}}$$

$$S_{i,j,\tau} = \left(s_{i,j,\tau,2},\ s_{i,j,\tau,3},\ \ldots,\ s_{i,j,\tau,N}\right) \in R^{(N-1)}$$

$$s_{i,j,\tau,\rho} = b_{i,j,\tau}\,\lambda_{j,\rho}$$

for $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, m$, and $\tau = 1, 2, \ldots, n_{B_{i,j}}$

The steps of the identification scheme are summarized as follows:

1. choosing an initial values for the adaptation matrix,
2. acquiring the input and output of the system and form the vector data as shown in (11) using the present and past values of the input u, output y, and u power,
3. solving the estimate parameter A_i, B_i and S_i using the algorithm RLS,
4. solving $\lambda_{i,j,\rho}$ using the estimated values $b_{i,j,\tau}$ and $s_{i,j,\tau,\rho}$ as:

$$\lambda_{i,j,\rho} = \left(\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}^{2}\right)^{-1}\left(\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}\, s_{i,j,\tau,\rho}\right) \quad (9)$$

2.2.2 A new coupled structure identification of MIMO Hammerstein system

The structure of this method is given in Figure 5. Each output of the system is depended on inputs and all other system outputs, Figure 4.

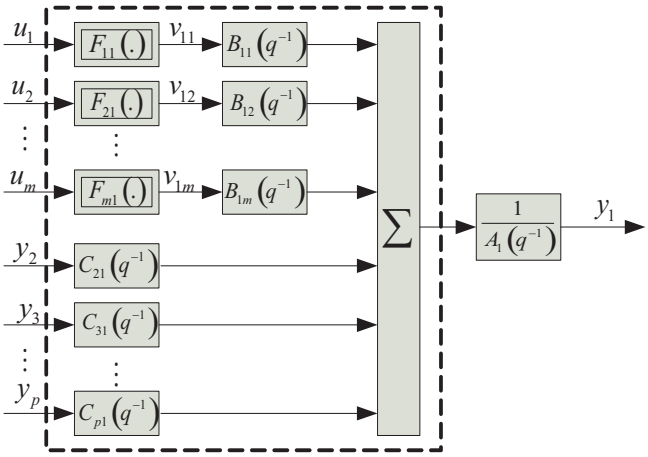


Figure 4: Structure of first model

Each output $y_{i,k}$, i = 1, 2, ...p, of the MIMO Hammerstein model is described by:

$$\begin{cases} A_i\left(q^{-1}\right)y_{i,k} = \sum_{j=1}^{m} B_{i,j}\left(q^{-1}\right)v_{j,k} + \sum_{\substack{l=1 \\ l\neq i}}^{p} C_{i,l}\left(q^{-1}\right)y_{l,k} \\ v_{i,j,k} = F_{i,j}\left(u_{j,k}\right) = u_{j,k} + \sum_{\rho=2}^{N} \lambda_{i,j,\rho}\, u_{j,k}^{\rho} \end{cases} \quad (10)$$

with:

$$A_i\left(q^{-1}\right) = 1 + a_{i,1}q^{-1} + a_{i,2}q^{-2} + \ldots + a_{i,n_{A_i}}q^{-n_{A_i}}$$
$$B_{i,j}\left(q^{-1}\right) = b_{i,j,1}q^{-1} + b_{i,j,2}q^{-2} + \ldots + b_{i,j,n_{B_{i,j}}}q^{-n_{B_{i,j}}}$$
$$C_{i,l}\left(q^{-1}\right) = c_{i,l,1}q^{-1} + c_{i,l,2}q^{-2} + \ldots + c_{i,l,n_{C_{i,l}}}q^{-n_{C_{i,l}}}$$

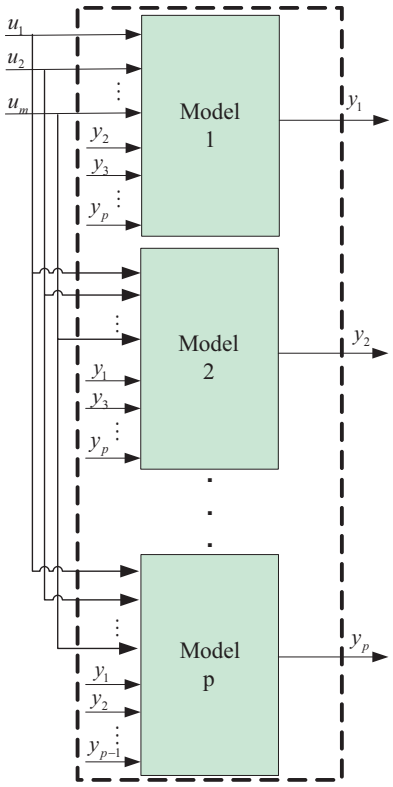


Figure 5: A proposed structure of MIMO Hammerstein model

System (10) can be rewritten as:

$$y_{i,k} = -\sum_{\tau=1}^{n_{A_i}} a_{i,\tau}\, y_{i,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}\, u_{j,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}}\sum_{\rho=2}^{N} b_{i,j,\tau}\,\lambda_{i,j,\rho}\, u_{j,k-\tau}^{\rho} + \sum_{\substack{l=1\\ l\neq i}}^{p}\sum_{\tau=1}^{n_{C_{i,l}}} c_{i,l,\tau}\, y_{l,k-\tau} \tag{11}$$

then:

$$y_{i,k} = -\sum_{\tau=1}^{n_{A_i}} a_{i,\tau}\, y_{i,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}} b_{i,j,\tau}\, u_{j,k-\tau} + \sum_{j=1}^{m}\sum_{\tau=1}^{n_{B_{i,j}}}\sum_{\rho=2}^{N} s_{i,j,\tau,\rho}\, u_{j,k-\tau}^{n} + \sum_{\substack{l=1\\ l\neq i}}^{p}\sum_{\tau=1}^{n_{C_{i,l}}} c_{i,l,\tau}\, y_{l,k-\tau} \tag{12}$$

Equation (12) can be written in the following form:

$$y_{i,k} = \Psi_{i,k}^{T}\theta_i^{new} \tag{13}$$

with:

$$\Psi_{i,k} = \begin{pmatrix} Y_{i,k} \\ U_k \\ \varphi_k \\ Y_{L,k} \end{pmatrix},\quad \theta_i^{new} = \begin{pmatrix} A_i \\ B_i \\ S_i \\ C_i \end{pmatrix}$$

$$\Phi_{i,k} \text{ and } \theta_i \in R^{n_R};\ n_R = n_{A_i} + \sum_{j=1}^{m} N n_{B_{i,j}} + n_{C_{i,l}} \times (p-1)$$

$$Y_{\substack{L,k\\ L\neq i}} = \left(Y_{1,k},\ Y_{2,k},\ \ldots, Y_{p,k}\right) \in R^{(p-1)\times n_{C_{i,l}}}$$

$$Y_{l,k} = \left(-y_{l,k-1},\ -y_{l,k-2},\ \ldots,\ -y_{l,k-n_{C_{i,l}}}\right) \in R^{n_{C_{i,l}}}$$

$$C_i = \left(C_{i,1},\ C_{i,2},\ \ldots,\ C_{i,p}\right) \in R^{(p-1)\times n_{C_{i,l}}}$$

$$C_{i,j} = \left(c_{i,j,1},\ c_{i,j,2},\ \ldots,\ c_{i,j,n_{C_{i,l}}}\right) \in R^{n_{C_{i,l}}}$$

for $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, m$, and $\tau = 1, 2, \ldots, n_{B_{i,j}}$, and $l = 1, 2, \ldots, p$, such as $l \neq i$.

Parametric identification consists to determine the parameters of the system based on vectors of inputs and outputs, using the recursive least squares algorithm.

3 Application: quadruple-tank process

A quadruple-tank process is used to illustrate the performance of the proposed structures of the MIMO Hammerstein model.

3.1 System description

The system setup is a model of a chemical plant fragment. Very often tanks are coupled through pipes and the reactant level and flow has to be controlled. The type of the experiments was performed on the 33-041 Coupled Tanks System of Feedback Instruments [24]. This plant, a variant of the quadruple tank process originally proposed in [25], is a model of a fragment of a chemical plant, Figure 6.

Figure 6: Plant of four coupled tanks [26]

The line diagram of the reservoir system is shown in Figure 7. The coupled tanks unit consists of four tanks placed on a rig. Another reservoir tank is placed at the bottom. In the reservoir two submersible pumps are placed, which pump the water on command to the tanks. The water flows freely to the bottom tanks through the configurable orifice. The way

the water flows through the setup can be configured in many ways with manual valves label led (MVA, MVB, MVC, MVD, MVE, MVF, MVG, MV1, MV2, MV3 and MV4).

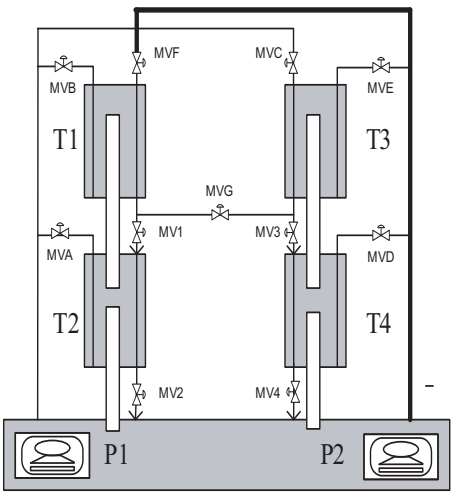


Figure 7: Line diagram for quadruple-tank process

3.2 Mathematical modelling

The quadruple-tank process admits the following nonlinear model [25] which has been assembled in simulink:

$$\begin{cases} \dot{h}_1(t) = \eta v_1(t) - \frac{a_1}{A}\sqrt{2gh_1(t)} - \frac{a_{13}}{A}\sqrt{2g(h_1(t)-h_3(t))} \\ \dot{h}_2(t) = \frac{a_1}{A}\sqrt{2gh_1(t)} - \frac{a_2}{A}\sqrt{2gh_2(t)} \\ \dot{h}_3(t) = \eta v_2(t) - \frac{a_3}{A}\sqrt{2gh_3(t)} - \frac{a_{13}}{A}\sqrt{2g(h_1(t)-h_3(t))} \\ \dot{h}_4(t) = \frac{a_3}{A}\sqrt{2gh_3(t)} - \frac{a_4}{A}\sqrt{2gh_4(t)} \end{cases} \quad (14)$$

where h_i, for $i = 1,2,3,4$, denote the water level in the corresponding tank and v_i, for $i = 1,2$, are voltages applied to the pumps. a_i, for $i = 1,2,3,4$, are the outlet area of the tanks, a_{13} is the outlet area betwixt tanks 1 and 3; η constant relating the control voltage with the water flow from the pump, A is the cross-sectional area of the tanks and g is the gravitational constant.

3.3 Simulation and results

The proposed structures has been tested with the model parameters presented in table 1.

	Value	Unit	Description
h_i	$0-25$	cm	Water level of tank i
v_i	$0-5$	V	Voltage level of pump i
S	0.014	m^2	Cross-sectional area
a_i	$5e-5$	m^2	Outlet area of tank i
a_{13}	$5e-5$	m^2	Outlet area betwixt $T1$ and $T3$
η	$2e-3$	$\frac{m^3}{V.s}$	Water level of tank i
g	9.81	m^{-2}	Gravitational constant

Table 1: Parameters of the plant

The inputs u_1 and u_2 are show in Figure 8. They are set to $[0\ldots+5V]$. The sample time used for all simulations is $T_e = 1s$.

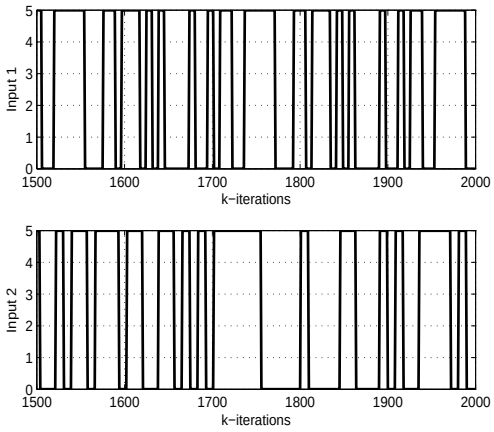


Figure 8: Inputs u_1 and u_2

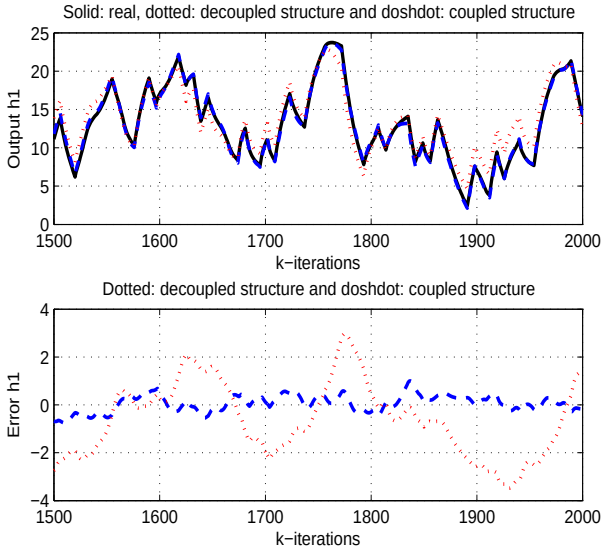


Figure 9: Response of the real and estimated h_1

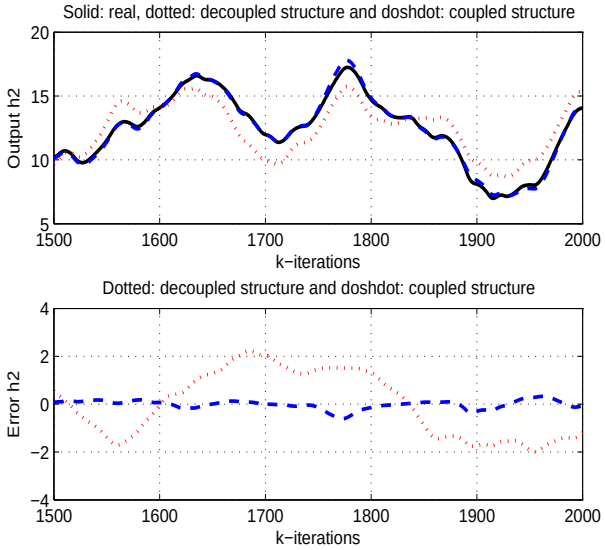


Figure 10: Response of the real and estimated h_2

Figures 9 to 12 show the results of the experiments with two method from a comparative point of view. Each of them shows a superposition of the actual output and the two outputs estimated as their two error curves. Solid line represents the real output, dotted line represents the estimated systems output signals with the decoupled structure and the dash dot lines represent the estimated systems output signals with the coupled structure in all four graphs. The responses of the original system and the results of the proposed method are very similar. It is clear that the error corresponds to the proposed structure is smaller than that corresponding to the decoupled structure. On inspection, the proposed structure is seen to work better than the decoupled structure. Thus, the feasibility and superiority of this proposed identification method are validated.

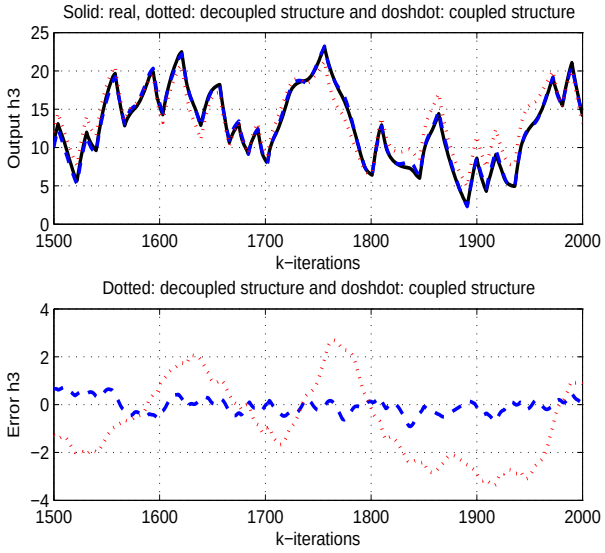


Figure 11: Response of the real and estimated h_3

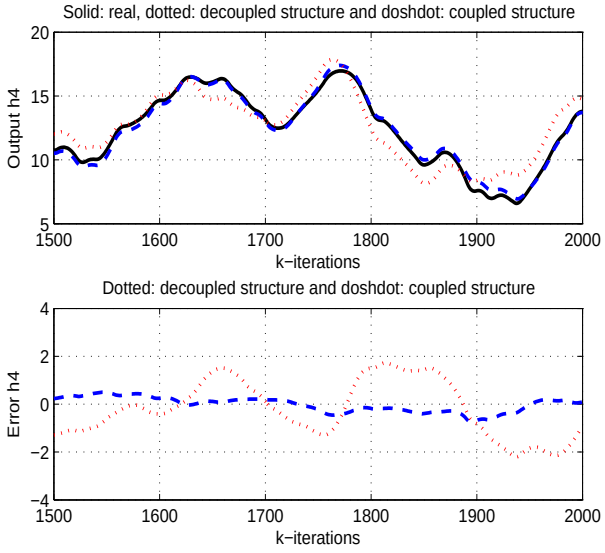


Figure 12: Response of the real and estimated h_4

4 Conclusion

In this study, a new coupled structure identification of MIMO Hammerstein model with separate nonlinearities has been developed. The method is based on RLS algorithm. A comparative study using simulation results for quadruple-tank process, between the proposed structure and the decoupled structure is discussed. Simulation results reveal the performance and effectiveness of the proposed method.
As a perspective, we will judge the performance of the method used in the presence of perturbations

References

1. Z. Rayouf, C. Ghorbel, and N. B. Braiek, "An extended identification methods for multivariable Hammerstein system" in 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, Hammamet Tunisia, 2016. https://doi.org/10.1109/SETIT.2016.7939921
2. Z. O. U. Zhiyun, Y. U. Meng, W. A. N. G. Zhizhen, L. I. U. Xinghong, G. U. O. Yuqing, F. ZHANG, and G. U. O. Ning, "Nonlinear model algorithmic control of a pH neutralization process" Chinese Journal of Chemical Engineering, **21**(4), 395–400, 2013. https://doi.org/10.1016/S1004-9541(13)60479-6
3. D. Z. Li, Y. X. Jia, Q. S. Li, and Q. B. Jin, "Identification and nonlinear model predictive control of MIMO Hammerstein system with constraints" Journal of Central South University, **24**(2), 448-458, 2017. https://doi.org/10.1007/s11771-017-3447-3
4. R. Kanthasamy, H. Anwaruddin, and S. K. Sinnadurai, "A new approach to the identification of distillation column based on hammerstein model" Modelling and Simulation in Engineering, **2014**(11), 1-7, 2014. http://dx.doi.org/10.1155/2014/813757
5. D. K. Rollins, N. Bhandari, A. M. Bassily, G. M. Colver, and S. T. Chin, "A continuous-time nonlinear dynamic predictive modeling method for hammerstein processes" Industrial and engineering chemistry research, **42**(4), 860-872, 2003. https://doi.org/10.1021/ie020169
6. D. Wang, and F. Ding, " Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems" Computers and Mathematics with Applications, **56**(12), 3157–3164, 2008. https://doi.org/10.1016/j.camwa.2008.07.015
7. F. Ding, "Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling" Applied Mathematical Modelling, **37**(4), 1694–1704, 2013. https://doi.org/10.1016/j.apm.2012.04.039
8. R. Su, B. Wang, and S. Liu, "The iterative identification method for hammerstein nonlinear channel" in 10th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), Beijing, China, 2014. https://doi.org/10.1049/ic.2014.0075
9. J. Wang, Q. Zhang, and L. Ljung. "Revisiting Hammerstein system identification through the two-stage algorithm for bilinear parameter estimation" Automatica, **45**(11), 2627-2633, 2009. https://doi.org/10.1016/j.automatica.2009.07.033
10. D. Westwick, and R. Kearney, "Separate least squares identification of nonlinear Hammerstein models: application to stretch reflex dynamics" Annals of Biomedical Engineering, **29**(8), 707-718, 2001. https://doi.org/10.1114/1.1385806
11. E. Bai, and M. Fu, "Identification of linear systems with hard input nonlinearities of known structure" IEEE Transactions on Signal Processing, **38**(5), 853-860, 2002. https://doi.org/10.1016/S0005-1098(01)00281-3

12. E. Bai, and M. Fu, "A blind approach to Hammerstein model identification" IEEE Transactions on Signal Processing, **50**(7), 1610-1619, 2002. https://doi.org/10.1109/TSP.2002.1011202

13. E. Bai, "Frequency domain identification of Hammerstein models" IEEE Transactions on Automatic Control, **48**(4), 530-542, 2003. https://doi.org/10.1109/TAC.2003.809803

14. Y. Liu, J. Sheng, and R. Ding, "Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems" Computers and Mathematics with Applications, **59**(8), 2615-2627, 2010. https://doi.org/10.1016/j.camwa.2010.01.030

15. J. Abonyi, R. Babuška, M. A. Botto, F. Szeifert, and L. Nagy, "Identification and control of nonlinear systems using fuzzy Hammerstein models" Industrial and engineering chemistry research, **39**(11), 4302–4314, 2000. https://doi.org/10.1021/ie990629e

16. L. Jia, M. S. Chiu, and S. S. Ge, "A noniterative neuro-fuzzy based identification method for Hammerstein processes" Journal of Process Control, **15**(7), 749-761, 2005. https://doi.org/10.1016/j.jprocont.2005.03.006

17. A. Jawad, A. M. Butt, M. T. Riaz1, S. Bhutta1, M. Z. khan1, and I. U. Haq "Multiclass Myoelectric Identification of Five Fingers Motion using Artificial Neural Network and Support Vector Machine" Journal of Advances in Science, Technology and Engineering Systems, **2**(3), 1026–1033, 2017. https://doi.org/10.25046/aj0203130

18. K. H. Chan, J. Bao, and W. J. Whiten, "Identification of MIMO Hammerstein systems using cardinal spline functions" Journal of Process Control, **16**(7), 659-670, 2006. https://doi.org/10.1016/j.jprocont.2006.01.004

19. I. Goethals, K. Pelckmans, J. A. Suykens, and B. De Moor, "Identification of MIMO Hammerstein models using least squares support vector machines" Automatica, **41**(7), 1263–1272, 2005. https://doi.org/10.1016/j.automatica.2005.02.002

20. H. Liang, and B. Wang, "Identification of MIMO Hammerstein Models Using Support Vector Machine" In International Symposium on Neural Networks, Springer, Berlin, Heidelberg, 2007. https://doi.org/10.1007/978-3-540-72395-0-51

21. Y. M. Hlaing, M. S. Chiu, and S. Lakshminarayanan"Modeling and control of multivariable process using generalized Hammerstein model" Chemical Engineering Research and Design Transactions, **85**(4), 445-454, 2007. https://doi.org/10.1205/cherd06210

22. F. Ding, and T. Chen, "Identification of Hammerstein nonlinear ARMAX system" IEEE Transactions Automatic Control, **41**(9), 1479-1489, 2005. https://doi.org/10.1016/j.automatica.2005.03.026

23. K. H.Chan, J. Bao, and W. J. Whiten, "Identification of MIMO Hammerstein systems using cardinal spline functions" Journal of Process Control, **16**(7), 659–670, 2006. https://doi.org/10.1016/j.jprocont.2006.01.004
http://dx.doi.org/10.1080/00207721.2014.951419

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 63-69 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

ARMA Feeding Techniques for Isoflux Coverage from a Micro Satellite

Ali Siblini [1,2,*], Hussein Abou Taam[2], Bernard Jecko[1], Mohammad Rammal[2], Anthony Bellion[3], Eric Arnaud[1]

[1]*Systems-RF University of Limoges, 12, rue Gemini 87068 Limoges Cedex 03, France*

[2]*Lebanese University, Doctoral School of Science and Technology, Beirut*

[3]*CNES, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 9, France*

ARTICLE INFO

Article history:
Received : 26 September 2017
Accepted : 18 October 2017
Online : 04 November 2017

Keywords :
Agile Radiating Matrix antenna
Multimode radiating beams
Spatial Telemetry

ABSTRACT

This paper deals with the design of a new reconfigurable beam antenna used to improve the efficiency of spatial telemetry links on Nano-Satellite. This agile beam antenna is not built on the well-known array concept AESA (Agile Electronically Scanned Array) but using a new one called ARMA (Agile Matrix Radiating Antenna); MARPEM in French. In this paper there is the design of the circularly polarized matrix antenna, the generation of two beam forming modes, the design of the distribution circuit and the polarization circuits.

1. Introduction

The needs in high data rate of future Nano-satellite missions are incompatible with telemetry applications in VHF or S bands unlike the X-band. Moreover, most of applications require earth wide angle coverage in addition to a long time visibility, implying a theoretical isoflux radiation pattern particularly for LEO satellites. But this theoretical limited gain radiation pattern in nadir direction reduces the global efficiency of the PDTM (Payload Data Telemetry). A solution to maximize the PDTM is to use an agile beam antenna to increase the data rate efficiency by introducing several kinds of radiation patterns depending on the satellite position.

These links are produced by beams of various shapes delivered by an agile antenna working on several radiation modes. To limit the complexity of the architecture and to be accommodated on a Nano-satellite, these requirements lead to a dual mode circularly polarized antenna: one mode characterized by isoflux earth coverage and another one introducing high gain directive beam around the axial direction (nadir). To realize this PDTM link in X-band, a very compact antenna using dual mode radiation pattern has been developed with the following requirements:

- [8 – 8.4]GHz.
- To be accommodated on a 100mm x 100mm surface.
- Gain > 2dB for θ [-60°: 60°] and φ [-180°:180°].
- Axial Ratio < 4dB for θ [-60°:60°] and φ [-180°:180°].

The beam forming procedure and the switching process between two or more working modes need to use agile beam antennas. There are two kinds of electronically agile beam antennas; the first one is the well-known AESA (Agile Electronically Scanned Array) technique and the second, a new one, called ARMA (Agile Radiating Matrix Antenna) which introduces some advantages for such applications [1], [2] as larger angles beam scanning, higher power balance, side lobes minimization [3]-[6], lower couplings between adjacent elements and high gain solutions [7], [2] with smaller number of elements and less feeding network complexity.

In the context of this paper, the new ARMA concept is defined and the circularly polarized pixel with the polarization circuits design is presented. Then the ARMA matrix with circular polarization is used to generate two beams, the wide Isoflux and the high gain axial beams. Beam forming law is optimized to solve

*Corresponding Author: Ali Siblini, Lebanese University Beirut Lebanon.
Email: ali.siblini.84@hotmailcom

the problem of Isoflux dissymmetry in the different azimuthal planes,

2. The ARMA Concept

ARMA principle introduces new sampling procedure for the radiating surface of the antenna that is sampled or pixelated using the pixel. Brief comparison of AESA and ARMA principles is found in [3]. ARMA approach uses jointed pixels as shown in Figure 1 to sample the radiating surfaces instead a lot of punctual sources for the array technique [4]. Each pixel generates a uniform radiating surface that never overlaps outside its dimensions [3].

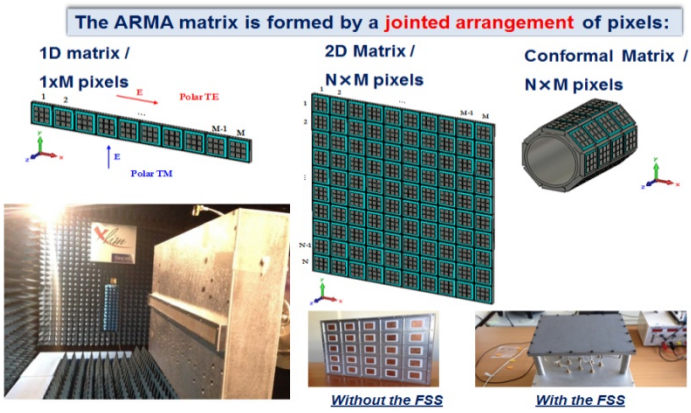


Figure 1. ARMA antennas architectures.

The most important point in ARMA approach is to be able to design a pixel which can generate a uniform *EM* field on its surface. This technique, deduced from the knowledge of EBG antennas properties, was extensively developed and validated in many papers [8]-[11] for linear polarization; it is the key point of the ARMA solution. Example of a pixel with one feeding probe is shown in Figure 2 with the corresponding radiating surface on the top.

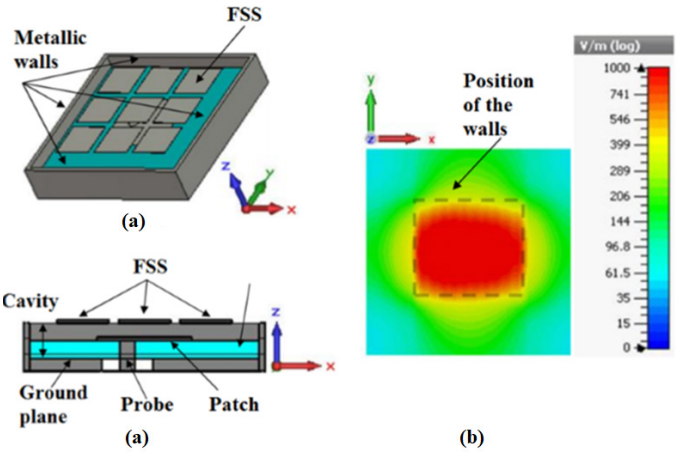


Figure 2. Pixel design: (a) FSS on top, patch inside the cavity and metallic walls, (b) the corresponding radiating surface.

For low profile antennas, a planar radiating surface *S* is built with pixels with a square radiating surfaces as shown in Figure 3.

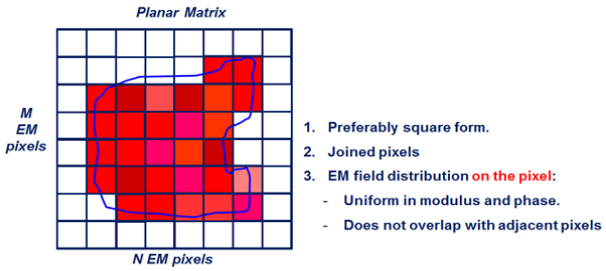


Figure 3. Radiating surface sampled by jointed square pixels.

3. Circularly Polarized Pixel Design

3.1. Pixel Design and Axial Ratio

For circular polarization (two orthogonal linear polarizations), the EBG cavity is usually fed by a patch probe generating a local circularly polarized field in the cavity. A good circular polarization can be obtained by two or four ports feeding circuit, shown in Figure 4. Of course, for a pixel considered alone, the axial ratio is better [12] [13] with the four ports technique (Figure 5 and Figure 6), but with higher manufacturing complexity and price.

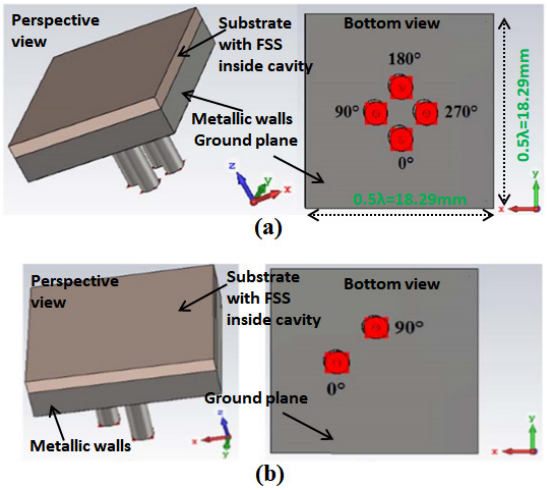


Figure 4. Pixel with two and four feeding ports for circular polarization generation.

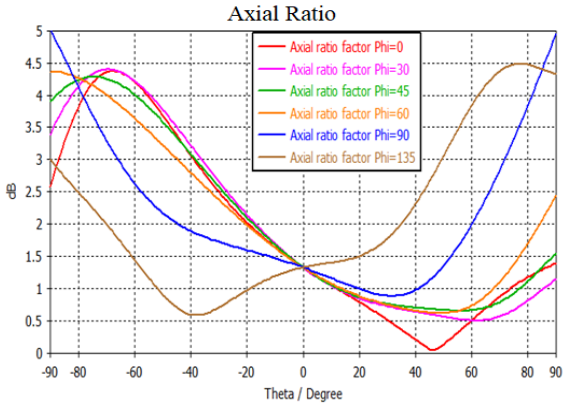


Figure 5. Axial ratio comparison obtained with a 2 ports ideal feeding procedure.

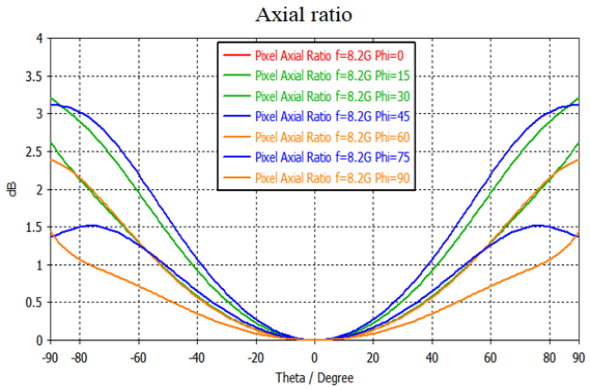


Figure 6. Axial ratio comparison obtained with a 4 ports ideal feeding procedure.

3.2. Polarization Circuits and Area Limitations

The main limitation in the design of the polarization circuit is the small area behind the pixel which is 18.29mmx18.29mm (0.5λx0.5) (Figure 4). This small area must include also the pad to solder the connector for the input power. For that the substrate TMM10i with permittivity 9.8 and thickness 0.381mm was used to print the circuit on it. The results of the axial ratio are not very modified by introducing the circuit to the pixel design: one hybrid coupler for the two ports case and 3 hybrid couplers for the 4 ports case as shown in Figure 7 and Figure 8.

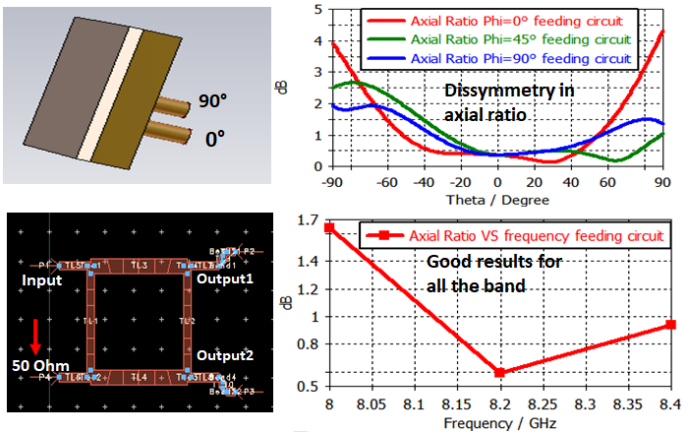


Figure 7. Axial ratios obtained with a one hybrid coupler for the two ports.

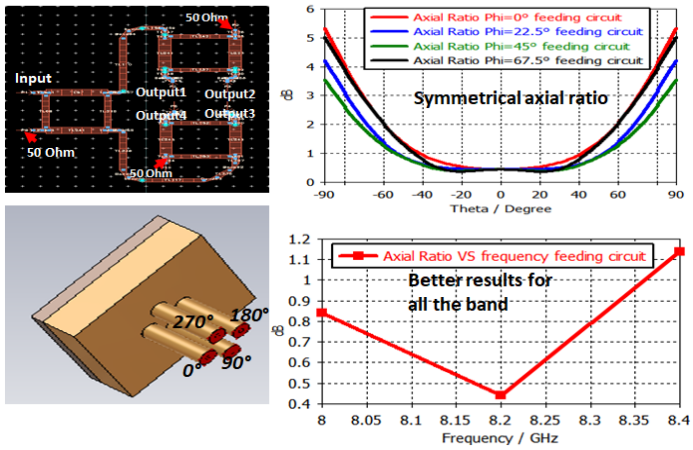


Figure 8. Axial ratios obtained with a d a 3 hybrid couplers for the 4 ports.

4. PDTM Application

4.1. Antenna Design

On this way, a square low profile ARMA antenna was built with 25 pixels where only nine are fed (Figure 9) by the BFN (Beam Forming Network); the other ones, located at the edges of the antenna, are only used to keep a good axial ratio for the whole matrix antenna.

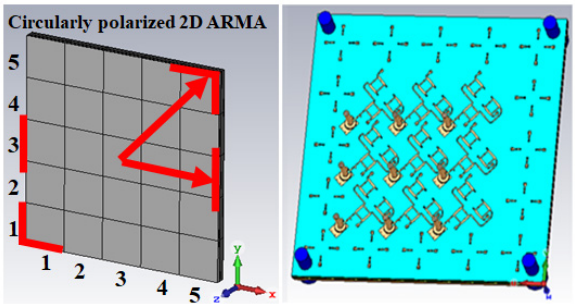


Figure 9. 25 pixels matrix antenna (left) with nine feeding ports and nine circular polarization circuits (Right).

4.2. Isoflux Pattern and Dissymmetry Problem

The isoflux coverage (first mode) is obtained by feeding the pixels with a weighting law deduced from a Sinc (sinus cardinal) function [14] [15]. This approach leads to weighting tables and to the corresponding matrix radiation patterns and axial ratios shown in Figure 10 and Figure 11.

Table of amplitudes $A_{i,j}$

	1	2	3	4	5
1	0	0	0	0	0
2	0	0.03	0.18	0.03	0
3	0	0.18	1	0.18	0
4	0	0.03	0.18	0.03	0
5	0	0	0	0.002	0

Table of $\varphi_{i,j}$

	1	2	3	4	5
1	0	0	0	0	0
2	0	180°	180°	180°	0
3	0	180°	0°	180°	0
4	0	180°	180°	180°	0
5	0	0	0	0	0

Figure 10. Sinc circular feeding law.

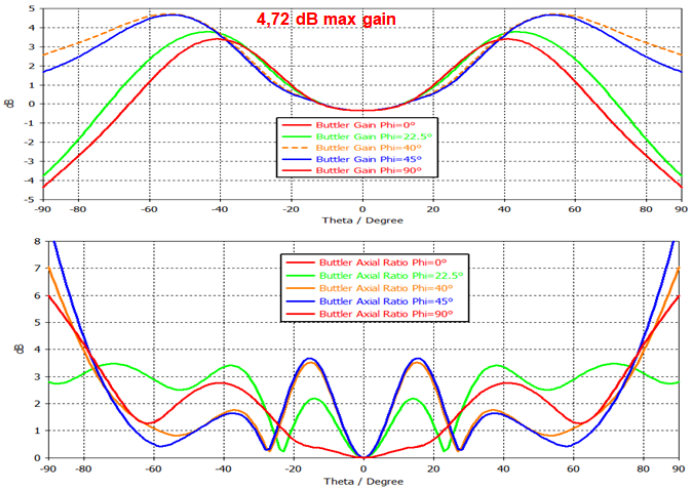


Figure 11. Gain an axial ratio patterns.

The pattern in Figure 11 shows a dissymmetry in the isoflux maximum gain positions (e.g. planes φ=0° and φ=45°); this is due to the non-symmetrical physical structure of the matrix antenna, where the distance from the center to the edges is not the same for all directions (Figure 9).

4.3. Feeding Law Optimization for Symmetrical isoflux

To compensate the dissymmetry resulted from the square matrix structure, the feeding law is optimized to modify the isoflux radiation pattern [15]. The value "A" is the weight applied on each pixel surrounding the central one, "A" is changed as shown in Figure 12, Figure 13 and Figure 14. This procedure, called Sinc-circular symmetrical, will be very useful to match easily the two modes together. It is also important to notice that the part of the following curves located between θ=±20° is not useful in the isoflux mode because for theses directions the matrix is working on the axial mode.

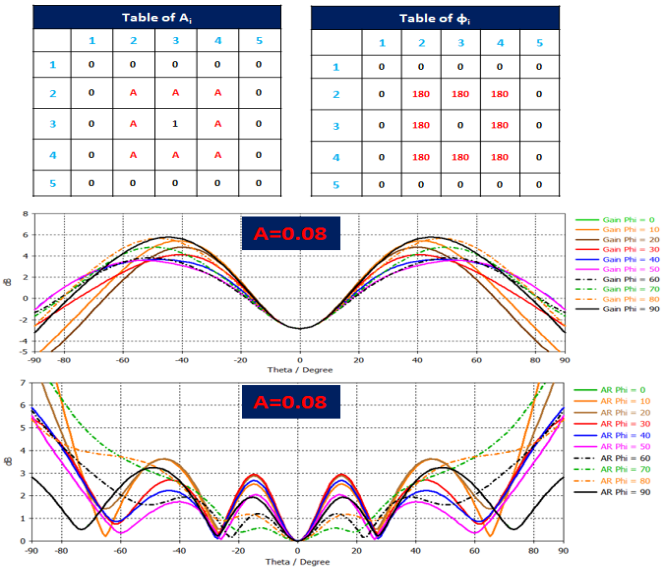


Figure 12. Gain and axial ratio evolution as a function of the elevation angle "θ" for different azimuth planes with A = 0.08.

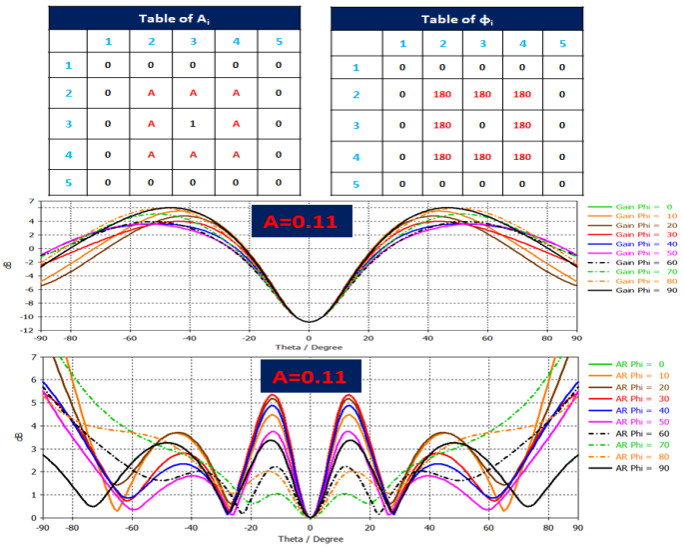


Figure 13. Gain and axial ratio evolution as a function of the elevation angle "θ" for different azimuth planes with A = 0.11.

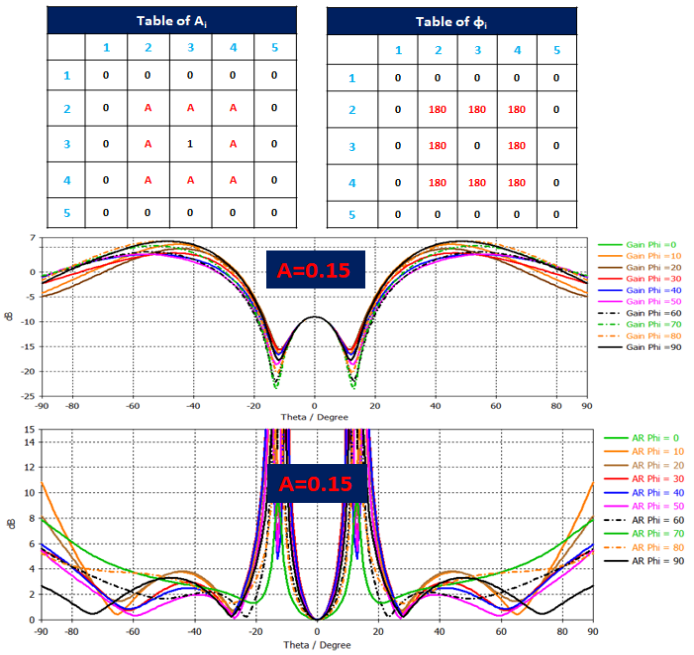


Figure 14. Gain and axial ratio evolution as a function of the elevation angle "θ" for different azimuth planes with A = 0.15.

Then with this symmetrical feeding solution the Isoflux gain pattern is more symmetrical with the position of the maximum gain located in the worst case at θ=±50° for the solution with A=0.15 (Figure 14). The axial ratio in this case is going very high for θ around ±20°, but these directions are not useful in this mode as mentioned above.

5. Directive Beam

The directive beam around the axial direction (second mode) is easy to obtain with the highest gain with all the weights equal to 1 (Figure 15). But, like for an array, it is also possible to optimize the radiation pattern and the axial ratio by modifying theses weights.

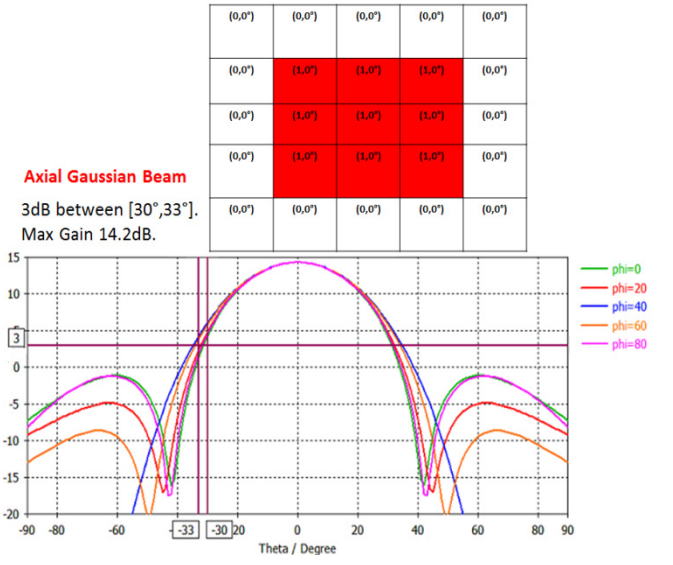


Figure 15. Radiation pattern of the axial mode as a function of the elevation angle θ for different azimuth planes φ.

6. Junction between Two Modes

The BFN (Beam Forming Network) architecture used to switch between the two modes is presented in Figure 16 for a 1D ARMA antenna.

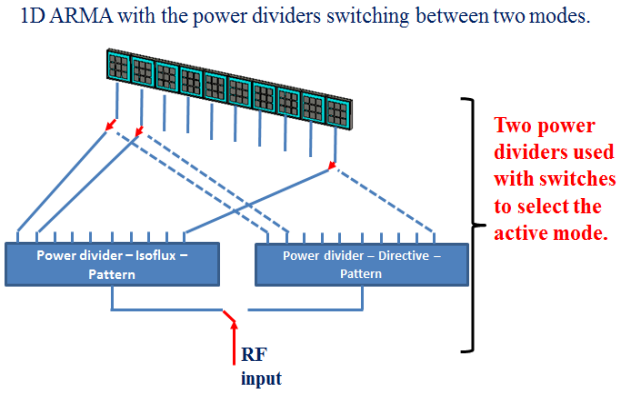


Figure 16. Switching procedure between the two modes.

The switching procedure [16] between the 2 modes can be easily performed for the elevation angle and the gain values expected by the users. For example it can be chosen to keep the gain above 3dB between -60° and +60° as shown in Figure 17.

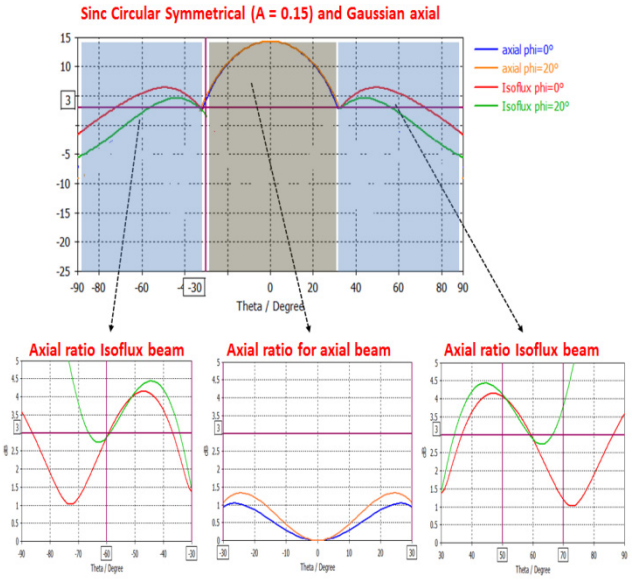


Figure 17. Two modes radiation patterns and axial ratios for the whole agile antenna for φ = 0° and φ = 20° planes.

7. Power Divider Design for Isoflux

The symmetrical feeding law with A=0.15 is considered for the Isoflux beam forming. To achieve this law a power divider was designed with one input and nine outputs as shown in Figure 18. The port labeled 5 is to be connected to the central pixel in the matrix antenna; the other output ports (1, 2, 3, 4, 6, 7, 8 and 9) will be connected to the other pixels around the central one. The return loss at the input of the power divider is shown in Figure 19. Using of the hybrid couplers at the outputs insure good matching (50Ω) when connecting the power divider to the polarization circuits that feed the patches inside the pixels cavities.

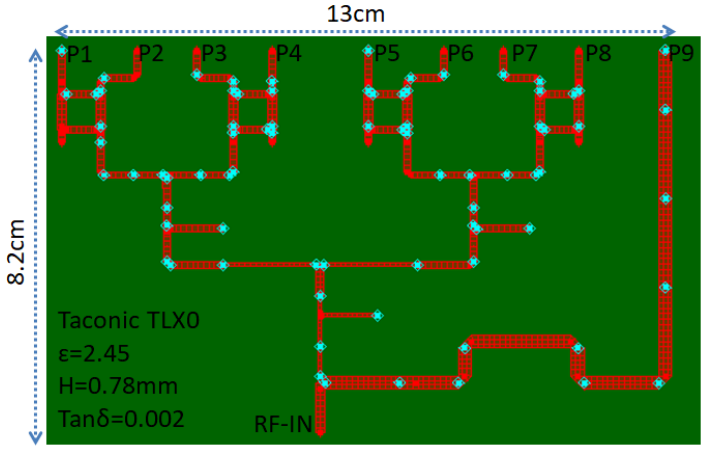


Figure 18. Power divider Momentum design for Isoflux beam forming.

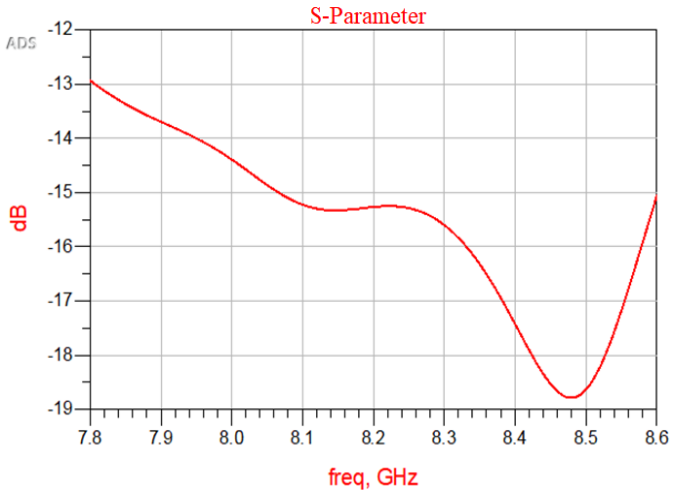


Figure 19. Return loss at the input of the power divider.

The output weights in magnitude and phase at 8.2GHz is presented in the table of Figure 20. These magnitudes are taken normalized with respect to the central port5 and the phases are the phase shift with the phase of port 5.

Port	P1	P2	P3	P4	P5	P6	P7	P8	P9
Magnitude	0.154	0.152	0.157	0.158	1	0.158	0.157	0.152	0.155
Phase	181°	182°	184°	181°	0°	181°	184°	182°	181°

Figure 20. Outputs of the power divider at 8.2GHz.

8. Pattern of the Whole Matrix

Polarization circuits and the power divider results files are taken from Momentum software and connected to the matrix antenna in CST (Figure 21). The realized gain and the axial ratio patterns are shown in Figure 22 at 8.2GHz. The position of the isoflux maximum is attained at approximately at fixed elevation angles and the difference in the gain between the different azimuth planes is 2.9dB.

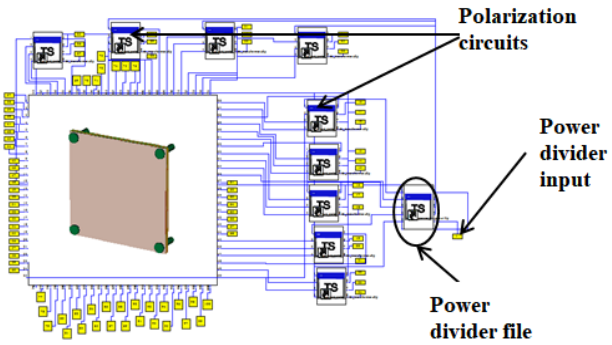


Figure 21. Matrix antenna with the polarization circuits and power divider.

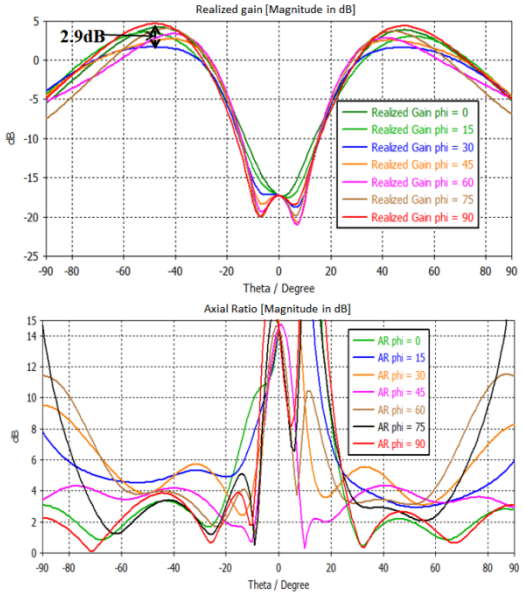


Figure 22. Realized gain and axial ratio patterns at 8.2GHz (whole antenna).

9. Testing the Antenna with the Nano-Satellite Support

The antenna is placed on the top of the Nano-satellite, so to check the effect of the satellite support the antenna was placed on the top of a metallic cube in CST (Figure 23) (cube with 130mm height and 100mm surface). In this case the antenna is tested as the real case for the satellite to be placed on it. Results of the antenna with and without the cube support show no significant modifications in the gain and axial ratio patterns (Figure 24).

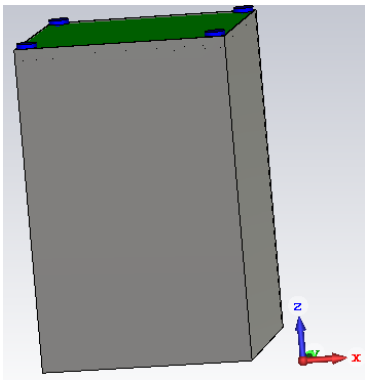


Figure 23. Antenna placed on the cube-sat support.

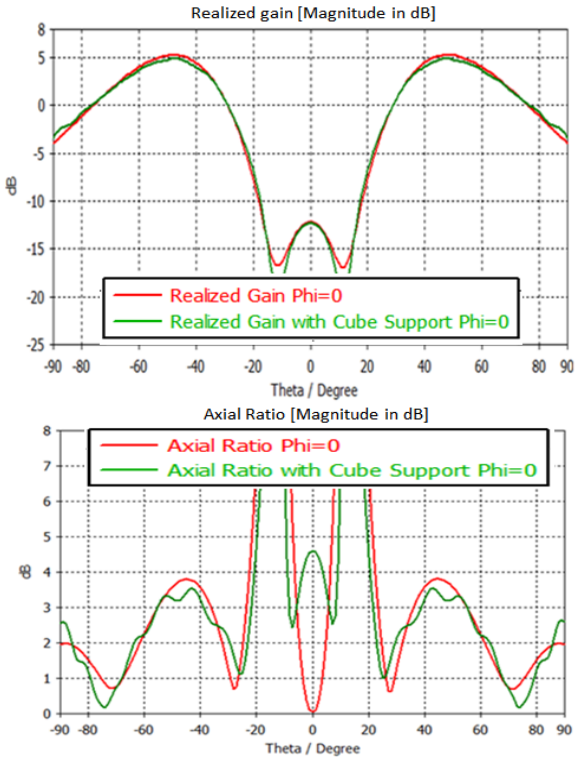


Figure 24. Realized gain and axial ratio patterns with and without cube-sat support.

10. Conclusion

A dual mode agile antenna based on the ARMA principle [1] [2] was presented in this paper. A prototype is under manufacturing and will be tested next month. The pixel with four ports with insures high stability of the axial ratio over wide elevation angles using the polarization circuit build by hybrid couplers. This pixel also insures the uniformity characteristics that are the basis of the new ARMA sampling procedure. The matrix ARMA was used to generate two different beams with and axial ratio for both below 3dB in their interesting regions. Also the antenna is tested on the Nano-satellite support with no problems.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work has been funded by the French Spatial Agency (CNES).

References

[1] Patent: JECKO, B.; HAJJ, M.; CHANTALAT, R.; SALAH TOUBET, «Antenne élémentaire et antenne réseau mono ou bidimensionnelle correspondante» PCT Patent: PCT/EP2012/076509; French Patent: FR 11 62141.

[2] ABOU TAAM, H., ZAKKA EL NASHEF, G., SALAH TOUBET, M., et al. Scan performance and re-configurability of agile radiating matrix antenna prototype. International Journal of Antennas and Propagation, 2015, vol. 2015.

[3] ABOU TAAM, Hussein, SALAH TOUBET, Moustapha, MONEDIERE, Thierry, et al. A new agile radiating system called electromagnetic band gap

matrix antenna. International Journal of Antennas and Propagation, 2014, vol. 2014.

[4] TAAM, Hussein Abou, EL NASHEF, Georges Zakka, ARNAUD, Eric, et al. Design development and experimental validation of an EBG matrix antenna for tracking application. International Journal of Microwave and Wireless Technologies, 2015, p. 1-9.

[5] TAAM, Hussein Abou, NASHEF, El, ZAKKA, Georges, et al. Feeding technique for the experimental validation of EBG matrix antenna. Microwave and Optical Technology Letters, 2015, vol. 57, no 3, p. 533-537.

[6] TAAM, Hussein Abou, TOUBET, Moutapha Salah, MONEDIERE, Thierry, et al. Interests of a 1D EBG Matrix compared to a patch array in terms of mutual coupling and grating lobe. In: Antennas and Propagation (EuCAP), 2013 7th European Conference on. IEEE, 2013. p. 1045-1048.

[7] TAAM, Hussein Abou, TOUBET, Moutapha Salah, MONEDIERE, Thierry, et al. Radiation control of an agile Matrix antenna by using specific algorithm. In: Microwave Symposium (MMS), 2013 13th Mediterranean. IEEE, 2013. p. 1-4.

[8] TAAM, Hussein Abou, EL NASHEF, Georges Zakka, JECKO, Bernard, et al. Agile radiating matrix antenna system for radar applications. In : Radar Conference (Radar), 2014 International. IEEE, 2014. p. 1-4.

[9] TAAM, Hussein Abou, SIBLINI, Ali, EL NASHEF, Georges Zakka, et al. An agile electronically scanned EBG matrix antenna for monitoring target activity. In: Microwave Conference (EuMC), 2015 European. IEEE, 2015. p. 1435-1438.

[10] TAAM, H. Abou, EL NASHEF, G. Zakka, ARNAUD, E., et al. Experimental validation of an agile electromagnetic band gap matrix antenna. In: Antennas and Propagation (EuCAP), 2015 9th European Conference on. IEEE, 2015. p. 1-4.

[11] B. Jecko, E. Aranud, H. Abou Taam, A. Siblini The ARMA concept: Comparison of AESA and ARMA technologies for agile antenna design. Fermat Journal vol.20, 2016

[12] SIBLINI, Ali, TAAM, Hussein Abou, JECKO, Bernard, Rammal, Mohammad. Pixel and Patch Comparison. SENSET 2017 Sensors, Networks, Smart and Emerging Technologies, Beirut Lebanon.

[13] SIBLINI, Ali, TAAM, Hussein Abou, JECKO, Bernard, et al. Wideband circularly polarized pixel antenna using special EBG approach: Building of agile antenna ARMA for space applications. In: Antennas and Propagation (MECAP), 2016 IEEE Middle East Conference on. IEEE, 2016. p. 1-4.

[14] SIBLINI, Ali, TAAM, Hussein Abou, JECKO, Bernard, Rammal, Mohammad. Pixel and Patch Comparison. New Circularly Polarized Matrix Antenna for Space Applications. Wireless testing symposiums 2016 London.

[15] SIBLINI, Ali, TAAM, Hussein Abou, JECKO, Bernard, et al. New agile EBG matrix antenna for space applications. In: Microwave Conference (EuMC), 2016 46th European. IEEE, 2016. p. 874-877.

[16] SIBLINI, Ali, JECKO, Bernard, et ARNAUD, Eric. Multimode reconfigurable Nano-satellite antenna for PDTM application. In: Antennas and Propagation (EUCAP), 2017 11th European Conference on. IEEE, 2017. p. 542-545.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 70-75 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Selective Electron Beam Melting Manufacturing of Electrically Small Antennas

Saad Mufti[*, 1], Christopher Smith[2], Alan Tennant[1], Luke Seed[1]

[1]*Department of Electronic and Electrical Engineering, University of Sheffield, S10 2TN, UK*

[2]*Department of Materials Science and Engineering, University of Sheffield, S10 2TN, UK*

A R T I C L E I N F O

Article history:
Received : 01 September, 2017
Accepted: 03 November, 2017
Online: 08 November, 2017

Keywords :
3D Printing
Additive Manufacturing
Conformal Antennas
Electrically Small Antennas
Inverted-F Antennas

A B S T R A C T

Real estate pressures in modern electronics have resulted in the need for electrically small antennas, which have subsequently garnered interest amongst researchers and industry alike. These antennas are characterized by their largest dimensions translating to a fraction of the operating wavelength; such a diminutive size comes at the expense of reduced gain and efficiency, and a worse overall match to a corresponding power source. In order to compensate for this deterioration in performance, antenna designers must turn towards increasingly complex and voluminous geometries, well beyond the capabilities of traditional manufacturing techniques. We present voluminous metal antennas, based on a novel inverted-F design, and fabricated using the emergent selective electron beam melting manufacturing technique, a type of powder bed fusion process. As predicted by small antenna theory, simulation results presented show in increase in the antenna's efficiency as it is voluminously expanded into the third dimension. Measurement results illustrate that key trends observed from simulations are upheld; however, further understanding of the electromagnetic properties of raw materials, in particular how these change during the printing process, is needed. Nevertheless, this type of additive manufacturing technique is suitable for rapid prototyping of novel and complex antenna geometries, and is a promising avenue for further research and maturation.

1. Introduction

With the continued trend of miniaturization in the design and manufacture of modern, 'smart' electronics, antenna designers are challenged with delivering compact, low-cost, and easy to fabricate antennas, often operating at wavelengths many times their largest dimension. A further key requirement is that these antennas are made efficient, so as to prolong battery life. This paper is an extension of work originally presented in the 2017 International Workshop on Antenna Technology in Athens, Greece [1]. The discourse on additively manufactured (AM) voluminous antennas is expanded to provide an overview of electrically small antennas (ESA) in literature fabricated using novel techniques, the three-dimensional (3D) powder bed fusion (PBF) printing process, the antenna design process, as well as the implications of simulated and measured results on practical antenna design.

Formally, an antenna is classed as electrically small if the product of the wavenumber, k (= $2\pi/\lambda$), and the radius of an imaginary sphere circumscribing the maximum dimension of the antenna, a, is less than or equal to 0.5; note that the parameter ka is unit-less [2-4]. It is widely accepted that as the electrical size of an antenna is reduced, its performance (in terms of gain, efficiency, and bandwidth) deteriorates. Generally, the radiation resistance decreases, while the reactive component of its impedance increases, leading to a poor match with the feed line or network. As such, there is a compromise between miniaturization and performance; small antenna theory dictates that a favourable compromise is reached when the antenna fully occupies a volume defined by the radius a [2-8]. Consequently, there is tremendous potential in the successful manufacture of novel, 3D antenna geometries – hitherto too complex to fabricate with traditional processes – using emergent technologies such as holographic photolithography, 3D printing, direct-write printing, direct transfer patterning, thermal transfer printing, and aerosol jet printing [9-18].

[*]Saad Mufti, Department of Electronic and Electrical Engineering, University of Sheffield, S10 2TN, UK, Email: saadmuftism@sheffield.ac.uk

PBF-AM is a particularly attractive technology, as it has the capability to print with metallic raw materials, simplifying the production line compared to some of the other additive techniques cited above. Techniques such as stereolithography and aerosol jet printing entail first printing a plastic or ceramic component, and then subsequently coating it with a layer of metal. Direct-write techniques are serial in nature, and thus slower and unsuitable for mass manufacture. Patterning techniques are suspect to wear and tear of the mechanical stamping parts. All of these processes require a substrate for subsequent metal patterning, whereas no such limitation applies to PBF-AM.

Simulated results for radiation efficiency are presented for a type of planar inverted-F antenna (PIFA), introduced in [19], as it is voluminously extended into the 3rd dimension. The results are validated with three PBF-AM fabricated antennas: flat, part-spherical (~66% fill-factor), and full-spherical (~100% fill-factor), using the Wheeler cap method [20-22]. Radiation pattern measurements are provided for two full-spherical prototypes to compare the effects of fabrication tolerances. The main aims of this research are to present measurement results on pioneering 3D antennas fabricated using AM, and to assess the viability of this emergent technique, in its current state, as a possible fabrication method for inherently 3D antennas. In particular, two popular PBF-AM techniques – which make use of a laser and an electron beam as the heat source, respectively – are compared and the relative strengths and weaknesses assessed.

2. Background on Additive Manufacture

Additive manufacture, also known as 3D printing, is the umbrella term given to a range of processes where raw materials are formatively shaped, layer by layer, to form a desired object or part [23, 24]. A generic illustration of a PBF printer is provided in Figure 1; the technology can be used on a wide range of plastics and metals, and has the obvious advantages of speed, reduced waste, and greatly simplified production lines. In particular, PBF-AM is an attractive technology for antenna manufacture as it is one the few subcategories of 3D printing which can handle metallic raw materials. Compared to other AM techniques, PBF is also advantageous with respect to operating costs, albeit at the expense of higher power consumption and lower build volumes. Two common heat sources used in PBF are lasers and electron beams, leading to the respective processes termed as selective laser melting (SLM) and selective electron beam melting (SEBM).

For this research, the laser based Renishaw AM250 and the electron beam based Arcam A2 printers were used. The difference in heat source leads to some inherent differences in the printing process. In an SLM printer, the chamber is typically filled with inert gas in order to minimize oxidation and degradation of the powdered material. A pressurized vacuum environment is necessary in the SEBM printer, otherwise the bombarded electrons would interact with atoms in the gas and thus get deflected. The electron beam can also be moved a lot faster than the laser, leading to shorter build times. Energy delivered by the electron beam is, however, more diffuse than a laser, leading to a greater area affected by the heat. Consequently, the minimum feature size, resolution, and surface-finish of the SEBM process are typically worse off compared to SLM.

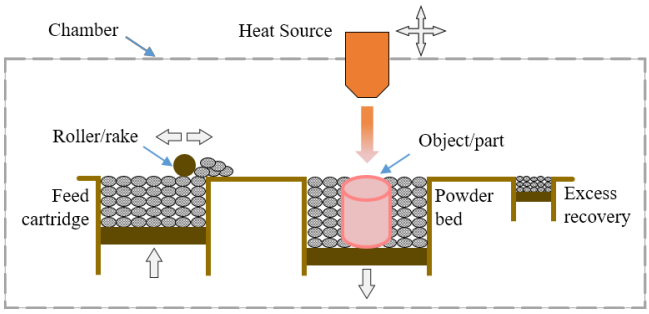


Figure 1. Illustration of generic PBF-AM printer.

In both processes, the powder in the build platform is maintained at an elevated temperature, typically just below its melting point and/or glass transition temperature; the build platform itself may be heated with the help of resistive heaters. For SLM, this elevated temperature allows for lowering the power rating of the laser, and prevents faults that might arise due to non-uniform thermal expansion and contraction. In the SEBM process, the powder is further heated to slightly sinter the particles prior to bombardment from the electron beam. This is done primarily to increase conductivity throughout the powder, to allow the negatively charged electrons to swiftly affect the desired object region. This eliminates the need for any sacrificial support structures under overhanging features, which are sometimes required in SLM printers.

3. Experimental Procedure

3.1. Antenna Design

Figure 2 shows the front profile view of the PIFA variant, based on the original design reported in [19]; the metallization is approximately 0.7 mm thick, resting on a hollow substrate with the same wall thickness, with the entire back side metallized to form the 'ground plane'. The centrally placed digitated structure contributes to a lowering of the resonance frequency of the antenna through additional capacitance and inductance. In fact, there are two current paths on this structure, contributing to orthogonally polarized modes, which may be tuned closer together with the relative positioning of the feed point and shorting pin, thereby increasing the effective bandwidth of the antenna. However, this was not a focus of the current research, as uncertainties in material properties and manufacturing tolerances meant that the focus was shifted purely to comparative efficiency measurements. Note than an extensive parametric study on antenna design and performance is presented in the reference [19], and is applicable to the voluminous antennas presented here.

This template is used for three variations of antennas with different voluminous expansions: flat, part-spherical, and full-spherical, as shown in Figure 3. For each antenna, the maximum radius, and parameter *a*, is 15.0 mm. Simulated results for the resonance frequency and efficiency as a function of the parameter *h* (as described in Figure 3) are provided in Figure 4 and Figure 5, respectively. It can be seen from Figure 5 that the 2nd resonance varies more as a function of the voluminous expansion, *h*, as compared to the 1st resonance, with the general trend that the respective resonance frequencies reduce as *h* increases, since the currents have a longer effective length on a more curved surface. The corresponding radiation efficiency at resonance also exhibits

a general trend, where the efficiency increases as h is increased and the antenna occupies more and more of the Chu sphere. For the 2nd resonance, the radiation efficiency of the flat antenna is 14%, increasing to 68% for the full-spherical antenna. Establishing an overall trend for this particular antenna geometry, at the dimensions specified, a unit (1 mm) change in height h results in an increase in radiation efficiency of 1.6 percentage points. Thus, the PIFA based antennas may be made more efficient, subsequently delivering power savings in appliances, at the cost of increased cross-sectional profile.

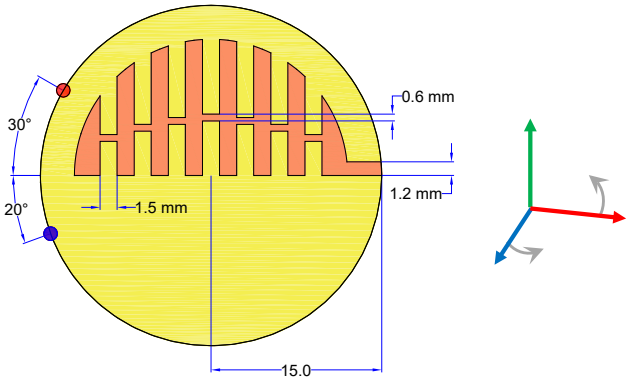


Figure 2. Schematic of top metallization for flat profile. Red dot is feed point; blue dot is shorting point. Back is fully metallicized.

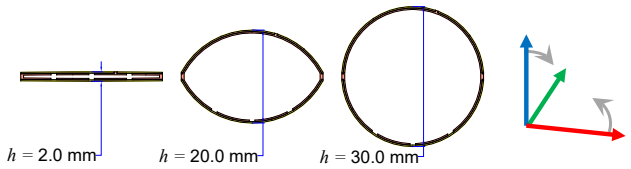


Figure 3. Cross-sectional view of CAD models simulated in CST Microwave Studio ®, for progressively greater voluminous expansion.

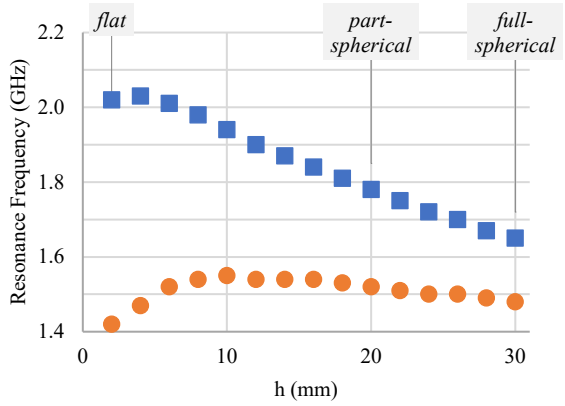


Figure 4. Simulated resonance frequency of (orange) 1st and (blue) 2nd resonances as a function of h. Fabricated antennas are annotated.

3.2. Fabricated Prototypes

The SLM process was trialed first, on a Renishaw AM250 printer, in order to take advantage of its superior surface finish. However, the process is much less capable of manufacturing overhanging features such as the centrally populated digits. Where this limitation is usually overcome with the use of sacrificial support structures analogous to scaffoldings used in building construction, the removal of such structures proved too damaging to the delicate digits.

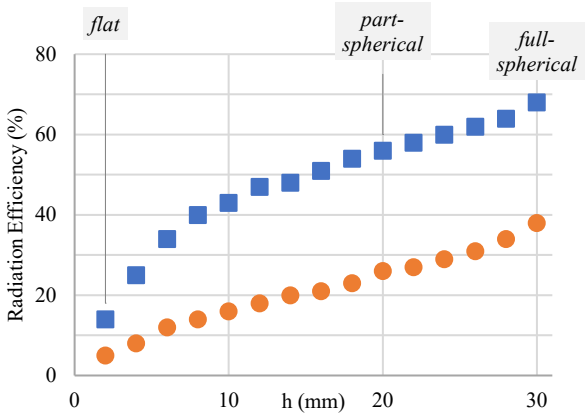


Figure 5. Simulated radiation efficiency of (orange) 1st and (blue) 2nd resonances as a function of h. Fabricated antennas are annotated.

A photograph of an antenna part fabricated with SLM is shown in Figure 6a, with the support structures partly removed. To alleviate this limitation, the antenna parts were subsequently fabricated using the SEBM based Arcam A2 printer. There is no need for support structures in this technique, as the metallic powder is pre-sintered before being melted by the electron beam. However, the poorer surface finish is evident from the photograph in Figure 6b. This is due to the use of coarser powder layers with the SEBM process (40-105 µm compared to 15-45 µm for the laser process), and the comparatively larger area affected at any one time by the electron beam compared to a heating laser.

(a) (b)

Figure 6. Photographs of metallic antenna parts fabricated using (a) SLM, and (b) SEBM.

In total, three variations of the same antenna design were fabricated: flat, part-spherical, and full-spherical; two of these are pictured in Figure 7. The antenna radiator and ground plane were fabricated separately using a Titanium alloy, Ti-6Al-4V ($\sigma \approx 10^5$ Sm^{-1}). Hollow plastic substrates fabricated from the plastic nylon12 were used to provide support. Copper tape was subsequently used to short the top metallization to the ground plane, and to provide contact pads for soldering Sub Miniature version A (SMA) connectors for interfacing to a power source. For this batch of antennas, a total of 18 parts were printed in just under two and a half hours; the build specifications for these antennas using the Arcam A2 printer are as follows:

- the printer was run under a vacuum of about 0.0001 mbar, with a slight amount of helium inserted into the chamber to help dissipate excess charge induced into the powder bed by the electron beam,
- a stainless-steel base plate was heated to 730 °C by the electron beam before any powder was deposited,
- for each layer of powder deposited (50 µm thick), the beam was first defocused and scanned across the layer at a high speed (30 ms^{-1}) with 2.3 kW power, for around 15 seconds, to sinter the powder,
- each part area was subsequently melted using a focused beam with a back and forth raster pattern at a speed of around 0.35 ms^{-1} with 300 W power,
- thereafter, successive powder layers were deposited and the previous sintering and melting steps repeated until the antenna parts were fully realized.

(a) (b)

Figure 7. Photographs of SEBM PBF-AM fabricated antenna prototypes, shown assembled; (a) flat, and (b) full-spherical.

4. Results and Discussion

Return loss measurements for the flat, part-spherical, and full-spherical antennas were conducted using the Agilent E5071B vector network analyser. Radiation efficiency (for 20 MHz bandwidth) was subsequently computed using the using the constant-loss-resistor Wheeler cap method described in [22]. Note that efficiency results were computed for the second resonance only as this is the dominant resonance from measurements, and that repeat measurements per prototype exhibit an error margin of ±10 %.

The corresponding simulations were carried out using CST Microwave Studio ®, using the frequency-domain solver. The titanium alloy was approximated using copper (Cu) metal properties in the CST material library, with reduced conductivity, σ. A notable limitation of using PBF-AM for manufacturing antenna parts is the relative lack of knowledge regarding the electromagnetic properties of the raw materials; this is in fact an active area of current research in the AM community [23, 25-27]. As such, these were iteratively refined in simulations. However, the general trends in performance observed can be interpreted as a validation of small antenna theory, as well as the usefulness of AM techniques in the swift prototyping of non-conventional antenna geometries.

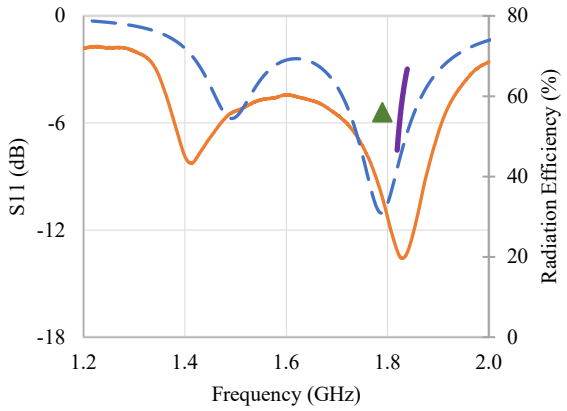


Figure 8. S_{11} results for PBF-AM part-spherical prototype; (orange) measured, (blue) simulated. Radiation efficiency at respective resonance frequencies, for the 2nd resonance; (purple) measured, (green) simulated.

Figure 9-11 compare the measured and simulated S_{11} parameter for the flat, part-spherical and two full-spherical prototypes, respectively. Note that S_{11} is the scattering parameter as obtained from a vector network analyser and is relatable to other reflectivity figures of merit such as return loss and reflection coefficient, and is typically presented in decibels (dB) [28]. Simulated efficiency at resonance and measured efficiency for 20 MHz bandwidth about the resonance frequency are also displayed. Despite differences in the absolute values for S_{11} and radiation efficiency, it can be seen that general trends are corroborated, as expected from small antenna theory. Computed radiation efficiency values for the 2nd resonances exhibit an increase as the voluminous expansion of the PIFA is increased.

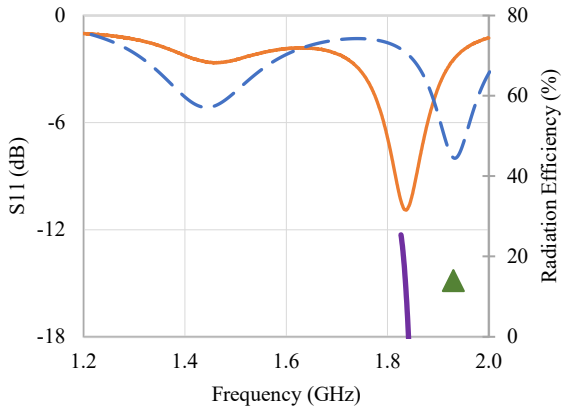


Figure 9. S_{11} results for PBF-AM flat prototype; (orange) measured, (blue) simulated. Radiation efficiency at respective resonance frequencies, for the 2nd resonance; (purple) measured, (green) simulated.

For the two full-spherical prototypes, despite the S_{11} curves following a similar curve, there is a difference of 50 MHz in the 2nd resonance. The mean radiation efficiency at resonance for the two prototypes is 68 %, with a discrepancy of 16 percentage points. Again, these differences are primarily due to inherent manufacturing differences in the antenna parts, as well as manual tolerances for the application of solder and copper tape contact pads. The differences in absolute values between simulations and measurements can be attributed to a number of factors, including

unknown changes in material properties during the printing process, the presence of solder and copper tape for electrical contact to the SMA connector, as well as the rough surface finish on the antenna parts.

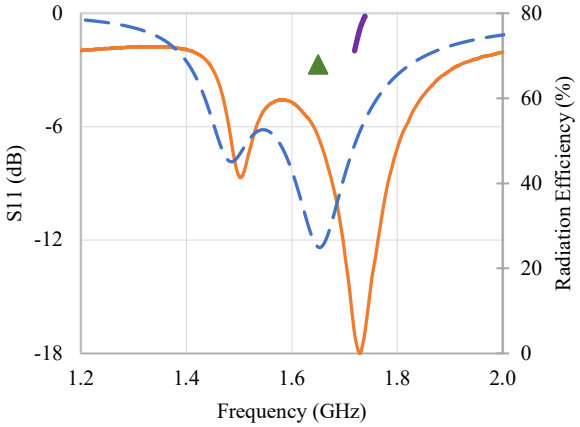


Figure 10. S_{11} results for PBF-AM full-spherical prototype-A; (orange) measured, (blue) simulated. Radiation efficiency at respective resonance frequencies, for the 2nd resonance; (purple) measured, (green) simulated.

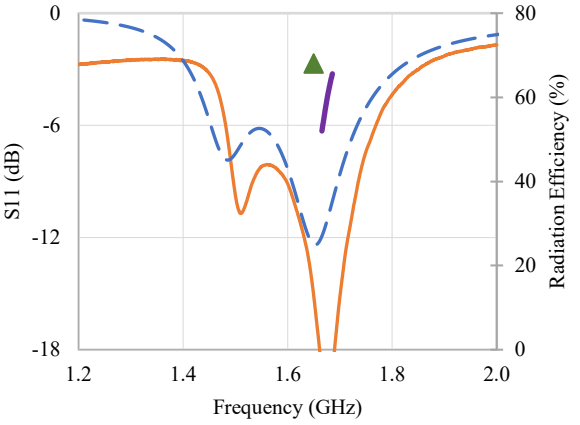


Figure 11. S_{11} results for PBF-AM full-spherical prototype-B; (orange) measured, (blue) simulated. Radiation efficiency at respective resonance frequencies, for the 2nd resonance; (purple) measured, (green) simulated.

Lastly, the realized gain for the full-spherical antenna prototypes was measured in an anechoic chamber, with an Agilent E5071C vector network analyser. Due to the particular feed location for this antenna, partial power gains were measured with respect to vertical and horizontal polarizations, and the gain transfer method [29] used subsequently to compute the total realized gain. Fig 6 and Fig. 7 plot the results for azimuth and elevation planes, respectively, measured at 1.7 GHz; the absolute gain is presented here in units of decibels-over-isotropic (dBi). It can be seen from these plots that simulation and measurement results largely agree, and are expectedly low directivity; there are some artifacts due to the setup (antenna holder) in the azimuth plane towards the rear of the antenna. Key results for the 2nd resonance of each antenna are summarized in Table 1.

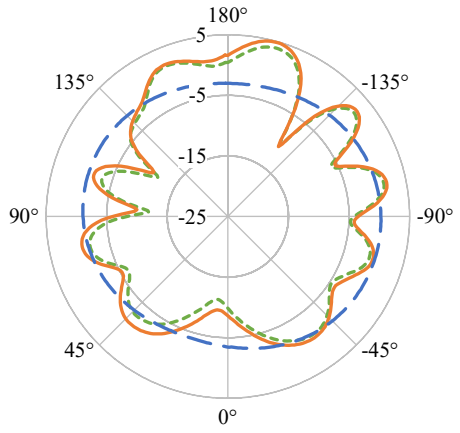


Figure 12. Realized gain (dBi) for the full-spherical ESA at 1.7 GHz; (orange) measured – sample A, (green) measured – sample B, (blue) simulated; azimuth (*z-x*) plane. Combined absolute gain (*φ and θ* components).

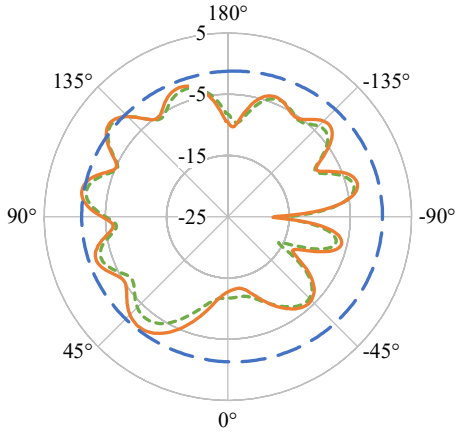


Figure 13. Realized gain (dBi) for the full-spherical ESA at 1.7 GHz; (orange) measured – sample A, (green) measured – sample B, (blue) simulated; elevation (*x-y*) plane. Combined absolute gain (*φ and θ* components).

Table 1. Simulated and measured resonant frequency and radiation efficiency at 2nd resonance for PBF-AM fabricated antennas. Measured and simulated gain at 1.7 GHz is provided for the two full-spherical antennas.

	Flat	*Part-spherical*	*Full-spherical (A)*	*Full-spherical (B)*
Simulated frequency (GHz)	1.93	1.79	1.65	-
Measured frequency (GHz)	1.84	1.83	1.73	1.68
Simulated efficiency (%)	14	56	68	-
Measured efficiency (%)	11	59	76	60
Simulated average gain; azimuth plane (dBi)	-	-	–1.7	-
Measured average gain; azimuth plane (dBi)	-	-	–2.2	–3.2
Simulated average gain; elevation plane (dBi)	-	-	–0.8	-
Measured average gain; elevation plane (dBi)	-	-	–5.9	–6.4

5. Conclusions and Future Work

This paper presents work done on voluminous, additive manufactured antennas based on a novel PIFA design. Measured results for return loss, radiation efficiency, and radiation pattern are presented for four prototypes manufactured using the SEBM based Arcam A2 printer. The metallization was realized from a Titanium alloy, Ti-6Al-4V, and hollow support substrates printed from nylon12. Copper tape was subsequently used to solder SMA connectors, and to short the top metallization to the respective antenna ground planes.

Radiation efficiency results corroborate the trends expected from simulations and small antenna theory. The mean radiation efficiency of 68% for the full-spherical antennas represents an improvement of 57 percentage points over their flat counterpart. The next stage of this project is to obtain an improved understanding of the material properties of metallic powders used in the printing process, as it has been demonstrated that the PBF-AM technique is viable for quick prototyping of complex and voluminous antenna geometries too difficult to realize using traditional fabrication techniques. Finally, the use of laser as the heat source would greatly improve the surface finish, but comes at the cost of limitations on the complexity of antenna geometry

Conflict of Interest

The authors declare no conflict of interest.

References

[1] S. Mufti, A. Tennant, L. Seed, C. J. Smith, "Efficiency measurements of additive manufactured electrically small antennas" in 2017 International Workshop on Antenna Technology: Small Antennas, Innovative Structures, and Applications, Athens, Greece, 2017. https://doi.org/10.1109/IWAT.2017.7915397

[2] S. R. Best, "Small and fractal antennas" in Modern Antenna Handbook, C. A. Balanis, Ed. John Wiley & Sons, 2008, pp. 475–528.

[3] J. L. Volakis, C. C. Chen, K. Fujimoto, Small Antennas: Miniaturization Techniques and Applications, McGraw-Hill, 2010.

[4] K. Fujimoto, H. Morishita, Modern Small Antennas, Cambridge University Press, 2014.

[5] L. J. Chu, "Physical limitations of omni-directional antennas" J. Appl. Phys., **19**(12), 1163–1175, 1948. https://doi.org/10.1063/1.1715038

[6] J. S. McLean, "A re-examination of the fundamental limits on the radiation Q of electrically small antennas" IEEE T. Antenn. Propag., **44**(5), 672–676, 1996. https://doi.org/10.1109/8.496253

[7] A. D. Yaghjian, S. R. Best, "Impedance, bandwidth, and Q of antennas" IEEE T. Antenn. Propag., **53**(4), 1298–1324, 2005. https://doi.org/10.1109/TAP.2005.844443

[8] C. Pfeiffer, "Fundamental Efficiency Limits for Small Metallic Antennas" IEEE T. Antenn. Propag., **65**(4), 1642–1650, 2017. https://doi.org/10.1109/TAP.2017.2670532

[9] A. Maiden, "Lithography in three dimensions using computer-generated holograms," Ph.D. Thesis, Durham University, 2005.

[10] D. E. Anagnostou, A. A. Gheethan, A. K. Amert, K. W. Whites, "A direct-write printed antenna on paper-based organic substrate for flexible displays and WLAN applications" J. Disp. Technol., **6**(11), 558–564, 2010. https://doi.org/10.1109/JDT.2010.2045474

[11] J. J. Adams *et al*., "Conformal printing of electrically small antennas on three-dimensional surfaces" Adv. Mater., **23**(11), 1335–1340, 2011. https://doi.org/10.1002/adma.201003734

[12] J. Hoerber, J. Glasschroeder, M. Pfeffer, J. Schilp, M. Zaeh, J. Franke, "Approaches for additive manufacturing of 3D electronic applications" Proc. CIRP, **17**, 806–811, 2014. https://doi.org/10.1016/j.procir.2014.01.090

[13] J. J. Toriz-Garcia *et al*., "Fabrication of a 3D electrically small antenna using holographic photolithography" J. Micromech. Microeng., **23**(5), no. 055010, 2013. https://doi.org/10.1088/0960-1317/23/5/055010

[14] C. Pfeiffer, X. Xu, S. R. Forrest, A. Grbic, "Direct transfer patterning of electrically small antennas onto three-dimensionally contoured substrates" Adv. Mater., **24**(9), 1166–1170, 2012. https://doi.org/10.1002/adma.201104290

[15] M. Kgwadi, C. J. Vourch, D. J. Harrison, T. D. Drysdale, "On-demand printing of antennas for TV white-space communications" in 2014 Loughborough Antennas and Propagation Conference, Loughborough, UK, 2014. https://doi.org/10.1109/LAPC.2014.6996448

[16] I. T. Nassar, T. M. Weller, "An electrically-small, 3-D cube antenna fabricated with additive manufacturing" in 2013 IEEE Radio and Wireless Symposium, Austin, USA, 2013. https://doi.org/10.1109/RWS.2013.6486708

[17] O. S. Kim, "3D printing electrically small spherical antennas" in 2013 IEEE Antennas and Propagation Society International Symposium, Orlando, USA, 2013. https://doi.org/10.1109/APS.2013.6711047

[18] H. Yuxiao, C. Oakley, P. Chahal, J. Albrecht, J. Papapolymerou, "Aerosol Jet printed 24 GHz end-fire quasi-Yagi-Uda antenna on a 3-D printed cavity substrate" in 2017 International Workshop on Antenna Technology: Small Antennas, Innovative Structures, and Applications, Athens, Greece, 2017. https://doi.org/10.1109/IWAT.2017.7915352

[19] S. Mufti, A. Tennant, L. Seed, "Electrically small modified planar inverted-F antenna" in 9th European Conference on Antennas and Propagation, Lisbon, Portugal, 2015.

[20] E. Newman, P. Bohley, C. Walter, "Two methods for the measurement of antenna efficiency" IEEE T. Antenn. Propag., **23**(4), 457–461, 1975. https://doi.org/10.1109/TAP.1975.1141114

[21] D. M. Pozar, B. Kaufman, "Comparison of three methods for the measurement of printed antenna efficiency" IEEE T. Antenn. Propag., vol. **36**(1), 136–139, 1988. https://doi.org/10.1109/8.1084

[22] D. Agahi, W. Domino, "Efficiency measurements of portable-handset antennas using the Wheeler cap" Applied Microwave and Wireless, **36**(1), 34–42, 2000.

[23] I. Gibson, D. W. Rosen, B. Stucker, Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing, Springer, 2010.

[24] F. Calignano *et al*. "Overview on Additive Manufacturing Technologies" P. IEEE, **105**(4), 593–612, 2017. https://doi.org/10.1109/JPROC.2016.2625098

[25] Y. Huang, M. C. Leu, J. Mazumder, A. Donmez, "Additive manufacturing: current state, future potential, gaps and needs, and recommendations" J. Manuf. Sci. E-T ASME, **137**(1), no. 014001, 2015. https://doi.org/10.1115/1.4028725

[26] J. A. Slotwinski, E. J. Garboczi, "Metrology needs for metal additive manufacturing powders" JOM, **67**(3), 538–543, 2015. https://doi.org/10.1007/s11837-014-1290-7

[27] S. K. Everton, M. Hirsch, P. Stravroulakis, R. K. Leach, A. T. Clare, "Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing" Mater. Design, **95**, 431–445, 2016. https://doi.org/10.1016/j.matdes.2016.01.099

[28] D. M. Pozar, Microwave Engineering, 4th ed, John Wiley & Sons, 2011.

[29] C. A. Balanis, Antenna Theory: Analysis and Design, 4th ed, John Wiley & Sons, 2016.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 76-86 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Real-Time Flux-weakening Control for an IPMSM Drive System Using a Predictive Controller

Tian-Hua Liu*, Shao-Kai Tseng, Yi Chen, Mao-Bin Lu

Department of Electrical Engineering, National Taiwan University of Science and Technology, 106, Taiwan

ARTICLE INFO

Received: 07 October, 2017
Accepted: 29 October, 2017
Online: 10 November, 2017

Keywords :
predictive controller
real-time tuning
flux-weakening control
IPMSM

ABSTRACT

This paper proposes extended-range high-speed control for an IPMSM drive system. A simple real-time tuning flux-weakening control algorithm is proposed and implemented to control an IPMSM drive system in a wide variable speed range, from 3 r/min up to 2700 r/min. This flux-weakening control algorithm does not require any motor parameters and only needs simple mathematical computations. The proposed drive system adjusts the angle between the d-axis and q-axis current to reach flux-weakening control. In addition, a multiple sampling predictive controller is implemented to enhance the dynamic responses of the proposed drive system, which yields improved overall transient responses, superior load responses, and good tracking responses. A detailed analysis of the proposed drive system's stability is discussed as well. A 32-bit digital signal processor, TMS-320F-28335, is used to execute the predictive controller and the flux-weakening control algorithm for the IPMSM drive system. Experimental results can validate the theoretical analysis.

1. Introduction

Use of the interior permanent magnet synchronous motor (IPMSM) has greatly increased due to its advantages over other motors, which include high efficiency, wide operating speed range, high power density, and robustness. The IPMSM has two different torque components: electromagnetic torque and reluctance torque. By strategically controlling the d-axis and q-axis currents, the maximum torque per ampere (MTPA) control or flux weakening control can be achieved at different operating speeds.

This paper is an extension of work originally presented in 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE) Conference [1]. Several researchers have proposed different flux-weakening control methods to increase the high-speed operating range of an IPMSM. For example, Bolognani et al. proposed an adaptive flux-weakening controller for an IPMSM drive system. The adaptation algorithm, however, is very complicated to implement [2]. Zhu et al. implemented a flux-weakening control for a PMSM drive system [3]. The proposed method accounted for resistance voltage drop and inverter nonlinearities. This method, however, is only suitable for a PMSM drive system, not an IPMSM drive system. Kwon et al. proposed a flux-weakening control for an IPMSM with quasi-six-step operation [4]. A feed-forward path, which consists of a one dimensional look-up table, is required. Pan et al. proposed a robust flux-weakening control strategy for a surface-mounted permanent-magnet motor drive [5]. A closed-form solution of the maximum available torque-producing current is generated to achieve both fast responses and real-time tuning flux-weakening control. This method, however, is suitable for a surface mounted PMSM but not an IPMSM. Uddin et al. proposed an on-line parameter-estimation-based high-speed control of an IPMSM drive. The method of estimating the d-axis inductance, q-axis inductance, and external load, however, is very complicated [6]. Pan et al. proposed a voltage-constraint-based flux-weakening control of an IPMSM drive system [7]. The method requires a d-axis current command and q-axis current command generator, which is rather complicated. Jung et al. proposed a hexagonal voltage limit to increase the voltage amplitude and to improve the DC-link voltage utilization in an IPMSM drive system. A torque control method based on the voltage angle was investigated. The idea is very good; however, the voltage vector selection algorithm is very complicated [8]. Chaithongsuk et al. proposed an optimal design of permanent magnet motors to improve flux-weakening performance in variable speed drives. It was shown that a salient-pole motor performed better at high speeds than a non-salient pole motor [9]. Kim proposed a novel magnetic flux-weakening method for permanent magnet synchronous motors in electric vehicles [10]. By adjusting the air gap between the stator and rotor, an increased maximum speed and maximum output power could be

*Corresponding Author: Tian-Hua Liu, Mobile: +886-22737-6108
Email: Liu@mail.ntust.edu.tw

obtained. This method, however, requires very complicated motor modifications. Jevremovic et al. proposed a simple robust integral controller to achieve field weakening control for a PMSM [11]. Harnefors et al. used a systematic analysis to determine the parameters of the current and speed combined controller to reduce the complicity [12]. To improve the performance of field weakening control, Wallmark et al. [13] and Bedetti et al. implemented a voltage regulator to achieve flux-weakening control for an electric vehicle and an IPMSM [14].

In this paper, a simple flux-weakening control method is proposed. This method has several advantages over previously proposed flux-weakening methods [1]-[14]. For one thing, the parameters of the IPMSM are not required. Also, the computations for the flux-weakening control are very simple. For example, only square root, multiplication, and addition are used. In addition, to achieve fast responses and good load disturbance rejection capability from the low-speed range to high-speed range, a predictive controller is employed. In contrast to previous research on predictive control [15]-[16], this paper proposes a multiple-loop predictive control motor drive system. The sampling interval of the current-loop is 100 , and the sampling interval of the speed-loop is 1 ms. Due to the quick current regulation, the proposed system has better performance than previous one-loop predictive control drive systems [15]-[16]. Unlike previous methods of flux-weakening control without using motor parameters [2], [16]-[17], the proposed method provides a simple control algorithm for voltage regulation without an adaptive controller or PI controller. To the authors' best knowledge, the proposed flux-weakening control method is a new idea. In addition, the idea of the predictive controller applied in the simplified flux-weakening control of an IPMSM is original.

2. Predictive Controller Design

The theory of predictive control was first proposed in the 1970s, and predictive controllers have been used in chemical and steel processing for more than 40 years. Recently, predictive control has been applied in motor drives due to the fast computation ability of the digital signal processor (DSP). The design objective of a predictive controller is to compute the control input u to optimize the future dynamic behavior for the output y of the plant. This optimization must be performed within a limited time window. Several researchers have investigated predictive control for motor drives. For example, Pacas et al. proposed a predictive direct torque control for PMSMs, which yielded a faster torque settling time than the classic field-oriented control with a PI controller [15]. Bolognani et al. investigated the design and implementation of model predictive control for electric motor drives [16]. Fuentes et al. implemented a predictive speed control for a two-mass system driven by a PMSM. Their proposed method allowed feedback of all the mechanical and electrical state variables into a single control input and provided a higher bandwidth closed-loop speed control [18]. Siami et al. proposed an improved predictive current control using prediction error correction for PMSMs [19]. Tarczewski et al. implemented a constrained state feedback speed control for PMSMs, which calculated the boundaries of control signals to provide permissible values of the future state variables [20].

In this paper, a systematic predictive controller design is proposed for an IPMSM drive system. The proposed predictive controllers, including a predictive current controller and a predictive speed controller, provide a multiple sampling rate control system and a closed-loop speed-control block diagram that are different from the predictive controllers proposed in previously published papers [15]-[16], [18]-[20].

The design of a predictive controller requires the following. First, a model of the uncontrolled drive system is developed. Next, a cost function that represents the desired behavior of the system is defined. Finally, a predictive controller for the motor drive system is derived.

2.1. Predictive Current Controller

The state-variable equation of an IPMSM can be expressed as

$$\frac{d}{dt}\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \begin{bmatrix} \frac{-r_s}{L_d} & 0 \\ 0 & \frac{-r_s}{L_q} \end{bmatrix}\begin{bmatrix} i_d \\ i_q \end{bmatrix} + \begin{bmatrix} \frac{1}{L_d} & 0 \\ 0 & \frac{1}{L_q} \end{bmatrix}\begin{bmatrix} v_d \\ v_q \end{bmatrix} + \begin{bmatrix} \frac{1}{L_d} & 0 \\ 0 & \frac{1}{L_q} \end{bmatrix}\begin{bmatrix} \omega_{re} L_q i_q \\ -\omega_{re}(L_d i_d + \lambda_m) \end{bmatrix} \quad (1)$$

Where d/dt is the differential operator, i_d and i_q are the d-axis and q-axis stator currents, r_s is the stator resistance, v_d and v_q are the d-axis and q-axis stator voltages, L_d and L_q are the d-axis and q-axis self-inductances, λ_m is the permanent magnetic flux leakage, and ω_{re} is the electric motor speed.

After transferring the continuous-time domain into the discrete-time domain, (1) can be expressed as

$$i_d(z) = \frac{\beta_{cd} z^{-1}}{1-\alpha_{cd} z^{-1}}(v_d - \omega_{re} L_q i_q) \quad (2)$$

and

$$i_q(z) = \frac{\beta_{cq} z^{-1}}{1-\alpha_{cq} z^{-1}}(v_q - \omega_{re}(L_d i_d + \lambda_m)) \quad (3)$$

The related parameters in (2) and (3) are shown as follows

$$\alpha_{cd} = e^{\frac{-r_s T_{cur}}{L_d}} \quad (4a)$$

$$\beta_{cd} = \frac{1}{r_s}(1-\alpha_{cd}) \quad (4b)$$

$$\alpha_{cq} = e^{\frac{-r_s T_{cur}}{L_q}} \quad (4c)$$

$$\beta_{cq} = \frac{1}{r_s}(1-\alpha_{cq}) \quad (4d)$$

where T_{cur} is the sampling interval of current-loop control. Because the sampling time interval is shorter than the time constant of the stator current-loop, Euler approximation can be used here. As a result, the parameters shown in (4a)-(4d) can be described as follows

$$\alpha_{cd} \approx 1-\frac{r_s T_{cur}}{L_d} \tag{5a}$$

$$\beta_{cd} \approx \frac{r_s T_{cur}}{L_d} \tag{5b}$$

$$\alpha_{cq} \approx 1-\frac{r_s T_{cur}}{L_q} \tag{5c}$$

and

$$\beta_{cq} \approx \frac{r_s T_{cur}}{L_q} \tag{5d}$$

Then (2) and (3) can be rewritten as follows

$$X_{cur}(k+1) = A_{cur}X_{cur}(k) + B_{cur}U_{cur}(k) + B_{cur}F_{cur}(k) \tag{6}$$

and

$$X_{cur}(k) = \begin{bmatrix} i_d(k) & i_q(k) \end{bmatrix}^T \tag{7a}$$

$$F_{cur}(k) = \begin{bmatrix} \omega_{re}(k)L_q i_q(k) & -\omega_{re}(k)\left(L_d i_d(k) + \lambda_m\right) \end{bmatrix}^T \tag{7b}$$

$$U_{cur}(k) = \begin{bmatrix} v_d(k) & v_q(k) \end{bmatrix}^T \tag{7c}$$

$$A_{cur} = \begin{bmatrix} 1-\frac{r_s T_{cur}}{L_d} & 0 \\ 0 & 1-\frac{r_s T_{cur}}{L_q} \end{bmatrix} \tag{7d}$$

$$B_{cur} = \begin{bmatrix} \frac{T_{cur}}{L_d} & 0 \\ 0 & \frac{T_{cur}}{L_q} \end{bmatrix} \tag{7e}$$

According to (6), the (k+1)-th predictive current can be expressed as

$$\hat{X}_{cur}(k+1) = A_{cur}X_{cur}(k) + B_{cur}\left(U_{cur}(k) + F_{cur}(k)\right) \tag{8}$$

The performance index of the d-axis current control is defined as

$$J_{cur} = \left[\hat{i}_d(k+1) - i_d^*(k+1)\right]^2 \tag{9}$$

The following equation is the partial differential of (9) manipulated to equal zero.

$$\frac{\partial J_{cur\_d}}{\partial U_d} = \frac{\partial\left(\hat{i}_d(k+1) - i_d^*(k+1)\right)^2}{\partial v_d(k)} = 0 \tag{10}$$

Substituting (8) into (10), we can obtain

$$\frac{\partial J_{cur\_d}}{\partial v_d(k)} = \frac{\partial\left((1-\frac{T_{cur}r_s}{L_d})i_d(k) + (\frac{T_{cur}}{L_d})\left(v_d(k) + \omega_{re}(k)L_q i_q(k)\right) - i_d^*(k+1)\right)^2}{\partial v_d(k)}$$
$$= 2\left((1-\frac{T_{cur}r_s}{L_d})i_d(k) + (\frac{T_{cur}}{L_d})\left(v_d(k) + \omega_{re}(k)L_q i_q(k)\right) - i_d^*(k+1)\right)\left(\frac{T_{cur}}{L_d}\right) = 0 \tag{11}$$

After arranging the above, it is not difficult to obtain

$$\left((1-\frac{T_{cur}r_s}{L_d})i_d(k) + (\frac{T_{cur}}{L_d})\left(v_d(k) + \omega_{re}(k)L_q i_q(k)\right) - i_d^*(k+1)\right)$$
$$= \left((\frac{L_d}{T_{cur}} - r_s)i_d(k) + \left(v_d(k) + \omega_{re}(k)L_q i_q(k)\right) - \frac{L_d}{T_{cur}}i_d^*(k+1)\right) = 0 \tag{12}$$

Finally, from (12), we can derive the d-axis voltage command as

$$v_d^*(k) = r_s i_d(k) + \frac{L_d}{T_{cur}}\left(i_d^*(k+1) - i_d(k)\right) - \omega_{re}(k)L_q i_q(k) \tag{13}$$

By using the same principle, we can derive the q-axis voltage command as follows

$$v_q^*(k) = r_s i_q(k) + \frac{L_q}{T_{cur}}\left(i_q^*(k+1) - i_q(k)\right) + \omega_{re}(k)\left(L_d i_d(k) + \lambda_m\right) \tag{14}$$

2.2. Predictive Speed Controller

The differential equation of an IPMSM is

$$\frac{d}{dt}\omega_{rm} = \frac{1}{J_m}(T_e - T_L - B_m\omega_{rm}) \tag{15}$$

Where J_m ω_{rm} is the mechanical speed, $\frac{d}{dt}$ is the differential operator, T_e is the total output torque, T_L is the external load, and B_m is the viscous frictional coefficient.

According to (15) and assuming $T_L = 0$, the speed can be expressed as

$$\omega_{rm}(s) = \frac{T_e}{J_m s + B_m} \tag{16}$$

By converting (16) into discrete form, one can obtain

$$\begin{aligned}\omega_{rm}(k+1) &= e^{-(B_m/J_m)T_{sp}}\omega_{rm}(k) + \frac{1}{B_m}(1-e^{-(B_m/J_m)T_{sp}})T_e(k) \\ &= a_1\omega_{rm}(k) + b_1 u(k)\end{aligned} \tag{17}$$

and

$$a_1 = e^{-(B_m/J_m)T_{sp}} \tag{18}$$

$$b_1 = \frac{1}{B_m}(1-e^{-(B_m/J_m)T_{sp}}) \tag{19}$$

$$u(k) = T_e(k) \tag{20}$$

where T_{sp} is the sampling interval of the speed-loop control. Next, by setting the predictive window as 1 and the predictive control as 1, the performance index J_{sp} is expressed as

$$\begin{aligned}J_{sp} = &\sum_{i=1}^{1}\left[P(z)\hat{\omega}_{rm}^{pre}(k+i) - P(1)\omega_{rm}^{*}(k+i)\right]^2 \\ &+ \rho\sum_{i=1}^{1}\left[Q(z)u(k+i-1)\right]^2\end{aligned} \tag{21}$$

where $\hat{\omega}_{rm}^{pre}$ is the predictive speed, $P(z)$ and $Q(z)$ are polynomials of z, and ρ is the weighting factor. By computing $\frac{\partial J_{sp}}{\partial u(k)}$, one can obtain

$$\begin{aligned}\frac{\partial J_{sp}}{\partial u(k)} = 2\frac{P(z)\partial\hat{\omega}_{rm}^{pre}(k+1)}{\partial u(k)}&[P(z)\hat{\omega}_{rm}^{pre}(k+1) \\ &- P(1)\omega_{rm}^{*}(k+1)] + 2\rho Q^2(z)u(k) = 0\end{aligned} \tag{22}$$

Substituting (17) into (22), one can derive

$$\begin{aligned}P(z)b_1\left(P(z)\left(b_1u(k) + a_1\omega_{rm}(k)\right) - P(1)\omega_{rm}^{*}(k+1)\right) \\ + \rho Q^2(z)u(k) = 0\end{aligned} \tag{23}$$

The speed command at (k+1)th sampling interval can be expressed as a fixed command Sp and then can be defined as

$$\omega_{rm}^{*}(k+1) = S_p \tag{24}$$

By substituting (24) into (23), we can obtain

$$(P(z)b_1)^2u(k) + \rho Q^2(z)u(k) = P(z)b_1P(1)S_p - P^2(z)b_1a_1\omega_{rm}(k) \tag{25}$$

The control input can be expressed as

$$u(k) = R(MS_p - N\omega_{rm}(k)) \tag{26}$$

and

$$M = \frac{P(1)}{b_1P(z)} \tag{27}$$

$$N = \frac{a_1}{b_1} \tag{28}$$

$$R = \frac{1}{1 + \frac{\rho}{b_1^2}(\frac{Q(z)}{P(z)})^2} \tag{29}$$

The predictive controller is shown in Figure 1. The block diagram of the predictive control includes a command gradation M(z), a controller R(z), a plant P(z), and a feedback gain N. The structure is very similar to a standard control system; however, the parameters can be systematically derived. Unlike [16], this paper outlines a systematic closed-loop block diagram of the predictive controller, which provides a clear feedback control structure for an IPMSM drive system. This is an unprecedented method of predictive control and also a major aspect of this paper.

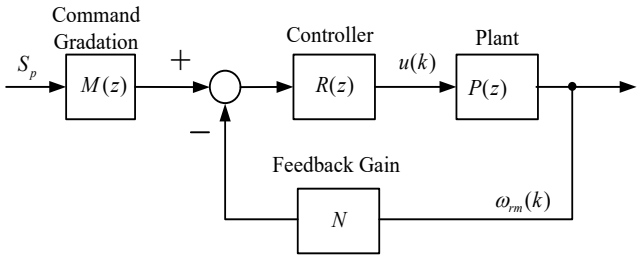


Figure 1 Block diagram of predictive speed-loop control.

2.3. Stability Analysis of the Proposed Drive System

The desired torque command of an IPMSM can be generated by the following equation

$$\frac{d}{dt}\omega_{rm}^{*} = \frac{1}{J_m}(T_e^{*} - T_L - B_m\omega_{rm}^{*}) \tag{30}$$

In the real world, there are parameter variations of the IPMSM and the external load measuring error. To compensate for that, the total equivalent load with uncertainty can be expressed as T_D, which includes the summation of the external load, the load influenced

by the parameter variations, torque tracking errors of both MTPA and flux weakening, and external load measuring error. Then (30) can be rewritten as

$$\frac{d}{dt}\omega_{rm}^{*} = \frac{-B_m}{J_m}\omega_{rm}^{*} + \frac{1}{J_m}T_e - T_{\mathrm{D}} \tag{31}$$

and

$$T_D = \frac{1}{J_m}(\Delta J_m \frac{d}{dt}\omega_{rm}^{*} + \Delta B_m \omega_{rm}^{*} + T_L) \tag{32}$$

where T_D is the total load disturbance, ΔJ_m is the variation of inertia, and ΔB_m is the variation of the viscous coefficient. Next, the dynamic speed equation of the IPMSM is

$$\frac{d}{dt}\omega_{rm} = \frac{-B_m}{J_m}\omega_{rm} + \frac{1}{J_m}T_e - \hat{T}_D \tag{33}$$

Where $\hat{T}_D$ is the estimated total load disturbance, which includes the influence of the motor parameters' uncertainty. The tuning step of the θ_{adv}, which is 0.18°, creates a varying torque that is less than 1% of the output torque T_e. As a result, the influence of the uncertainty is bound. The speed error, which is the difference between the speed command and the real speed, can be expressed as

$$e = \omega_{rm}^{*} - \omega_{rm} \tag{34}$$

Taking the differential of the speed error, one can obtain

$$\dot{e} = \frac{d}{dt}(\omega_{rm}^{*} - \omega_{rm}) = \frac{d}{dt}\omega_{rm}^{*} - \frac{d}{dt}\omega_{rm} \tag{35}$$

Subtituting (31) and (33) into (35), one can obtain

$$\begin{aligned}\dot{e} &= (\frac{-B_m}{J_m}\omega_{rm}^{*} + \frac{1}{J_m}T_e - T_{\mathrm{D}}) - (\frac{-B_m}{J_m}\omega_{rm} + \frac{1}{J_m}T_e - \hat{T}_D) \\ &= \frac{-B_m}{J_m}(\omega_{rm}^{*} - \omega_{rm}) - \tilde{T}_D\end{aligned} \tag{36}$$

We can define the load estimation error as $\tilde{T}_D = T_D - \hat{T}_D$. Assuming the external load changes slowly when compared to the quick responses of the speed control loop, it is possible to let $\dot{T}_D = 0$. The differential of the estimated load, therefore, can be expressed as follows

$$\dot{\tilde{T}}_D = -\dot{\hat{T}}_D \tag{37}$$

Define a lyapunov function as

$$V = e^2 + \frac{1}{k_1}(\tilde{T}_D)^2 \tag{38}$$

where k_1 is a weighting factor. By taking the derivative of the Lyapunov function, one can derive

$$\dot{V} = 2e\dot{e} + \frac{2}{k_1}\tilde{T}_D\dot{\tilde{T}}_D \tag{39}$$

Substituting (36) into (39), one can obtain

$$\dot{V} = 2e(\frac{-B_m}{J_m}e - \tilde{T}_D) + \frac{2}{k_1}\tilde{T}_D\dot{\tilde{T}}_D = 2\frac{-B_m}{J_m}e^2 + 2\tilde{T}_D(\frac{1}{k_1}\dot{\tilde{T}}_D - e) \tag{40}$$

According to (40), it is possible to choose an adaption law as follows

$$\dot{\hat{T}}_D = k_1 e \tag{41}$$

Substituting (41) into (40), we can obtain the differential of Lyapunov function as

$$\dot{V} = 2\frac{-B_m}{J_m}e^2 \le 0 \tag{42}$$

From (42), the differential of Lyapunov function is negative semi-definite. Then, Barbalet lemma can be applied. First, a function y_1 is set, which is equal to $\dot{V}$. Next, by integrating the y_1, we can obtain

$$\int_0^{\infty} y_1(\tau)d\tau = V(0) - V(\infty) < \infty \tag{43}$$

From (43), we can conclude

$$\lim_{t\to\infty} y_1(t) = 0 \tag{44}$$

According to (44), the speed error converges to zero as time approaches infinity.

3. MTPA Control and Flux-weakening Control

3.1. Basic Principle

The dynamic d-q axis voltages of an IPMSM can be expressed as

$$v_d = r_s i_d + L_d\frac{di_d}{dt} - \omega_{re}L_q i_q \tag{45}$$

and

$$v_q = r_s i_q + L_q \frac{di_q}{dt} - \omega_{re} L_d i_d + \omega_{re} \lambda_m \quad (46)$$

The total output torque of an IPMSM is

$$T_e = \frac{3}{2}\frac{P}{2}[\lambda_m + (L_d - L_q)\ i_d]\ i_q$$
$$= \frac{3}{2}\frac{P}{2}[\lambda_m i_q - (L_d - L_q)\sqrt{I_s^{\ 2} - i_q^{\ 2}}]i_q \quad (47)$$

Where I_s is the current amplitude. Two major control algorithms can be derived from (47) and can be used to adjust the torque of an IPMSM: field-oriented control and MTPA control. The field-oriented control sets a fixed d-axis current and adjusts the q-axis current to obtain linear proportional torque. The MTPA control sets a variable d-axis current, which is related to the q-axis current, to obtain maximum torque/ampere. The method can generate more torque than field-oriented control. However, the method creates a nonlinear relationship between the torque and q-axis current [3], which is a disadvantage of MTPA control. While it is a fact that MTPA's influence on torque linearity is a drawback, achieving maximum torque is a greater benefit that outweighs that disadvantage. The issue will be researched further by the authors of this paper.

Taking $\frac{d}{di_q} T_e = 0$ and substituting it into (47), we can obtain

$$\frac{3}{2}\frac{P}{2}[\lambda_m + (L_d - L_q) i_d - (L_d - L_q) i_q^2 \frac{1}{i_d}] = 0 \quad (48)$$

From (48), the d-axis current under MTPA control can be expressed as

$$i_d|_{MTPA} = \frac{-\lambda_m}{2(L_d - L_q)} - \sqrt{\frac{\lambda_m^2}{4(L_d - L_q)^2} + i_q^2} \quad (49)$$

The advance angle under MTPA control is

$$\theta_{MTPA} = \sin^{-1} \frac{i_d\ |_{MTPA}}{I_s} \quad (50)$$

and the current amplitude I_s is expressed as

$$I_s = \sqrt{i_d^2 + i_q^2} \quad (51)$$

In steady-state and neglecting the voltage drops of the resistance and inductance, it is not difficult to obtain

$$v_d = -\omega_{re} L_q i_q \quad (52)$$

and

$$v_q = \omega_{re} L_d i_d + \omega_{re} \lambda_m \quad (53)$$

The voltage amplitude is

$$v_s = \sqrt{(v_d^2 + v_q^2)} \quad (54)$$

If V_{om} is the maximum voltage amplitude, the voltage constraint can be shown as

$$(L_q i_q)^2 + (L_d i_d + \lambda_m)^2 = (\frac{V_{om}}{\omega_{re}})^2 \quad (55)$$

The maximum current amplitude I_{sm} can be expressed as

$$i_d^2 + i_q^2 = I_{sm}^2 \quad (56)$$

From (56), it is easy to obtain

$$i_q = \sqrt{I_{sm}^2 - i_d^2} \quad (57)$$

Combining (49), (54), and (57), the constraints of the IPMSM drive system can be obtained. The drive system operates under MTPA control when the motor speed is below the rated base-speed. After the drive system reaches its base speed, there are two constraints: the current constraint of the maximum amplitude, I_{sm}, and the voltage constraint of the maximum voltage amplitude, V_{om}. However, the $\frac{V_{om}}{\omega_{re}}$ is decreased as the motor speed increases, which is shown in Figure 2. As a result, the axes of the ellipse are reduced when the motor speed increases. When $\omega_{re} < \omega_{base}$, the motor operates under MTPA control, which is indicated as OCBA in Figure 2. When $\omega_{base} < \omega_{re} < \omega_c$, the motor operates in region II, in which the motor can operate either under MTPA control or flux-weakening control. If $i_d < i_d\ |_{MTPA}$, MTPA control is used; on the other hand, if $i_d > i_d\ |_{MTPA}$, flux-weakening control is used. When $\omega_c < \omega_{re}$, the motor operates in region III, in which only flux-weakening control can be used.

3.2. Proposed Real-Time Flux-weakening Control

In the real world, flux-weakening control is very complicated because it requires λ_m, L_d, and L_q, which are very difficult to measure [15].

To solve this difficulty, in this paper, a real-time tuning flux-weakening control is proposed. First, the voltage amplitude of the IPMSM, v_s, which is expressed in (54) is compared to the maximum voltage amplitude, V_{om}. The θ_{FW}, which is shown in Figure 3(a), is tuned as follows

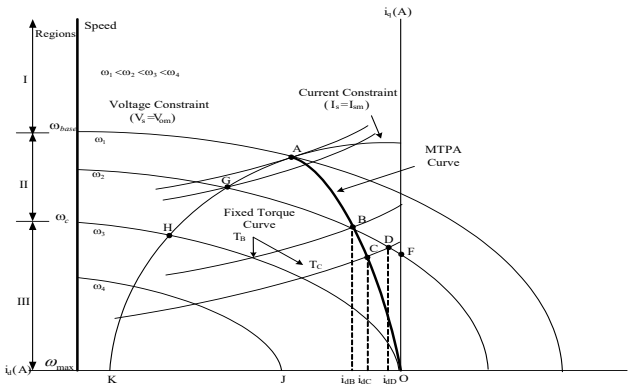


Figure 2 Operating regions of an IPMSM.

$$\theta_{FW}(\mathrm{k}) = \theta_{FW}(\mathrm{k}-1) + \alpha \cdot \mathrm{sgn}(v_s - V_{om}) \tag{58}$$

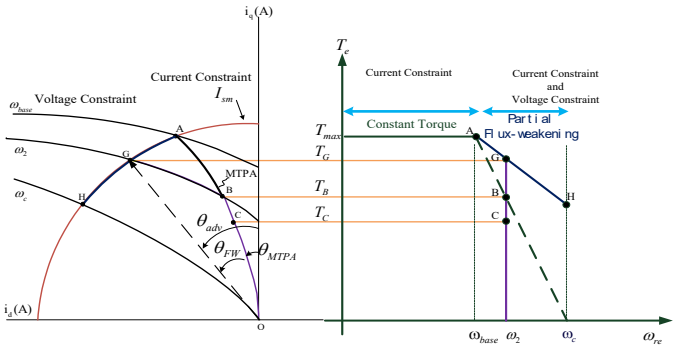
and

$$\mathrm{sgn}\ (v_s - V_{om}) = -1 \qquad \text{when } (v_s - V_{om}) < 0 \tag{59a}$$

and

$$\mathrm{sgn}\ (v_s - V_{om}) = 1 \qquad \text{when } (v_s - V_{om}) \geq 0 \tag{59b}$$

where α is the step size of the real-time tuning control. By using (58), (59a), and (59b), the θ_{FW} can be real-time tuned, as shown in Figure 3(a)(b). Finally, a block diagram of the proposed real-time tuning flux-weakening control is shown in Figure 4. First, the amplitude of the desired output voltage v_s is calculated. Next, the desired output voltage, v_s, is compared with the measured output voltage, V_{om}. After that, (58) is used to compute the $\theta_{FW}(k)$. Similar to previous research [2]-[5], [21], the proposed flux-weakening control method reduces the torque at high-speed operating range, and is not guaranteed the torque linearity. This is a common characteristic of flux-weakening control algorithms. However, the stability analysis of the whole drive system is discussed in the previous section of this paper.



(a) (b)

Figure 3 Operating curves and torque-speed curves of an IPMSM
(a) operating curves (b) torque-speed curves.

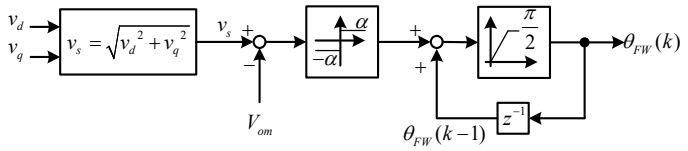


Figure 4 Block diagram of the proposed flux-weakening control.

4. Implementation

The implemented IPMSM drive system is shown in Figure 5(a) and (b). Figure 5(a) is a block diagram of the proposed IPMSM. The system consists of two main parts: the DSP system and the hardware circuit. The DSP system uses a TMS-320F-28335, which is a 32-bit, digital signal processor manufactured by Texas Instruments. The sampling interval of the current-loop control is 100 μs, while that of the speed-loop control is 1 ms.

Taking the MTPA operating region as an example, the controller works as follows. First, as shown in Figure 5(a) and based on the predictive speed controller, after the speed command is input and the motor speed is feedback, the q-axis current control i_{qp} is computed. Second, the load compensator compensates for the load disturbance $\hat{T}_D$ and then generates the compensation current $i_{q_{\hat{T}_D}}$. Then the total q-axis current command, which is the summation of the i_{qp} and $i_{q_{\hat{T}_D}}$, is computed and expressed as i_q^*. Next, the d-axis current is computed through (49) when the motor is operating in the MTPA region. After that, the advance phase angle θ_{MTPA} and stator current amplitude I_s^* are computed by using (50) and (51). Finally, the flux-weakening is obtained by adding θ_{FW}.

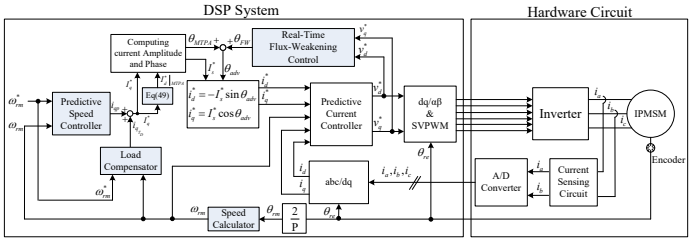
By using the MTPA table and the real-time tuning flux-weakening control, the advance angle of the current can be obtained. After that, the d-axis current command i_d^* and the q-axis current command i_q^* are calculated. A current regulated control and a space vector PWM modulation are executed. Finally, the switching states of the inverter are determined and output. A closed-loop system can thus be obtained. A DC motor is coupled with the IPMSM. The motor parameters are: $L_q = 31\ mH$, $L_d = 15\ mH$, $\lambda_m = 0.227\ v.\mathrm{sec}/rad$, and $R_s = 1.9\ \Omega$. The motor is 4-pole, 1-HP, and rated at 1500 r/min.

Figure 5(b) shows a photograph of the entire hardware circuit, which includes a DSP, gate drivers, peripheral devices, current sensing circuits, and an inverter. Figure 5(c) shows a photograph of the implemented IPMSM drive system, including an IPMSM and a dynamometer.

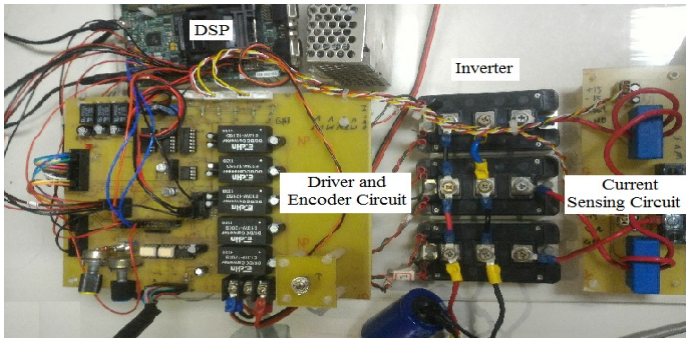
5. Experimental Results

The performance of the proposed flux-weakening control for an IPMSM drive system was evaluated according to experimental results. Figure 6(a) shows the measured steady-state torque-speed curves using the zero d-axis current control and the MTPA control. Figure 6(b) shows measured steady-state torque-speed curves

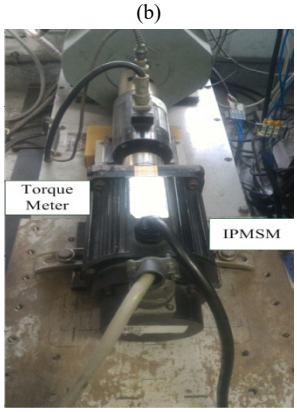
using the MTPA control and the hybrid control, which is a combination of the MTPA control and the flux-weakening control.



(a)



(b)



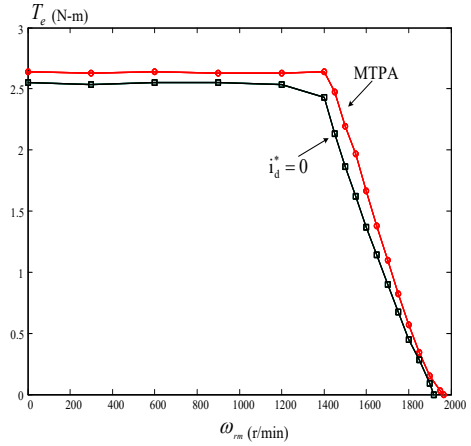
(c)

Figure 5 The implemented IPMSM drive system
(a) block diagram (b) hardware circuit (c) motor and dynamometer.

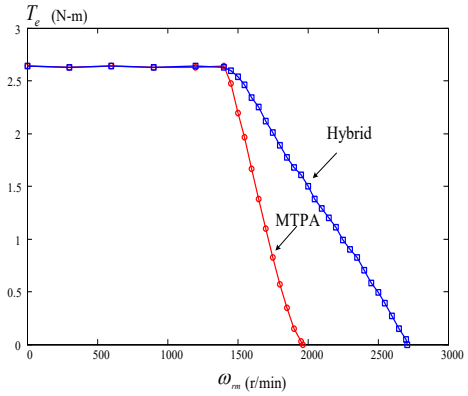
As you can observe, the hybrid control has better performance than the MTPA control. Figure 6 (c) shows the q-axis current to the d-axis current trajectory when the IPMSM is operating from the MTPA control region to the flux-weakening control region. The d-axis current increases slightly in the MTPA region; however, the d-axis current increases significantly in the flux-weakening region.

Figure 7(a)(b) and (c) show the measured waveforms of the IPMSM operating at 1800 r/min under loads of 0.57 N-m and 1.89 N-m. Figure 7(a) is the amplitude of the stator current I_s . Figure 7(b) shows the relative measured d-axis and q-axis currents. Figure 7(c) shows the measured advance angle θ_{adv} that is the sum of the θ_{FW} and θ_{MTPA}, which are the measured flux-weakening angle and the measured MTPA control angle. The delay of the measured responses is due to the lag between the load command and the external load that is generated by the dynamometer. Figure 8(a) shows the load disturbance responses when using the proposed predictive controller, the previous predictive controller that was proposed by [21], and the PI controller. The parameters of the PI controller are obtained by using the pole assignment technique. The locations of the desired poles for the closed-loop drive system are assigned as -4.06+j0.13 and -4.06-j0.13, respectively. The proposed predictive controller performs better than the previous predictive controller and the PI controller. At this 2000 r/min high speed, the field weakening algorithm is applied. The maximum allowed external load at this operating speed is 1.6 N.m. Experimental results show the proposed predictive controller performs the best. Figure 8(b) shows the repetitive 1.6 N.m load disturbance responses. The proposed predictive controller performs well again.

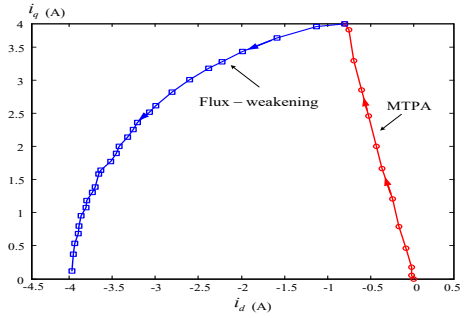
Figure 9(a) shows the measured step-input transient speed responses at the maximum speed, 2700 r/min. In this paper, the rated speed is 1500 r/min. The measured responses show the proposed predictive controller has a quicker rise time than the previous predictive controller proposed by [16] and the PI controller. Figure 9(b) shows the measured step-input transient speed responses at 3 r/min, which is the minimum operating speed. According to Figure 9(a) and Figure 9(b), we can conclude that the adjustable speed ratio of the maximum speed to minimum speed is 900:1. Figure 10 (a)(b)(c)(d) shows the current command and measured current errors using different controllers. Again, the proposed predictive controller has the smallest tracking errors. Figure 11(a)(b)(c) show the measured dynamic responses of the proposed flux-weakening control at 1800 r/min. A 5V step input at V_{om} is injected for every 1 second. The real voltage amplitude of the motor, which is expressed as v_s can track the voltage command V_{om} well. The rise time of the voltage control is about 50 ms. Figure 11(a), Figure 11(b) and Figure 11(c) show the responses of speed, voltage amplitude, and controlling angle θ_{FW} respectively. Figure 12(a) shows the measured efficiency of the whole drive system. The proposed method has higher efficiency than the linear flux-weakening method that was proposed by [21]. The major reason is that the proposed method provides greater output torque at the same motor speed. Figure 12(b) and (c) show the measured speed responses as the inertia of the drive system is increased to 5 times. The proposed predictive controller can provide a lower overshoot and a shorter settling time as the inertia increases.



(a)

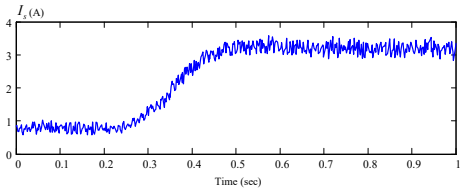


(b)

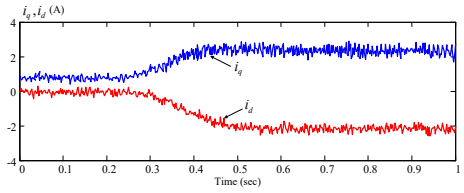


(c)

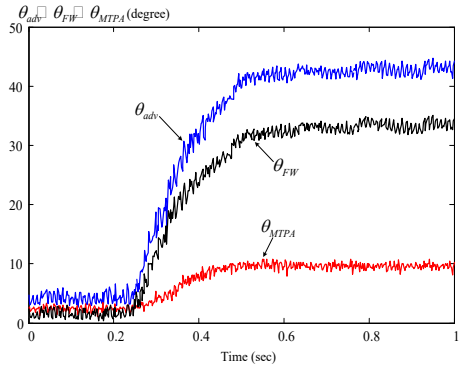
Figure 6 Measured steady-state characteristic curves
(a)MTPA and zero d-axis current (b) MTPA and hybrid control
(c) q-axis current to d-axis current curve.



(a)

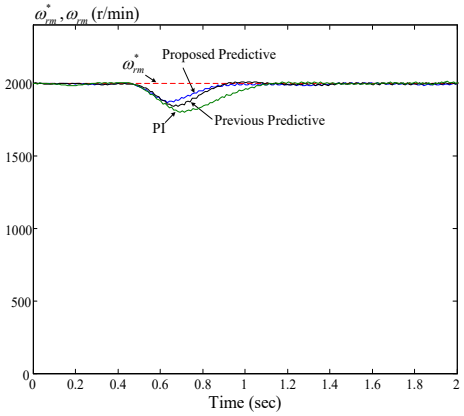


(b)

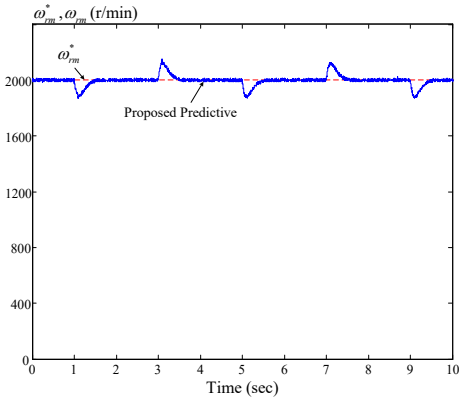


(c)

Figure 7 Measured 1800 r/min with loads of 0.57 N-m and 1.89 N-m
(a) I_s (b) measured d-q axis currents (c) advance angle.

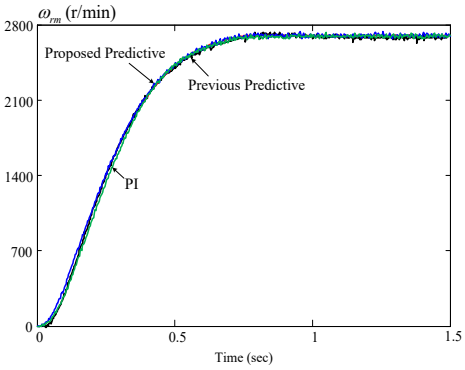


(a)

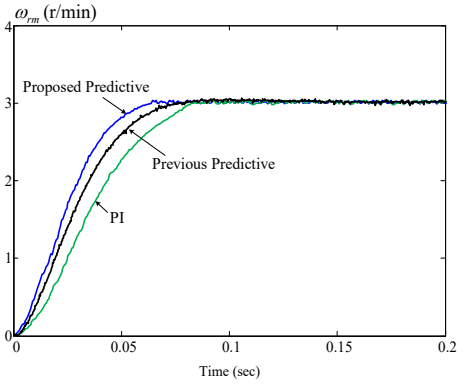


(b)

Figure 8 Measured field weakening operating speed with a large 1.6 N.m external load.
(a) step-input load (b) repetitive loads.

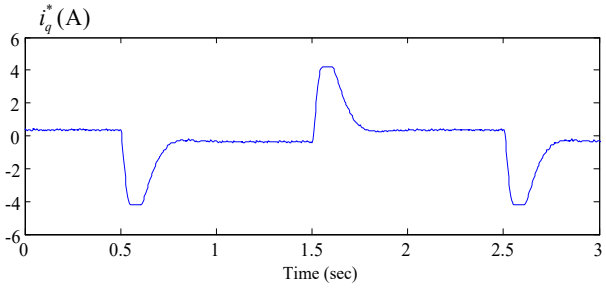


(a)

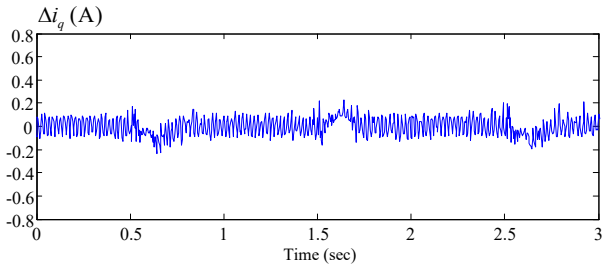


(b)

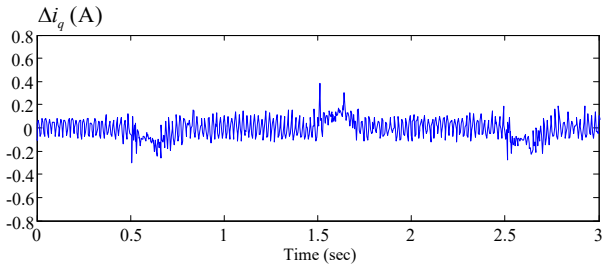
Figure 9 Measured step-input transient speed responses
(a) highest speed (b) lowest speed.



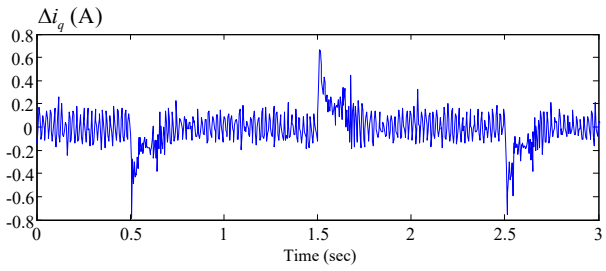
(a)



(b)

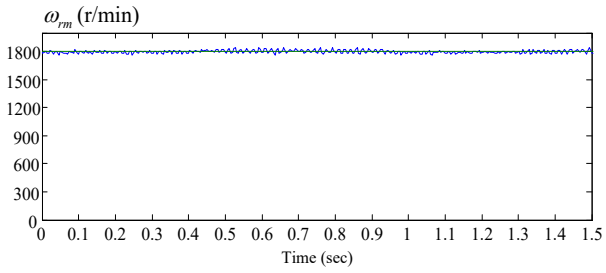


(c)

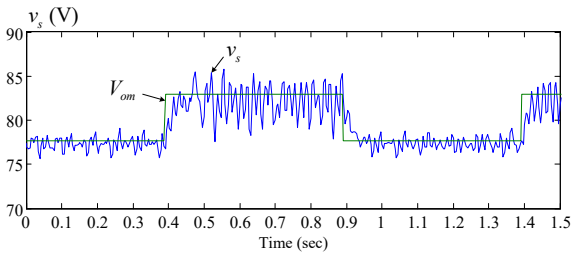


(d)

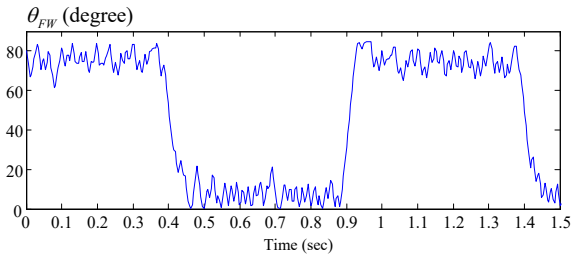
Figure 10 Comparison of measured current responses using different controllers (a) current command (b) proposed predictive (c) previous predictive (d) PI.



(a)

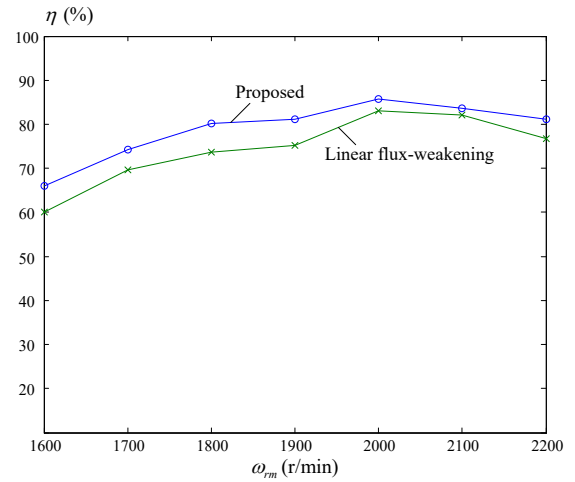


(b)

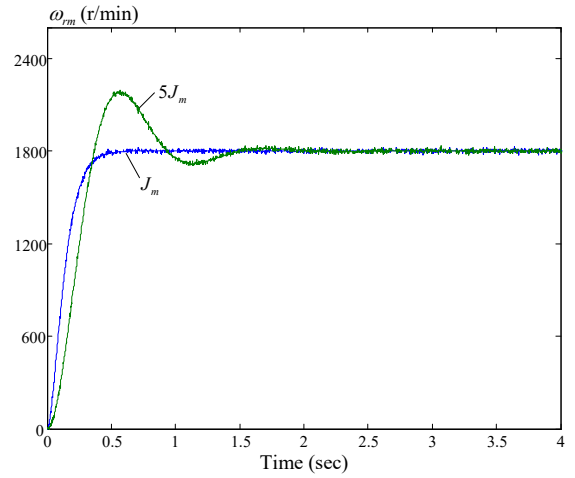


(c)

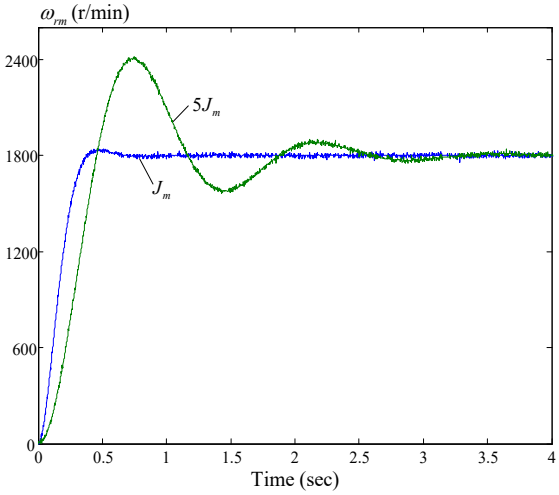
Figure 11 Measured dynamic responses of the flux-weakening control at 1800 r/min
(a) speed (b) voltage magnitude regulation (c) θ_{FW} .



(a)



(b)



(c)

Figure 12 Measured efficiency and varied inertia speed responses.
(a) efficiency (b) proposed predictive controller (c) PI controller.

6. Conclusions

In this paper, an MTPA control and a real-time flux-weakening control drive system using a predictive controller is proposed to improve the adjustable speed range and dynamic responses of an IPMSM. Experimental results show that the proposed drive system has good performance, including fast transient responses, good load disturbance rejection capability, and good tracking responses. Furthermore, the proposed method is easily implemented by using a DSP. Experimental results validate the theoretical analysis.

References

[1] T. H. Liu, S. K. Tseng, and M. B. Lu, "Auto-tunning flux-weakening control for an IPMSM drive system using a predictive controller," IEEE 26th International Symposium on Industrial Electronics (ISIE), 238-243, 2017.

[2] S. Bolognani, S. Calligaro, and R. Petrella, "Adaptive flux-weakening controller for interior permanent magnet synchronous motor drives," IEEE J. of Emerg. and Sel. Top. in Pow. Electron., **2**(2), 236-248, June 2014.

[3] H. Liu, Z. Q. Zhu, E. Mohamed, Y. Fu, and X. Qi, "Flux-weakening control of nonsalient pole PMSM having large winding inductance, accounting for resistive voltage drop and inverter nonlinearities," IEEE Trans. on Pow. Electron., **27**(2), 942-952, Feb. 2012.

[4] T. S. Kwon, G. Y. Choi, M. S. Kwak, and S. K. Sul, "Novel flux-weakening control of an IPMSM for quasi-six-step operation," IEEE Trans. on Ind. Appl., **44**(6), 1722-1731, Nov./Dec. 2008.

[5] C. T. Pan and J. H. Liaw, "A robust field-weakening control strategy for surface-mounted permanent-magnet motor drives," IEEE Trans. on Energy Conver., **20(**4), 701-709, Dec. 2005.

[6] M. N. Uddin and M. M. I. Chy, "Online parameter-estimation-based speed control of PM AC motor drive in flux-weakening region," IEEE Trans. on Ind. Appl., **44(**5), 1486-1494, Sep./Oct. 2008.

[7] S. M. Sue and C. T. Pan, "Voltage-constraint-tracking-based field-weakening control of IPM synchronous motor drives," IEEE Trans. on Ind. Electron., **55(**1), 340-347, Jan. 2008.

[8] S. Y. Jung, C. C. Mi, and K. Nam, "Torque control of IPMSM in the field-weakening region with improved DC-link voltage utilization,' IEEE Trans. on Ind. Electron., **62**(6), 3380-3387, June 2015.

[9] S. Chatithongsuk, B. Nahid-Mobarakeh, J. P. Caron, N. Takorabet, and F. Meibody-Tabar, "Optimal design of permanent magnet motors to improve field-weakening performances in variable speed drives," IEEE Trans. on Ind. Electron., **59**(6), 2484-2494, June 2012.

[10] K. C. Kim, "A novel magnetic flux weakening method of permanent magnet synchronous motor for electric vehicles," IEEE Trans. on Magn., **48**(11), 4042-4045, Nov. 2012.

[11] V. R. Jevremovic and D. P. Marcetic, "Closed-loop flux-weakening for permanent magnet synchronous motors," Conference on Power Electronics, Machines and Drives, PEMD 2008, 4th IET, 717-721, Apr. 2008.

[12] L. Harnefors, K. Pietilainen and L. Gertmar, "Torque-maximizing field-weakening control: design, analysis, and parameter selection," IEEE Trans. on Ind. Electron., **48**(1), 161-168, Feb. 2001.

[13] O. Wallmark, "On control of permanent-magnet synchronous motors in hybrid-electric vehicle applications," Technical Reports at the School of Electric Engineering, Technical Report, no. 495L, Chalmers University of Technology, Master Thesis, 2004

[14] N. Bedetti, S. Calligaro and R. Petrella, "Analytical design of flux-weakening voltage regulation loop in IPMSM drives," IEEE Energy Conversion Congress and Exposition (ECCE), ECCE 2015, 6145-6152, Sep. 2015.

[15] M. Pacas and J. Weber, "Predictive direct torque control for the PM synchronous machine," IEEE Trans. on Ind. Electron., **52**(5), 1350-1356, Sep. 2005.

[16] S. Bolognani, S. Bolognani, L. Peretti, and M. Zigliotto, "Design and implementation of model predictive control for electrical motor drives," IEEE Trans. on Ind. Electron., **56**(6), 1925-1936, June 2009.

[17] J. M. Kim and S. K. Sul, "Speed control of interior permanent magnet synchronous motor drive for the flux weakening operation," IEEE Trans. on Ind. Appl., **33**(1), 43-48, Jan./Feb. 1997.

[18] E. J. Fuentes and C. A. Silva, "Predictive speed control of a two-mass system driven by a permanent magnet synchronous motor," IEEE Trans. on Ind. Electron., **59**(7), 2840-2848, July 2012.

[19] M. Siami, D. A. Khaburi, A. Abbaszadeh, and J. Rodriguez, "Robustness improvement of predictive current control using prediction error correction for permanent-magnet synchronous machines," IEEE Trans. on Ind. Electron., **63**(6), 3458-3466, June 2016.

[20] T. Tarczewski and L. M. Grzesiak, "Constrained state feedback speed control of PMSM based on model predictive approach," IEEE Trans. on Ind. Electron., **63**(6), 3867-3875, June 2016.

[21] K. Chen, Y. Sun, and B. Liu, "Interior permanent magnet synchronous motor linear field-weakening control," IEEE Trans. on Energy Conver., **31**(1), 159-164, Mar. 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 87-93 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Discovering Interesting Biological Patterns in the Context of Human Protein-Protein Interaction Network and Gene Disease Profile Data

Rami Alroobi[*,1], William Perizzo[2]

[1]*Department of Math and Computer Science, Southern Arkansas University, 71753, USA*

[2]*Department of Computer Science, North Dakota State University, 58105, USA*

A R T I C L E I N F O

Article history:
Received: 11 October, 2017
Accepted: 08 November, 2017
Online :16 November, 2107

Keywords:
Organisms
Bioinformatics
Functional modules
Interaction networks
Disease
Expression profiles

A B S T R A C T

The current advances in proteomic and transcriptomic technologies produced huge amounts of high-throughput data that spans multiple biological processes and characteristics in different organisms. One of the important directions in today's bioinformatics research is to discover patterns of genes that have interesting properties. These groups of genes can be referred to as functional modules. Detecting functional modules can be accomplished by the deep analysis of protein-protein interaction (PPI) networks, gene expression profiles, or both. In this work the focus will be on Human protein-protein interaction network and genes expression data that represents genes behavior in a group of diseases. Two of the most well-established clustering methods that target the interaction networks and the expression data will be used in this analysis. In addition, and to have more insights, genes molecular functionality will be studied. Finally, I will introduce the relation of the extracted modules on biological pathways. This study mainly illustrates the importance of including protein interaction activities as part of any study that aims at discovering meaningful knowledge about the biological scene where many actors play different roles.

1. Introduction

The current advances in proteomic and transcriptomic technologies produced huge amounts of high-throughput data that spans multiple biological processes and characteristics in different organisms. One of the important directions in today's bioinformatics research is to discover patterns of genes or genes products that have interesting characteristics that can be related to specific activities or functions inside the living organism. Researchers believe that those groups of genes are can give deeper insights of what is happening inside a living cell than studying individual genes in isolation. These groups of genes are called functional modules (patterns). Different approaches can be employed to discovering functional modules. One of the most important approaches is to focus on protein-protein interaction (PPI) networks and try to find well connected sub-networks. Another approach is to consider gene expression data and try to find modules that show similar behavior according to their expression levels by calculating correlations. The last approach is to integrate both the PPI and the expression data in a complementary effort because both PPI and expression data suffer from incompleteness and inconsistency problems [1].

However, for the expression profiles that are related to diseases in specific, most of the available data is in the form of genome-wide expression profiles that extracted from microarray measurements. A number of methods have been proposed to clarify the biological mechanisms from this expression data [2,3]. Most of those methods are based on individual genes ability to be used as a strong indication of the disease under consideration. However, considering only individual genes will give limited insights about the molecular mechanisms and biology of an organism under different disease conditions. On the other hand, studying groups of genes that show specific behavior related to some disease will help in grasping more comprehensive views about the disease itself and what molecular functionality those genes may have and can be involved in this disease. The reason for that is the fact that any biological process, such as a disease, occurring inside the living organism is affected and affects multiple and different biological components and pathways of that organism. Patterns of genes,

*Corresponding Author: Rami Alroobi, Magnolia, Ar 71753, 870-235-4295, Email: rmalroobi@saumag.edu

or functional modules, can be discovered in different ways. One approach is by applying clustering techniques on PPI networks to find patterns of well-connected genes. One of the most established methods used in this field is the stochastic flow-based clustering (Markov Cluster Algorithm, MCL) [4, 5]. The Markov cluster algorithm (MCL) uses a graph's natural transition probability matrix to cluster a graph by combining random walks with two alternating operations (expansion and inflation). With these operations, the algorithm iteratively simulates the dissipation and elimination of stochastic flow across the graph structure to produce a non-negative matrix representation of a graph clustering. In other words, the nodes of the graph are clustered with each other based on the strength of connections between these nodes. The MCL method has been shown to be significantly robust and superior to other network-based methods when used to cluster benchmarking biological networks [6].

Another approach is to use gene expression profiles to cluster genes that show similar behavior based on their expression levels that are extracted from microarray measurements. In this case, any clustering method can be used like K-means [7]. K-means and its variations are among the most popular iterative methods used for clustering data and can be easily applied to biological data. The general idea of K-means is that a number of centers are specified and a distance metric from these centers is used as a dissimilarity measure. The points are grouped towards the closest center based on the criteria used. For example, one of the widely used metrics is the Euclidean Distance. The Euclidean Distance [8] between values x_i and $x_{i'}$ can be defined as follows:

$$d(x_i, x_{i'}) = \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = ||x_i - x_{i'}||^2$$

Then each of the points, genes in this case, will be assigned to cluster so that the distance between these points and the specified center is minimized. After that, a new set of centers are calculated and the assignment step is repeated. The process continues iteratively until the assignments stabilize so that no further assignments changes are possible.

2. Experimental Analysis

2.1 Datasets

In order to elucidate the work in this paper, a group of data sets were obtained and preprocessed in order to suite the setting required for this work. The following is a listing of the datasets used in the analysis.

- **Interaction Networks**
  The experiments were performed using Homosapiens (Humans) protein-protein interaction network. The Human network was obtained from the Human Protein Reference Database (HPRD), release 9 [9]. The Human PPI network has 9465 nodes with 37039 interactions.

- **Disease Gene profiles**
  In order to perform the planned analysis, gene expression profiles for 16 diseases were obtained from the GEO database [10]. Table 1 shows the names and identification numbers of these diseases.

- **OMIM dataset**
  The OMIM, Online Mendelian Inheritance in Man, database [11] was used to extract genes-disease data. The data includes 4022 different diseases. Each disease is represented by the genes causing it or in relation to that disease.

- **Human protein complexes**
  The human protein complexes are groups of proteins that have strong evidence that they are interact with each other. The protein complexes were extracted from HPRD (Human Protein Reference Database), release 9 [9]. This data set contains 1521 of manually curated protein complexes.

- **Molecular Functionality**
  For studying Molecular functions of the Human genes, the Molecular Functions from the GO annotation database were used (release date of the data base: 11/15/2011, version: 1.216) [12].

2.2 Performing Experiments

After obtaining all the datasets needed for this work, both MCL and K-means were employed to discover the interesting patterns.

Firstly, the MCL algorithm was applied to the HPRD interaction network. MCL uses a parameter that controls its operation. That parameter is called inflation. I used an inflation of 1.75 which is considered optimal according to [6]. I only considered the resulting patterns of sizes greater than or equal to 4 genes. The MCL produces connected modules when applied to the interaction network. The result contains 776 modules.

Table 1. Disease Datasets for expression profiles

GEO Series ID	GEO DataSet ID	Disease Name
GSE2503	GDS2200	Actinic keratosis
GSE1420	GDS1321	Adenocarcinoma of esophagus
GSE1297	GDS810	Alzheimer's disease
GSE5388	GDS2190	Bipolar disorder
GSE475	GDS289	Chronic obstructive lung disease
GSE1462	GDS1065	Chronic progressive ophthalmoplegia
GSE1629	GDS1850	Complex dental cavity
GSE3585	GDS2205	Congestive cardiomyopathy
GSE3365	GDS1615	Crohn's disease
GSE5370	GDS2153	Dermatomyositis
GSE2006	GDS1376	Essential thrombocythemia
GSE1751	GDS1331	Huntington's disease
GSE2018	GDS999	Lung transplant rejection
GSE3189	GDS1375	Malignant melanoma
GSE3868	GDS1746	Malignant neoplasm of prostate
GSE2549	GDS1220	Malignant pleural mesothelioma

2.3 Performing Experiments

After obtaining all the datasets needed for this work, both MCL and K-means were employed to discover the interesting patterns.

Firstly, the MCL algorithm was applied to the HPRD interaction network. MCL uses a parameter that controls its

operation. That parameter is called inflation. I used an inflation of 1.75 which is considered optimal according to [6]. I only considered the resulting patterns of sizes greater than or equal to 4 genes. The MCL produces connected modules when applied to the interaction network. The result contains 776 modules.

Secondly, the K-means algorithm was applied to the gene disease expression profile data. In order to have comparable number of modules to the one produced by the MCL, the number of modules resulting from the MCL was used as the number of K-means initial centers (clusters). Then, the analysis was performed with clusters of sizes greater than 4 genes (similar to what was done in MCL). K-means produced 1039 module of size $\geq$ 4 genes. The genes in these modules show similar behavior according to the disease expression data used in this study.

2.4 Human protien complexes

The following analysis was performed to assess the quality of the MCL and K-means produced modules in predicting known protein complexes. Those complexes are proved to have biological significance. Therefore, more matched complexes by the produced patterns means higher quality patterns from the biological point of view.

For assessing complex matching, I used the matching criteria introduced by [13]. The overlap score was computed by the formula $w = i^2/a*b$, where i is the size of the intersection set between the discovered patterns and the known complex, a is the size of the discovered protein pattern and b is the size of known protein complex. The overlap threshold between produced patterns and the protein complexes ranged from 0.1 to 1.0 with 0.1 increments; 1.0 implies a 100% match between the module and the protein complex. The range was chosen to make the overlapping abilities of the discovered patterns a tangible amount illustrating the quality of these patterns. Furthermore, we have not considered any overlap to be considered as a match which would may be seen as an overstatement of these results. More overlap indicates a better match with protein complexes, however knowing that more protein complexes are yet to be identified, even a lower match percentage can highlight some promising results. Figure 1 illustrates the MCL performance. While Figure 2 shows the performance of K-means prediction capability. Clearly, MCL has superiority over the K-means and was able to predict protein complexes even with higher overlap thresholds. This can be referred to the fact that many of the complexes are known to be well connected sub-graphs originally, and MCL aims at producing connected modules. Furthermore, K-means does not care about patterns connectivity. It is only concerned of expression similarity.

Another note about the results that I found interesting is that the larger the produced pattern the more penalty it will face in the complex matching process. This is because the average size of protein complexes is $\approx$ 5 genes and the criteria used for matching penalizes large patterns.

2.5 OMIM disease modules

The OMIM dataset provides a comprehensive collection of diseases that relates a large number of known diseases to the genes that are causing them or have a strong involvement in the cause and the mechanisms of the disease, in other words the Etiology and Pathogenesis of a disease. To evaluate the produced patterns and to have more insights that they can provide in the disease domain, both MCL and K-means produced patterns were tested against the OMIM disease dataset. This analysis aimed at discovering modules that might be related to known diseases.

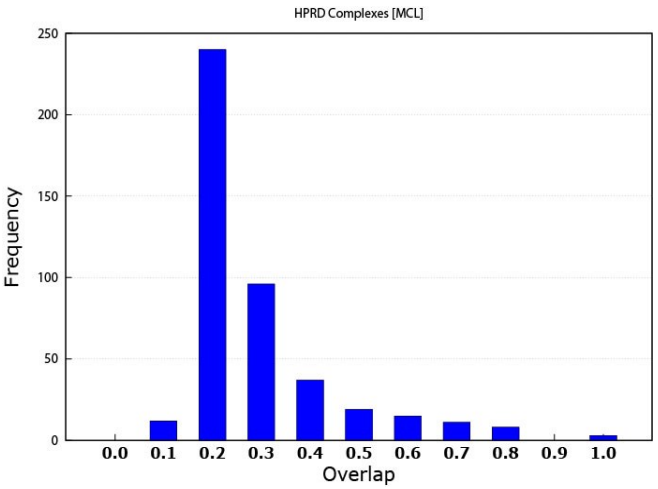


Figure 1. MCL protein complex prediction

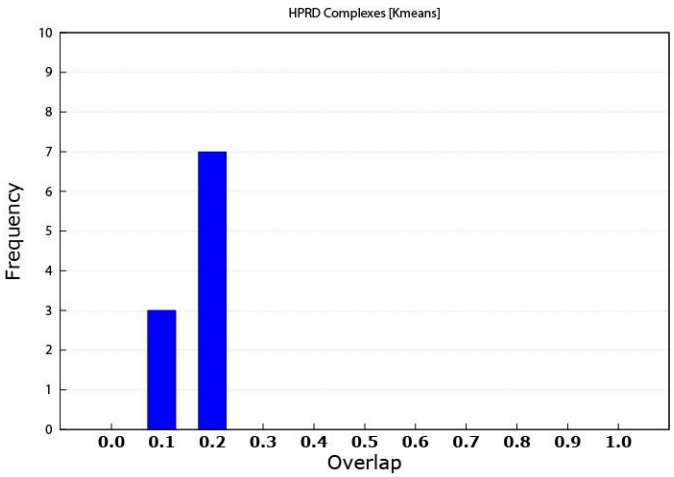


Figure 2. K-means protein complex prediction

As expected, the K-means slightly outperformed the MCL in this case. In this work, the patterns in the OMIM dataset are referred to as disease modules.

MCL produced patterns matched 68 disease modules while K-means matched 134 disease modules. The reason is that K-means patterns were originally based on data that is disease related. Some of the diseases that were used to create expression profile dataset were among the ones that were found from the OMIM dataset. As mentioned above, this study used only 16 diseases expression data, while the resultant matched diseases were multiples of that number. An explanation is that some genes can be involved in multiple diseases. This note means that some genes play the rule of a link between different diseases.

Thus, targeting these specific genes can illuminate hidden information that might lead interesting results; the cure for example. However, this statement needs more investigation by intensive research that is out of the scope of this work. In addition, a produced pattern can match more than one disease.

Figure 3 shows the performance of MCL when used with the OMIM dataset. Figure 4 is similar but for K-means. Table 2 and Table 3 show some of the disease matched by the MCL and K-means modules respectively. More specifically, Table 3 shows that some produced modules by K-means has matched multiple types of disease.

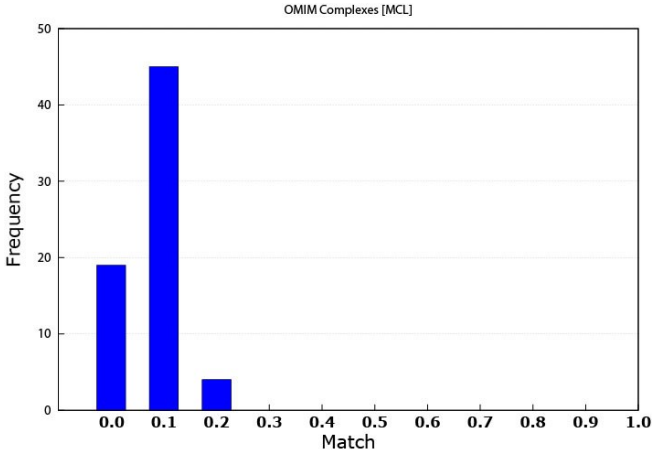


Figure 3. MCL matching performance against the OMIM data (0.0 means that the threshold is ≤ 0.1)

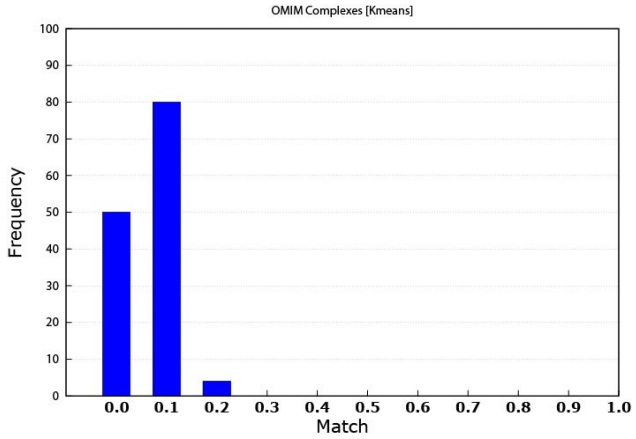


Figure 4. K-means matching performance against the OMIM data (0.0 means that the threshold is ≤ 0.1)

Table 2. Some of the MCL matched disease

Module's Genes	Related Disease
CDK2AP1 CDK2AP2 WARS PC GOT2 MDH2 CS ZDHHC6 FH NDUFS1	Thrombophilia due to protein C deficiency autosomal dominant
FANCC FANCA FANCE FANCG FANCF HES1 SAMD3 CYP19A1	Fanconi anemia complementation
GORASP2 RAB2A BLZF1 MIF NQO2 KCTD5 FAM71C	Persistent Mullerian duct syndrome type I
LYST SEMA4C DGCR14 MRPL17 NCDN DNAJC14 CNTROB	Chediak-Higashi syndrome
SHH PTCH1 PTCH2 IHH SMO DHH HHIP	Basal cell carcinoma somatic
NSF GABBR1 NAPG ATF5 GABBR2 PTPN9	Nicotine dependence susceptibility
CYP17A1 POR CYP2C19 CYP1A2 CYP2C9 CYP2E1	Mephenytoin poor metabolizer
NHEJ1 LIG4 XRCC4 APLF IRX5	Multiple myeloma resistance, Severe combined immunodeficiency with sensitivity to ionizing radiation
KDR ITGA9 FLT4 FIGF VEGFC	Hemangioma capillary infantile somatic
DHX9 SERPINB2 C6 MGEA5 PRPF8	Combined C6/C7 deficiency

3. Molecular functionality

As a further step, the Molecular Functionality (MF) of the genes can be considered as a useful source of information in many aspects. First, knowledge about the MF of the genes in the produced patterns is informative especially if some of the genes in the pattern are not well studied and their functions are not well known, but the association in the pattern may result in new data about these genes. Second, the combination between the MF of the patterns and the disease expression data can lead to meaningful indications about the disease mechanisms inside the living organism, the human in this case. Molecular Functions Gene Ontologies were extracted from the Gene Ontology project [12]. Studying the resulting modules showed that modules span multiple molecular functionalities and they did not show any tendency towards some particular biological processes. However, modules gene members have shown high similarities in what biological processes they share.

Table 3. Some of the K-means matched disease

Module's Genes	Related Disease
PKD2 IGBP1 TXNL4A NID1	Polycystic kidney disease
PKD2 IGBP1 TXNL4A NID1	Corpus callosum agenesis of with mental retardation ocular ,coloboma and micrognathia
SLC9A1 ADAM2 GMPS HIVEP2	Leukemia acute myelogenous
PCSK7 TES BSN COL10A1	Metaphyseal chondrodysplasia Schmid
PHEX INSL3 WDR61 APPL1	Cryptorchidism idiopathic
SI PDX1 BNIP2 DNAJA2	Sucraseisomaltase deficiency congenital
SI PDX1 BNIP2 DNAJA2	Lacticacidemia due to PDX1 deficiency
PTRF MPP3 TSPAN4 SERPIND1 GRIP1	Thrombophilia due to heparin cofactor II deficiency
AKAP4 MDK SH2D1A SLC12A3 RPP40	Mesomelic dysplasia Kantaputra
GADD45G TRO NPR1 BACH1	Breast cancer early-onset

Figure 5 has an example module that is enriched in a number of multiple molecular functions such as, GO:0008601: Modulation of the activity of the enzyme protein phosphatase type 2A. GO:0019888: Modulates the activity of a protein phosphatase, an enzyme which catalyzes of the removal of a phosphate group from a protein substrate molecule GO:0019208: Modulates the activity of a phosphatase, an enzyme which catalyzes of the removal of a phosphate group from a substrate molecule. GO:0005488: The selective, non-covalent, often stoichiometric, interaction of a molecule with one or more specific sites on another molecule. GO:0004512: Catalysis of the reaction: D-glucose 6- phosphate = 1D-myo-inositol 3-phosphate. This reaction requires NAD, which dehydrogenates the CHOH group to CO at C-5 of the glucose 6-phosphate, making C-6 into an active methylene, able to condense with the aldehyde at C- 1. Finally, the enzyme-bound NADH reconverts C-5 into the CHOH form. GO:0005515: Interacting selectively and non-covalently with any protein or protein complex. The above results were obtained using The Database for Annotation, Visualization and Integrated Discovery (DAVID) [14] [15].

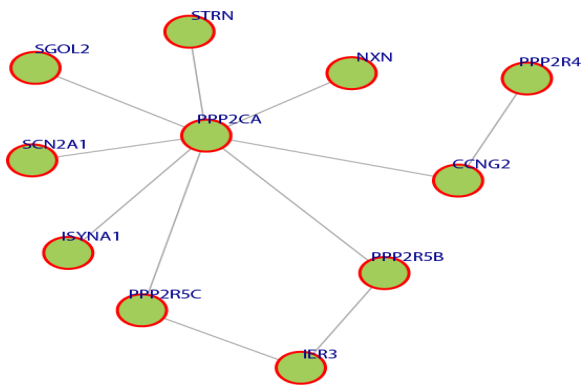


Figure 5. A module that was extracted from the network where all the genes have similar molecular functionality.

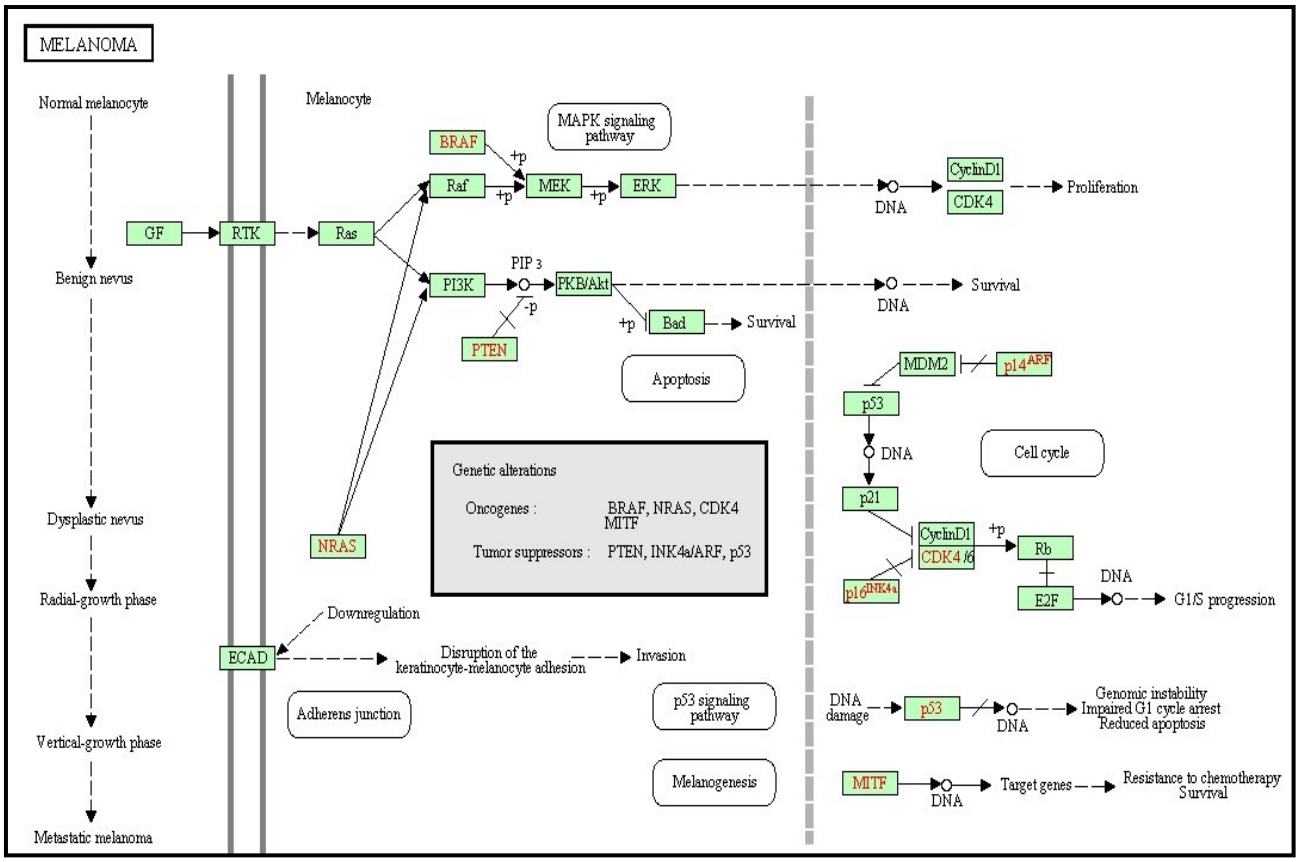


Figure 6. The prostate cancer pathway.

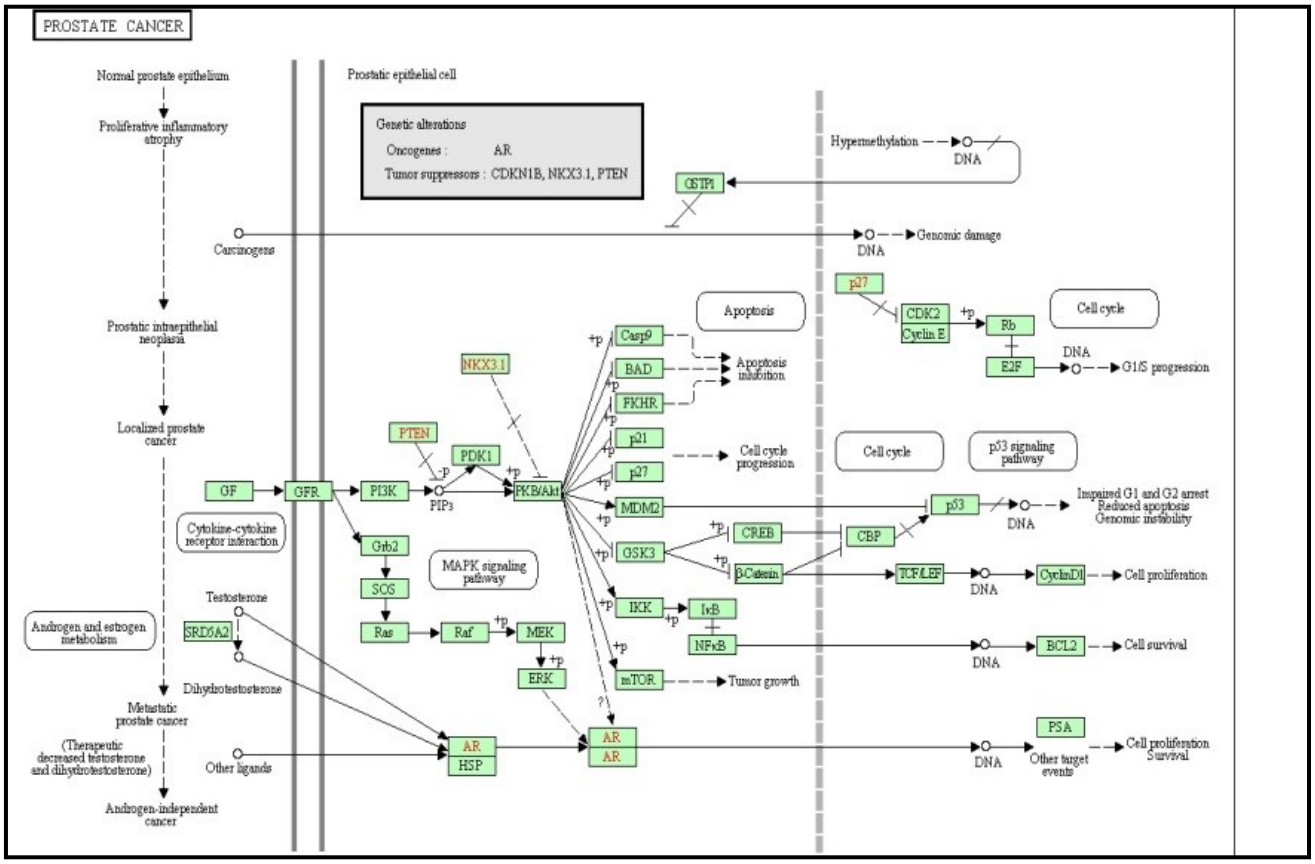


Figure 7. The Melanoma pathway.

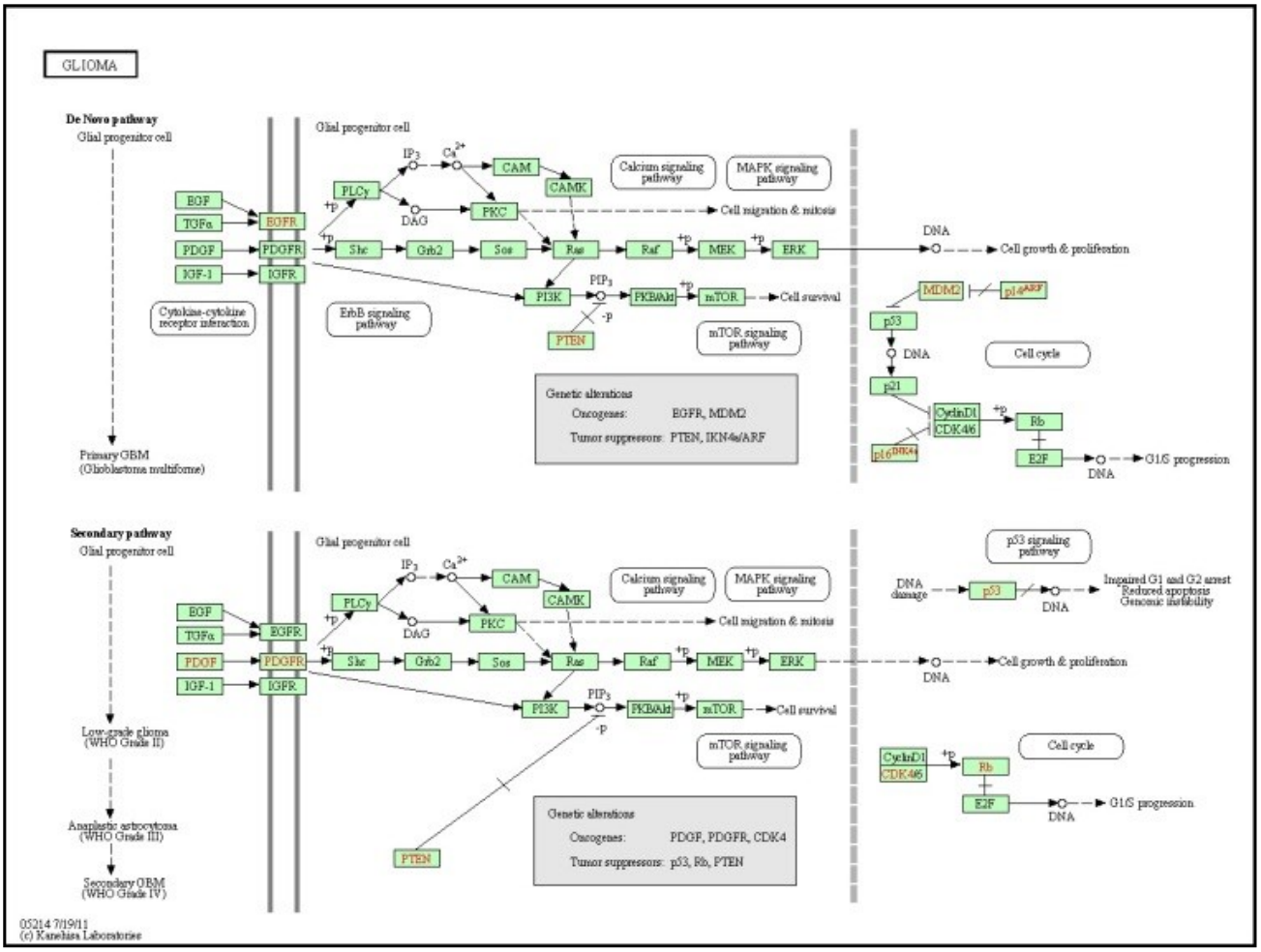


Figure 8. The Glioma pathway.

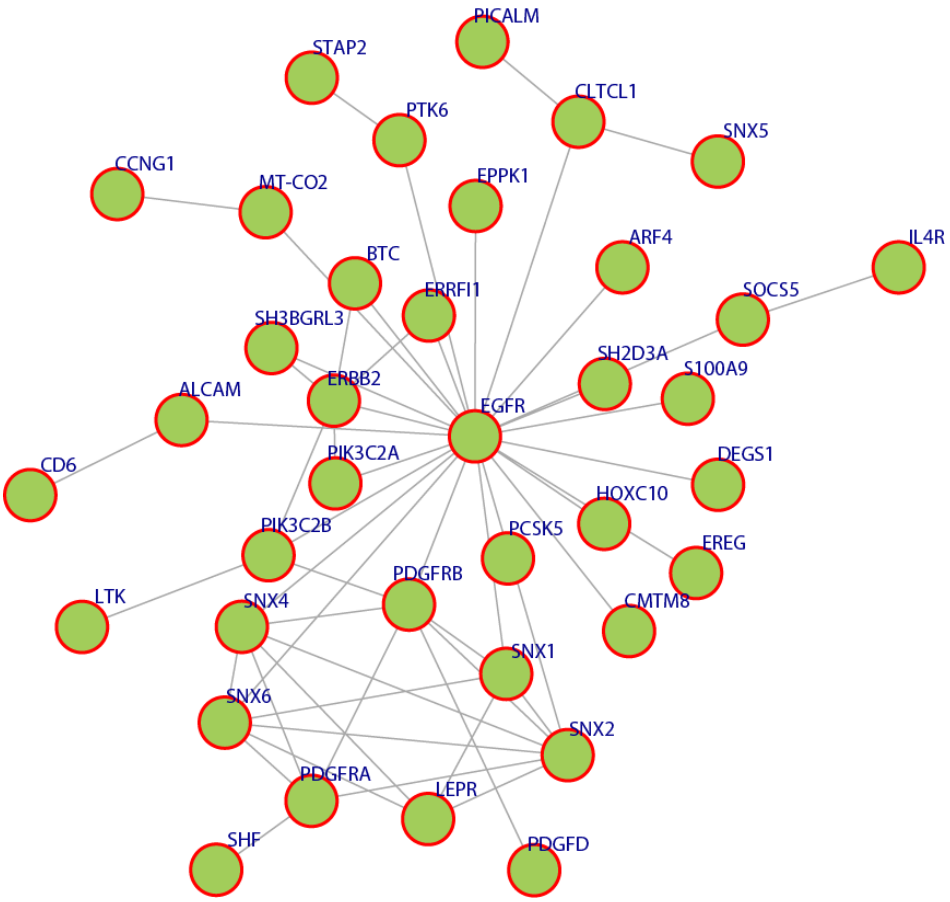


Figure 9. This discovered pattern above shows a module that is overlapped with multiple disease pathways

4. Pathways

Many of the discovered modules have biological pathways associations. Biological pathway is a group of actions that occur between the molecules in the cell that result in the formation of a new product or change in the cell. This association with pathways was checked against the Kyoto Encyclopedia of Genes and Genomes (KEGG) [16]. To illustrate the idea further, some example pathways chosen were very interesting in that they have a relation with different types of diseases. Figure 6, 7, 8, and 9 have some examples of disease-related pathways like: Prostate cancer pathway, Glioma pathway, and Melanoma pathway. In addition, another example (not shown here) matched the Notch Signaling Pathway. Notch signaling has been related to a number of malignancies including leukemia, lymphomas and carcinomas of the breast, skin, lung, cervix and kidneys [9]. It is clear from Figure 6, that one of the discovered modules has several genes that are highly incorporated in multiple disease pathways. Furthermore, we can see that not all of the module genes are present in those pathways; this raises an interesting question; can future research prove that those absent genes are related to the same pathway and therefore to the diseases because they were clustered with several of the pathway genes?

5. Conclusion

The availability of large amounts of microarray data in, both interaction networks and expression profiles, has been used to have deeper insights of what biological processes are taking place inside an organism. Much of this data is related to different aspects of the living cell activities. In addition, this expression data can be extracted from organism or tissues under a group of experimental, environmental, or any kind of stress conditions. Diseases expression data can be considered a very useful source of information regarding what genes are involved in or affected by them. In this study I tried to shed some lights about modules extracted from two separate sources of information and what relations they may have with known diseases and try to uncover some of new connections between genes and diseases. Both of the two methods employed here, MCL and K-means, have its strength sides and weakness sides. When it comes to interaction networks, MCL was superior. However, K-means has shown acceptable performance with expression profile data. The method has produced interesting patterns when they are tested against several biological concepts such as protein complexes, disease modules, molecular functionality, and biological pathways.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

I would like to thank the reviewers of this work for their help. This work was accomplished by the support of Southern Arkansas University, College of Science and Engineering.

References

[1] Y. Fang, M. Sun, G. Dai and K. Ramain, "The Intrinsic Geometric Structure of Protein-Protein Interaction Networks for Protein Interaction Prediction," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 1, pp. 76-85, Jan.-Feb. 1 2016. doi: 10.1109/TCBB.2015.2456876

[2] M . Bansal, V. Belcastro, A. Ambesi-Impiombato, D. di Bernardo." How to infer gene networks from expression profiles". Mol Syst Biol 3: 78, 2007. https://dx.doi.org/10.1038%2Fmsb4100120

[3] I. Ulitsky, I. Gat-Viks, R. Shamir. "MetaReg: a platform for modeling,analysis and visualization of biological systems using large-scale experimental data". Genome Biol 9: R1, 2008. https://dx.doi.org/10.1186%2Fgb-2008-9-1-r1

[4] S. van Dongen. "A new cluster algorithm for graphs". Report No. INS-R0012, Center for Mathematics and Computer Science (CWI), Amsterdam, 2000.

[5] S. van Dongen. "Graph clustering via a discrete uncoupling process". SIAM J Matrix Anal Appl, 30(1): 121–141, 2008. https://doi.org/10.1137/040608635

[6] S. Broh´ee and J. van Helden. "Evaluation of clustering algorithms for protein-protein interaction networks". BMC Bioinformatics, 7(488), 2006. https://doi.org/10.1186/s12859-014-0445-4

[7] J. B. MacQueen. "Some methods for classification and analysis of multivariate observations" Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967. https://projecteuclid.org/euclid.bsmsp/1200512992

[8] L. Liberti, C. Lavor, N. Maculan, and A. Mucherino. "Euclidean Distance Geometry and Applications". SIAM Review. 56:1, 3-69. 2014. DOI10.1137/120875909.

[9] L. Baolin and H. Bo. "Hprd: a high performance rdf database". In Proceedings of the 2007 IFIP international conference on Network and parallel computing , NPC'07, pages 364–374, Berlin, Heidelberg, 2007. Springer-Verlag. https://doi.org/10.1007/978-3-540-74784-0_37

[10] T. Barrett , D.B. Troup , S. E. Wilhite , P. Ledoux , D Rudnev, et al. "NCBI GEO: archive for high-throughput functional genomic data". Nucleic Acids Res 37: D885–890. 2009. https://doi.org/10.1093/nar/gkn764

[11] "Online Mendelian Inheritance in Man, OMIM". McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). 2011. URL: http://omim.org/. https://doi.org/10.1093/nar/gki033

[12] The gene ontology consortium. "Gene ontology: tool for the unification of biology". Nat. Genet., 25(1):25–29, May 2000. http://dx.doi.org/10.1038/75556

[13] G. D. Bader and W.V. Christopher. "An automated method for finding molecular complexes in large protein interaction networks". BMC Bioinformatics, 4 (2), 2003. https://dx.doi.org/10.1186%2F1471-2105-4-2

[14] D. W. Huang, B. T. Sherman, R. A. Lempicki. "Systematic and integrative analysis of large gene lists using david bioinformatics resources". Nat. Protocols, 4(1754-2189), 2008/12//print. http://dx.doi.org/10.1038/nprot.2008.211

[15] D. W. Huang, B. T. Sherman, and R. A.Lempicki. "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists". Nucleic Acids Research, 37(1):1–13, 2009. https://dx.doi.org/10.1093/nar/gkn923

[16] M. Kanehisa, S. Goto, M. Furumichi, M. Tanabe, and M. Hirakawa. "Kegg for representation and analysis of molecular networks involving diseases and drugs". Nucleic Acids Research, 38(suppl 1):D355–D360, 2010. https://dx.doi.org/10.1093/nar/gkp896

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 94-99 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

Modelling and Analysis of Radial Flux Surface Mounted Direct-Driven PMSG in Small Scale Wind Turbine

Theint Zar Htet *, Zhengming Zhao, Qing Gu, Jing Li

State Key Lab of Power Systems, Dept. of Electrical Engineering, Tsinghua University, Beijing,100084, China.

ARTICLE INFO

Article history:
Received: 13 October, 2017
Accepted: 02 November, 2017
Online: 15 November,2017

Keywords :
Airgap flux density
Cogging Torque
Maxwell 2D
PMSG
Wind Turbine

ABSTRACT

This paper presents the modelling and analysis of permanent magnet synchronous generator (PMSG) which are used in direct driven small scale wind turbines. The 3 kW PM generator which is driven directly without gear system is analyzed by Ansoft Maxwell 2D RMxprt. The performance analysis of generator includes the cogging torque in two teeth, induced coil voltages under load, winding current under load, airgap flux density distribution and so on. The modelling analysis is based on the 2D finite element techniques. In an electrical machine, an accurate determination of the geometry parameters is a vital role. The proper performance results of 3kW PMSG in small scale wind turbine can be seen in this paper.

1. Introduction

In this century, wind power generation has become the most powerful influences in all over the world[1]. The variable speed wind turbines have been more popular in nowadays because it has some advantages compare to fixed speed wind turbines such as less mechanical stress, higher efficiency, higher power quality and so on[2].

The several advantages of variable speed wind turbines are reduced mechanical stress and optimized power capture.In variable speed wind turbines, two types of generators are usually employed in the small scale wind turbine such as induction generators (IGs) and permanent magnet synchronous generators (PMSGs)[3].

Recently, in the WTs, the PMSGs have been more attractive the IGs because of direct drive and high efficiency. Moreover, the PMSGs have the high torque density and the power factor characteristics. The direct-drive train wind power on system generation can be seen in Figure 1.Four main components are contained in this system and the PMSG connects to the grid through a back to back converter, converting the mechanical energy produced by the wind turbine to the electrical energy[4].

Two type of rotor position can be used in PMSG family such as outer rotor and inner rotor. In this study, the small outer rotor surface mounted generator is adopted for modelling and analysis because of the easier construction. In conventional machines, the magnetic field is in the radial direction. These types of machines are referred to as radial flux machines (RFM) whereas the magnetic field of axial flux machine (AFM) is in the axial direction [5].

The small scale generators can be used for remote areas. The American Wind Energy Association (AWEA) defines small scale wind generator as those whose power reaches up to 100 kW.

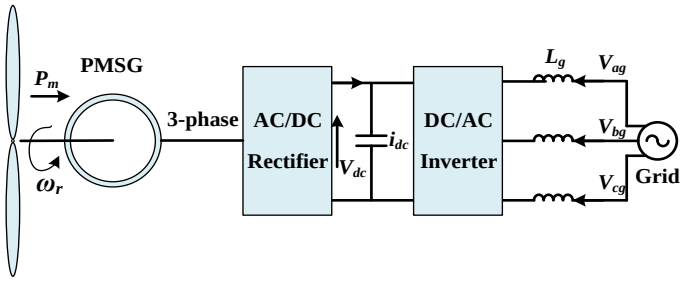


Figure 1. Configuration of a direct-drive PMSG system

The 2D analysis of outer rotor surface mounted PM generator is studied in this paper. The output power is 3 kW, rated

*Corresponding Author: Theint Zar Htet, Tsinghua University, +8618811515613
Email: theintzarhtet@gmail.com

mechanical speed is 250 rpm and number of poles are 34, respectively.

2. Modelling of PMSG by Maxwell 2D RMxprt

2.1. The geometry model of PMSG

In this study, outer rotor radial flux surface mounted direct-driven PMSG are employed. The radial flux outer-rotor PMSG, 17 pole pairs, are considered. The advantages of outer rotor PMSG compared to inner-rotor PMSG is that the magnets are easily attached to the rotor surface.Permanent magnet synchronous generators (PMSGs) have two parts; stator and rotor. Rotor has permanent magnets to create constant magnetic field in electrical machine[6]. Other time and space varying magnetic field is created with the stator windings. The geometry model of outer rotor radial flux surface mounted PMSG is shown in Figure 2.

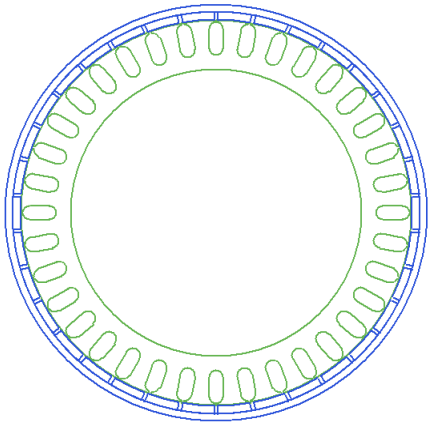


Figure 2. The geometry model of outer rotor PMSG

2.2. Maxwell 2D Model of PMSG

In this model, 3 kW small scale DDPMSG for wind turbine are designed. The main parameters of the DDPMSG are shown in Table 1.

Table 1: Major Parameters of DDPMSG

Parameters	Unit	Quantities
Rated output power	kW	3
Rated phase voltage	V	220
Control Type	-	AC
Rated power factor	-	0.9
Rated speed	rpm	250
Number of poles	-	34
Number of phases	-	3
Permanent magnet	-	NdFeB
Rotor Position	-	Outer Rotor

Parametrical analysis procedure are considered. Firstly, the stator steel type, rotor steel type and the magnet type should be determined from RMXprt library and the diameters of stator and rotor, the thickness of yoke and the airgap are roughly defined. And then, design calculations and modelling analysis are started. Table 2 represents the model design results of 3kW PMSG which are based on electrical machine theory.

Table 2: Design Results of DDPMSG

Parameters	Unit	Quantities
Phase		3
Pole		34
Stator Slot		36
Conductors per slot		24
Length of Rotor	mm	60
Outer Diameter of Rotor	mm	512
Inner Diameter of Stator	mm	358
Thickness of Magnet	mm	10
Rotor Pole Embrace	-	0.9

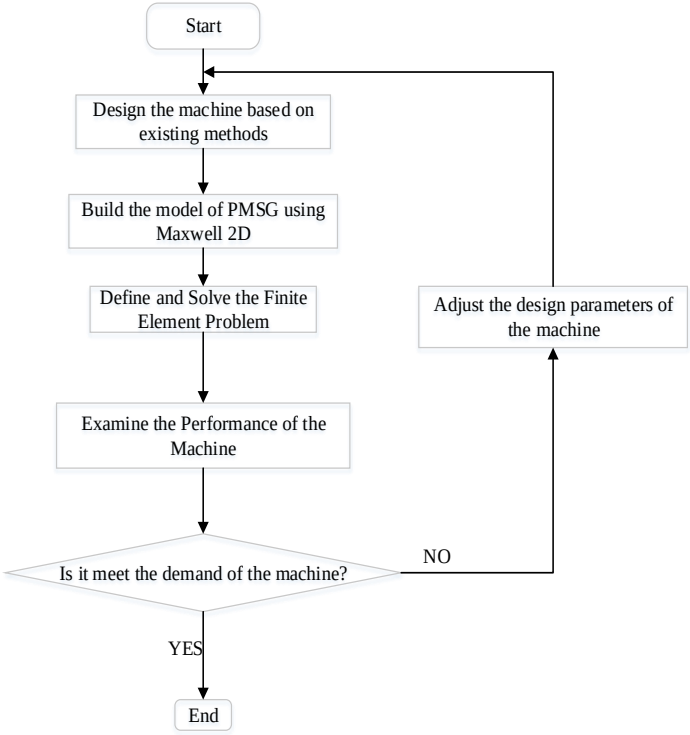


Figure 3. Design flow chart of the model for the PMSG

According to the proposed flow chart of the PMSG shown in Figure 3, a PMSG of 3kW is modelled. In this model, RMXprt (XG196/96) magnets are radially magnetized and are mounted on the surface of the rotor and the steel type of D21-50 is chosen for stator and rotor laminations. To reduce the risk of partial demagnetization, an accurate prediction of magnet loss at the design stage is an important role. It can not only give a better efficiency but also may prevent excessive temperature rise in magnets [6]. Considering the almost fixed magnetic field, the rotor core is made of solid steel and the stator core is made of electrical steel with 36 slots for three phase windings. The

prototype model of 3 kW DDPMSG in wind energy conversion system (WECS) can be seen in Figure 4.

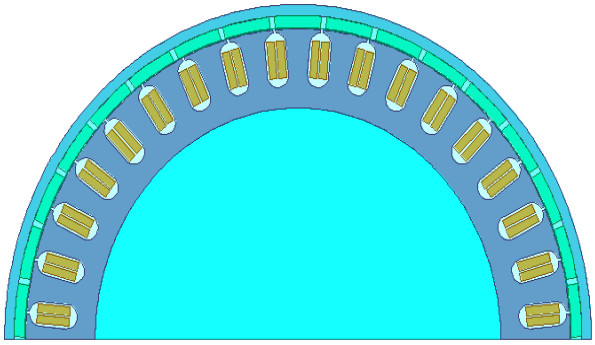


Figure 4. Maxwell 2D Model of PMSG Prototype

3. General Characteristics of PMSG

Nowadays, the DDPMSG is widely used in wind power application which has some advantages compare to other type of generators. The Maxwell 2D stator winding construction of this model is presented in Figure 5.

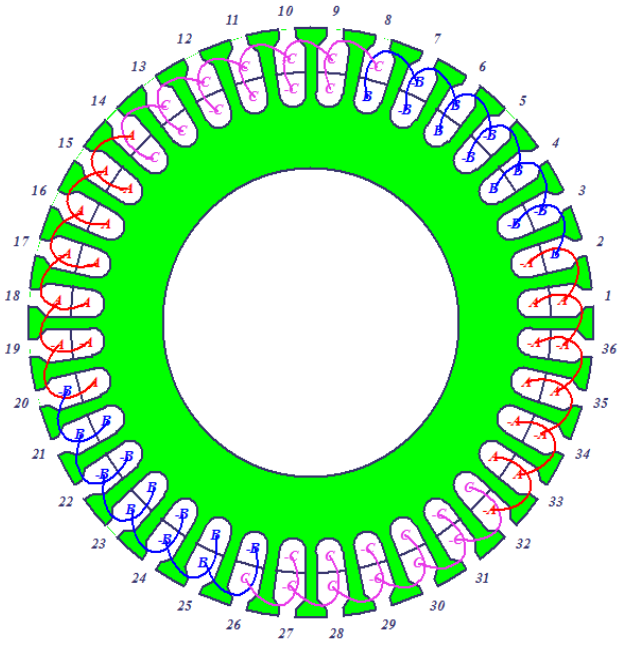


Figure 5. The stator winding connection of 3 kW PMSG

Moreover, the hysteresis (B-H) curves which express the relation between the flux density B and the magnetic field intensity H of the stator and rotor steel is an important characteristics of PMSG[7]. The stator and rotor steel type B-H curve of the FEM model is shown in Figure 6. Figure 7 represents the relationship between the flux density B and the magnetic field intensity H of magnet XG 196/96 from FEM analysis.

3.1. Magnetic Flux and Flux Density Distribution

With the application of high energy-product rare-earth PM to electric machines, many 2-D analytical models have been proposed for the direct driven PM machines and most of them focus on the field distribution and airgap flux density distribution. ANSYS Maxwell 2D RMxprt software is universal finite element analysis software of magnetic field calculation [8].

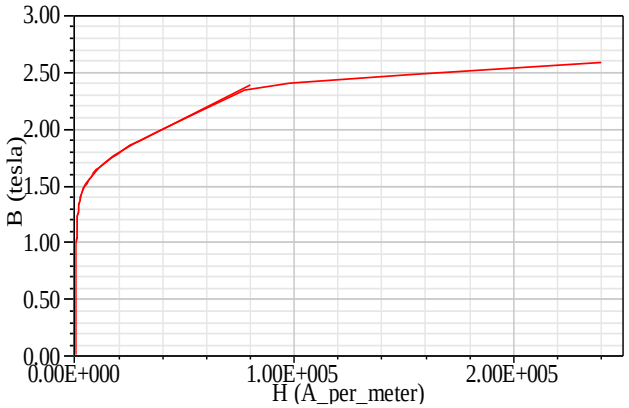


Figure 6. B-H Curve of Stator and Rotor Steel D21-50 RMxprt

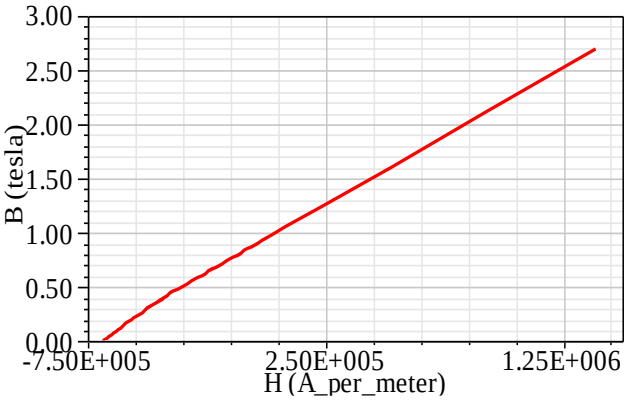


Figure 7. B-H Curve of Magnet (XG196/96) RMxprt

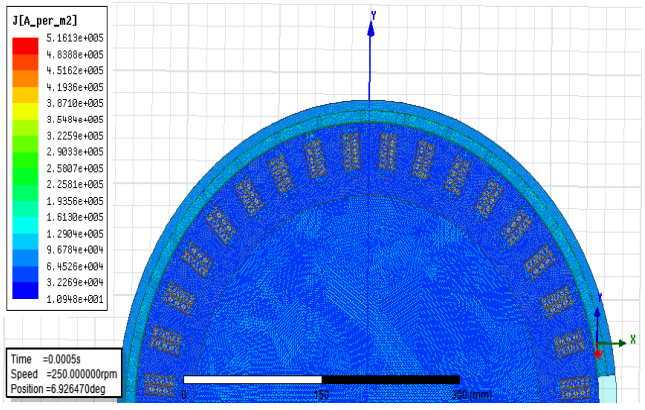


Figure 8. Meshing Plot of J Vector by Maxwell 2-D PMSG Model

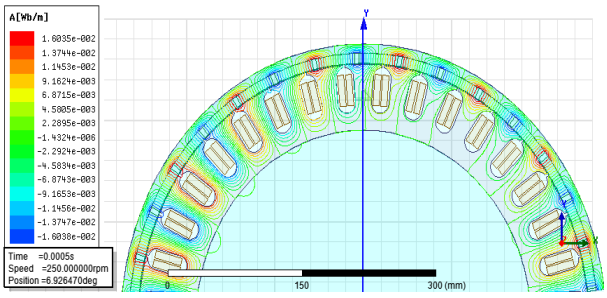


Figure 9. Flux Distribution in Maxwell Design

The model meshing plot of J vector can be seen in Figure *8. And* then, The Maxwell 2D analysis of the important parameters of PMSG such as the flux distribution, field distribution and magnetic flux density distribution are shown in Figure 9, Figure 10 and Figure 11, respectively.

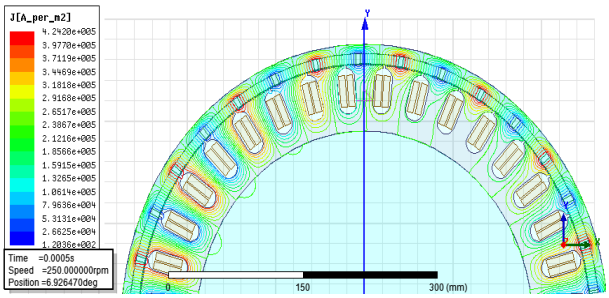


Figure 10. Field Distrubtion of J vector

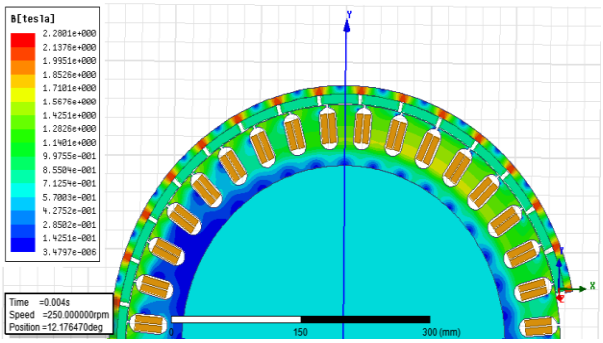


Figure 11. Magnetic flux density distribution of Maxwell 2-D

3.2. Mathematical Models of PMSG

Mathematical model of PMSG is usually given in rotating reference frame of machine. The quadrature (q) axis and the direct (d) axis PMSG equations are given below. Figure 12 and Figure 13 show equivalent circuit of PMSG in d-q frame [9].

$$v_{qs} = R_s i_{qs} + \frac{d\psi_{qs}}{dt} + \omega\psi_{ds} \quad (1)$$

$$v_{ds} = R_s i_{ds} + \frac{d\psi_{ds}}{dt} - \omega\psi_{qs} \quad (2)$$

$$\psi_{ds} = (L_{ld} + L_{md})i_{ds} + \psi_{mag} = L_{ds} i_{ds} + \psi_{mag} \quad (3)$$

$$\psi_{qs} = (L_{ld} + L_{mq})i_{qs} = L_{qs} i_{qs} \quad (4)$$

$$T_e = \frac{3}{2}\frac{P}{2}\psi_{mag} i_{qs} \quad (5)$$

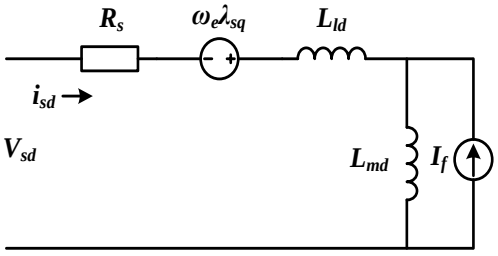


Figure 12. The d- axis equivalent circuit of PMSG

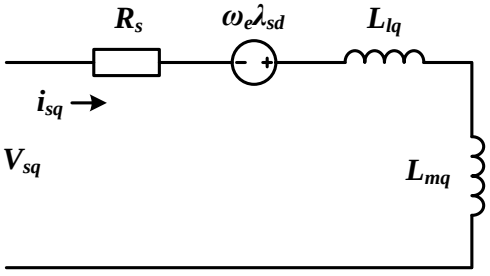


Figure 13. The q- axis equivalent circuit of PMSG

3.3. Finite Element Model Analysis

The two dimensional finite element model of the PMSG of 3kW with surface mounted PM is set up by using Maxwell 2D RMxprt. There are four main modelling processes such as inputting the data of the geometrical size of the PMSG, defining the stator, rotor, airgap, windings and PM properties, defining the boundary condition and generating finite element mesh.

After finishing these four steps, the proposed PMSG could be the modeled and solved though the FEA at different operation modes.

4. Performance Analysis of DDPMSG

The small scale wind turbine is directly connected to PMSG which receives wind turbine torque as a mechanical input and converts it into the electrical input; however, this electrical energy has a variable amplitude and frequency [10]. The FEA model of electromagnetic field is built by Ansoft Maxwell 2D and the simulation time is taken some hours [11]. To visualize magnetic fields and predict magnetic forces, Maxwell 2D program can be used.

The modelling analysis of 3 kW PM generator for wind energy conversion system is presented. The key characteristics of DDPMSG such as the cogging torque, air-gap flux density, induced coil voltage at rated speed and winding current under load are analyzed by Maxwell 2D RMxprt. Figure 14 shows the curve of cogging torque in two teeth. In PMSG, the cogging torque is generated by the sum of a series of cogging torque harmonics [12]. The cogging torque in PM machine can be expressed as:

$$T_c = \frac{1}{4\mu_0} L_{eff}(r_m^2 - r_s^2)\sum_{i=1}^{\infty} nNB_{nN}^2 \lambda_{nN}^2 \sin(nN\theta_r) \quad (6)$$

Where N is the least common multiple of the number of poles and stator slots of the PMSG. The cogging torque causes noise and vibration in PMSG at low speed and the direct drives applications. The frequency of cogging torque is the proportional to the least common multiple of the number of poles and the stator slots of the PMSG. The results obtained by using 2D Maxwell model provide sufficient accuracy with measured ones in a PMSG prototype.

In this FEM model, the proper simulation results of DDPMSG as shown in Figure 14 to Figure 19.The variation of cogging torque with respect to electrical degree can be seen in Figure 14 that generated the 3 kW FE model of Maxwell 2D.

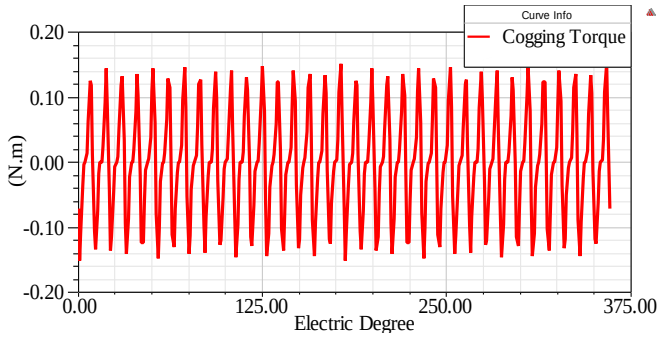


Figure 14. Cogging torque in two teeth of 3kW PMSG

The relationship between the winding voltage under load condition and electrical degree can be determined by the FEM analysis represented in Figure 15. The induced winding voltage of rated speed is shown in Figure 16.

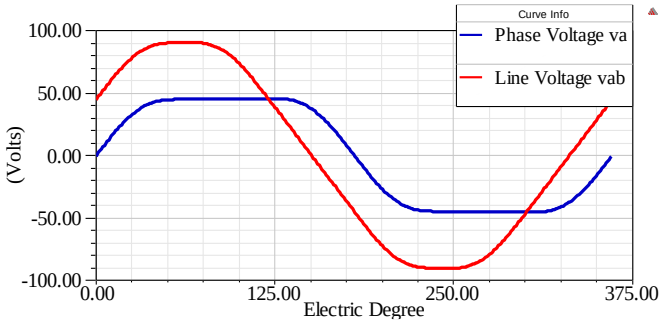


Figure15. Winding Voltage under Load

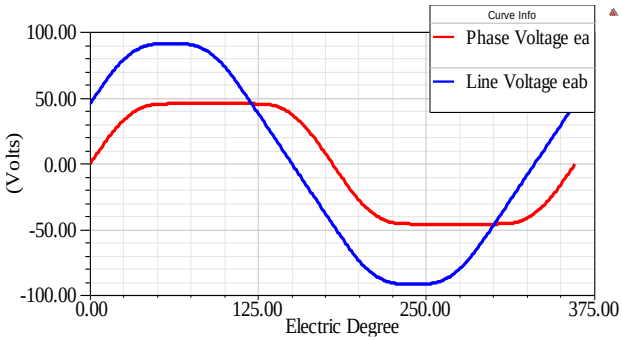


Figure 16. Induced Winding Voltage of Rated Speed

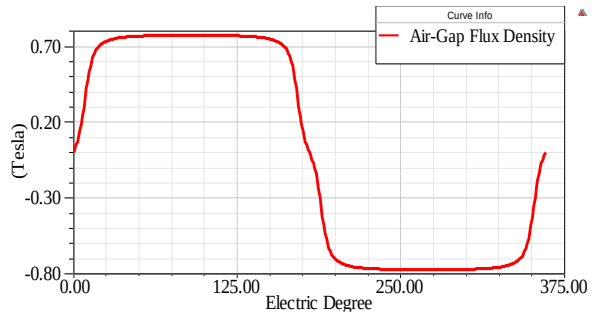


Figure 17. Air –Gap Flux Density

In this model, the maximum air-gap flux density distribution can be generated about 0.7 T by Maxwell 2D.The important parameter of air-gap flux density distribution of PMSG as shown in Figure 17. The performance curves of induced coil voltage at rated speed and the winding current under load are represented in Figure 18 and Figure 19,respectively.

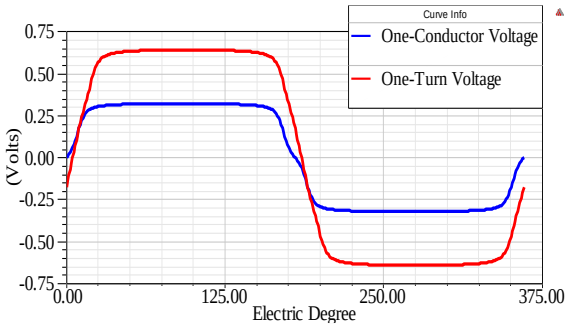


Figure 18. Induced Coil Voltage at Rated Speed

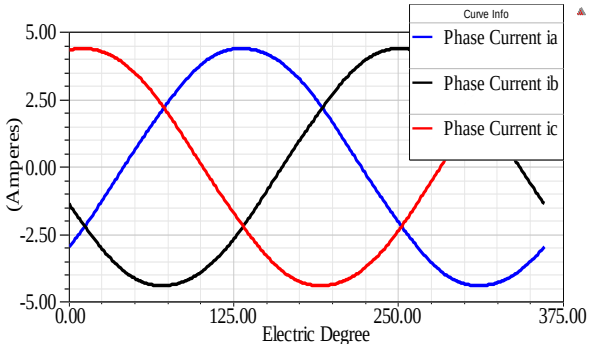


Figure 19. Wndinding Current under Load

Moreover, the transient analysis of 3 kW DDPMSG are discussed in this paper. The transient plots of main parameters in wind generator like the induced winding voltage of phase A,B and C,the comparison of solid losses and core loss,the moving torque of PMSG and the flux distributions in machine are generated by Ansoft Maxwell 2D RMxprt. All of the transient analysis curves are shown below.

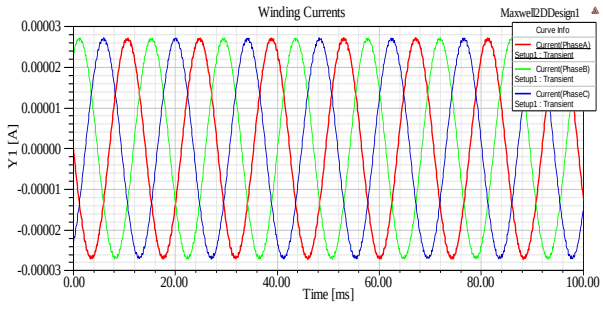


Figure 20. Induced Windind Voltages of Phase A, B, and C

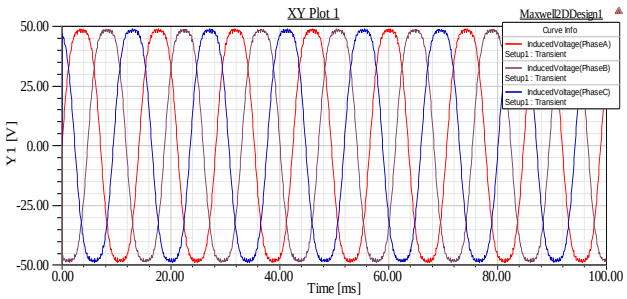


Figure 21 .Comparison of Core loss and Solid loss

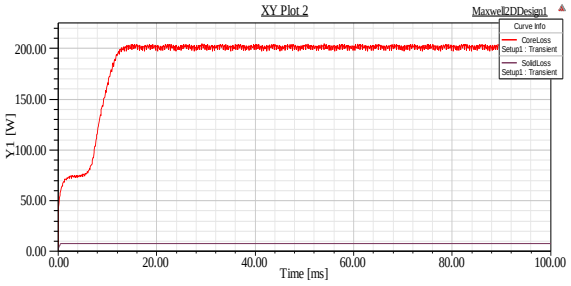


Figure 22. Transient Moving Torque of PMSG

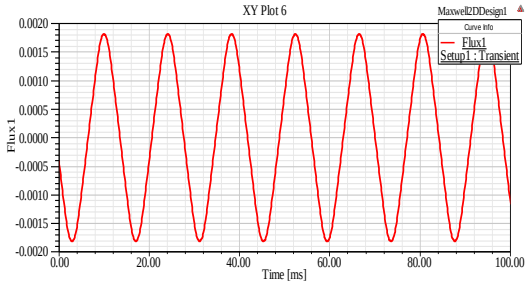


Figure 23. Flux analysis of Generator

5. Conclusions

The 2D modelling analysis of surface mounted PM generator in small scale wind turbine is studied in this paper. There are two main parts of the performance analysis in electrical machine such as the FEM performance curves and the transient curves. In this paper, the key machine parameters analysis are discussed by Ansoft Maxwell 2D RMxprt. To construct the machine model, Maxwell 2D RMxprt is to be a competent and valuable tool. Moreover, the analysis results obtained by using 2D finite element model provided sufficient accuracy with a PMSG prototype. According to the FEM modelling analysis and transient analysis, the design of 3 kW PMSG in this paper is suitable for small wind power applications.

Nomenclature

T_c	Cogging torque
L_{eff}	Effective length of the PMSG
r_m	Inner diameter of the magnet
r_s	Outer diameter of the stator
L_{mq}	Q-axis magnetizing inductance
L_{md}	D-axis magnetizing inductance
ω	Electrical angular speed
T_e	Electromagnetic Torque of PMSG
L_{ds}	D-axis stator inductance
L_{qs}	Q-axis stator inductance
R_s	Stator resistance
v_{qs}	Q-axis stator voltage
v_{ds}	D-axis stator voltage
ψ_{ds}	D-axis flux linkage
ψ_{qs}	Q-axis flux linkage
ψ_{mag}	Permanent magnet flux linkage
i_{ds}	D-axis stator current
i_{qs}	Q-axis stator current

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The lead author is grateful to her supervisor for supporting to develop this study by providing the suitable suggestions.

Reference

[1] Umashankar, D.P.K.a., *Wind Energy Systems and Applications*. 2013, Narosa.

[2] Yasa, Y. and E. Mese. *Design and analysis of generator and converters for outer rotor direct drive gearless small-scale wind turbines*. in *Renewable Energy Research and Application (ICRERA), 2014 International Conference on*. 2014.

[3] Alnasir, Z. and M. Kazerani, *An analytical literature review of stand-alone wind energy conversion systems from generator viewpoint.* Renewable and Sustainable Energy Reviews, 2013. **28**: p. 597-615.

[4] Yan, J., et al., *Control of a grid-connected direct-drive wind energy conversion system.* Renewable Energy, 2014. **66**(Supplement C): p. 371-380.

[5] Faiz, J., et al. *Design of a radial flux permanent magnet wind generator with low coercive force magnets*. in *2016 2nd International Conference on Intelligent Energy and Power Systems (IEPS)*. 2016.

[6] Nair, S.S., et al., *Computationally efficient 3D analytical magnet loss prediction in surface mounted permanent magnet machines.* IET Electric Power Applications, 2017. **11**(1): p. 9-18.

[7] Ducar, I.M. and C.P. Ion. *Design of a PMSG for micro hydro power plants*. in *Optimization of Electrical and Electronic Equipment (OPTIM), 2012 13th International Conference on*. 2012.

[8] Zhang, Z., et al. *Study on circle-annulus meshing methods using ANSYS software and its application in motor electromagnetic field analysis*. in *Electrical and Control Engineering (ICECE), 2011 International Conference on*. 2011.

[9] Ugalde-Loo, C.E., et al. *Analysis and efficient control design for generator-side converters of PMSG-based wind and tidal stream turbines*. in *Power Systems Computation Conference (PSCC), 2014*. 2014.

[10] Fazli, N. and J. Siahbalaee. *Direct torque control of a wind energy conversion system with permanent magnet synchronous generator and Matrix Converter*. in *2017 8th Power Electronics, Drive Systems & Technologies Conference (PEDSTC)*. 2017.

[11] Alam, F.R. and K. Abbaszadeh, *Magnetic Field Analysis in Eccentric Surface-Mounted Permanent-Magnet Motors Using an Improved Conformal Mapping Method.* IEEE Transactions on Energy Conversion, 2016. **31**(1): p. 333-344.

[12] Liu, T., S. Huang, and J. Gao. *Optimization of a direct-drive wind power generation system by using PMSG*. in *Power Engineering, Energy and Electrical Drives (POWERENG), 2011 International Conference on*. 2011.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 100-106 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

A Switched-Capacitor Low-Pass Filter with Dynamic Switching Bias OP Amplifiers

Hiroo Wakaumi*

Tokyo Metropolitan College of Industrial Technology, 140-0011, Japan

ARTICLE INFO

Article history:
Received : 28 September, 2017
Accepted : 8 November, 2017
Online : 18 November, 2017

Keywords :
Switched Capacitor Circuit
Filter
CMOS
Operational Amplifier
Dynamic Switching

ABSTRACT

A switched capacitor low-pass filter employing folded-cascode CMOS OP Amps with a dynamic switching bias circuit capable of processing video signals, which enables low power consumption, and operation in wide bandwidths and low power supply voltages, is proposed. In this filter, charge transfer operations through two-phase clock pulses during the on-state period of the OP Amps and a non-charge transfer operation during their remaining off-state period are separated. Through simulations, it was shown that the low-pass filter with an OP Amp switching duty ratio of 50 % is able to operate at a 14.3 MHz high-speed dynamic switching rate, allowing processing video signals, and a dissipated power of 68 % of that observed in the static operation of the OP Amps and a full charge transfer operation without separation of a cycle period. The gain below -31 dB in the frequency response, which is suitable, was obtained at over 6 MHz within a stop-band. Especially high attenuation in 5 MHz was achieved under the optimized condition of load capacitances (4 pF) of OP Amps.

1. Introduction

The switched capacitor (SC) techniques are appropriate for realizing various filters that can be integrated in monolithic ICs (Integrated Circuits) from using the CMOS (Complementary Metal-Oxide-Semiconductor) technology. In conventional active RC filters, monolithic ICs cannot be realized from using resistors. On the contrary, the CMOS SC techniques suitable for realizing analog signal processing ICs, have promising use in video signal bandwidth circuits in particular, because these can replace resistors to switched capacitor pairs with small capacitances. It has been demonstrated that SC techniques using CMOS operational amplifiers (OP Amps) are useful for implementing analog functions such as filtering [1-5]. Although CMOS OP Amps are suitable for such filter ICs, the use of several OP Amps results in large power consumption. Especially, the power consumption of OP amplifiers in high speed operation becomes large because they have wideband properties. Therefore, the use of them is currently limited to the use in low-frequency passband of at most a few hundred kHz (that is, applications of low speed signal processing, such as analog voice signals).

Until now, several approaches have been considered to decrease the power consumptions of OP Amps, including the development of ICs that work at low power supply voltages [6]. A clocked current bias scheme for folded-cascade OP Amps suitable for achieving a wide dynamic range has typically been proposed to decrease the power consumption of the OP Amp itself [7, 8]. Because the circuit requires complicated four-phase bias-current control pulses and biasing circuits, it results in a large layout area and is not suitable for the high-speed operation. A control method using power supply switching has been proposed for audio signal processing as another approach in decreasing the power consumption of OP Amps [9]. Because large capacitors for the power supply terminals are intrinsically loaded, the switching speed is limited to a low speed of 1 MHz at most. Therefore, this type of control circuit is not suitable for application to video signal processing ICs, which are required to operate at over 10 MHz switching frequency.

Recently, the author proposed a folded-cascade CMOS OP Amp with a dynamic switching bias circuit (DSBFC OP Amp), of simple configuration, to provide low power consumption while maintaining high speed switching operation suitable for processing video signals [10]. However, low-power signal processing circuits such as wideband filters with such DSBFC OP

*Corresponding Author: Hiroo Wakaumi, Tokyo Metropolitan College of Industrial Technology, Tokyo, 140-0011 Japan
E-mail: wakaumi@metro-cit.ac.jp ; waka781420j@ab.auone-net.jp

Amp, suitable for processing video signals, have not been developed yet.

In this paper, a configuration of SC Butterworth low-pass filter (LPF) with DSBFC OP Amps [11] as an example of application of the DSBFC OP Amp is proposed, which enables low power consumption and is suitable for achieving wide bandwidths and operation in low power supply voltages, and its availability is also evaluated for the performance of frequency response and power dissipation. Further, the effect of OP Amp load capacitances on the frequency response is evaluated. This paper is an extension of work originally presented in LASCAS 2017 [11].

2. SC Filter Theory

The discrete-time transfer function of second-order SC infinite impulse response (IIR) LPF is shown as follows using the z-transform.

$$H(z) = \frac{K\ (1+z^{-1})^2}{1+\sum_{k=1}^{2} b_k z^{-k}} \quad (1)$$

Here, K, b^k, and z^{-1} represent the gain constant, the filter coefficient in a recursive loop, and the one-step delaying operation, respectively. All of operation circuits are composed of active sampled data processing circuits with a sampling circuit, switching circuits and capacitors.

3. SC Filter Circuit Design

A second-order IIR LPF with the Butterworth frequency characteristic was designed because it is easy to design due to its flat gain characteristic in the passband. The Butterworth LPF is also superior to a Chebyshev filter for processing video signals owing to its ripple less characteristic within the passband. The filter order of second was selected to achieve a gain of -30 dB at a stop-band over 6 MHz. The other design condition was set as follows. That is, a sampling frequency fs=14.3 MHz, which is equal to four times as much as the NTSC color sub-carrier frequency 3.58 MHz, and a cutoff frequency fc=2 MHz, respectively, were chosen, that enable the LPF to process video signals. Under this condition, the inverted discrete-time transfer function is given by

$$H(z) = -\frac{0.1174(1+z^{-1})^2}{1-0.8252z^{-1}+0.2946z^{-2}} \quad (2)$$

The circuit configuration realizing this transfer function is shown in Figure 1. In order to enable easily to determine the capacitance value of each capacitor, the coefficient of A is set to be equal to that of B. Capacitors can be basically divided into two groups. In one group (C, D, E, and G), charges are supplied to OP Amp 1. In another group (A, B, and I), charges are supplied to OP Amp 2. Even if a capacitance of each capacitor group is multiplied by constant times, the transfer function does not change. Therefore, capacitances of integral capacitors B and D are here selected as a reference capacitance in each group and each coefficient of B and D is normalized to 1. At this time, 1 as every normalized coefficient of A, B, and D is obtained because A and B are the same coefficient. In Figure 1, when the coefficients of A, B, and D are normalized to 1, other coefficients are determined as follows.

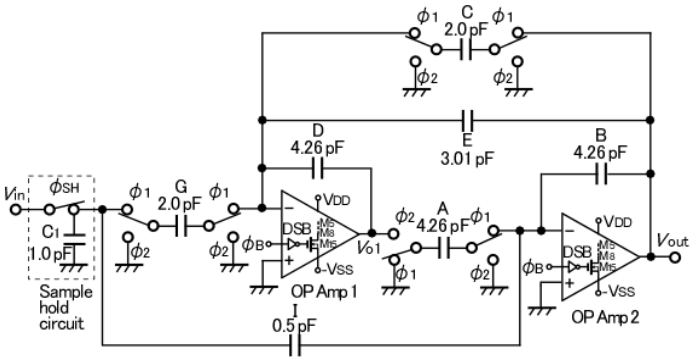


Figure 1: Configuration of the 2nd-order SC LPF with DSBFC OP Amps

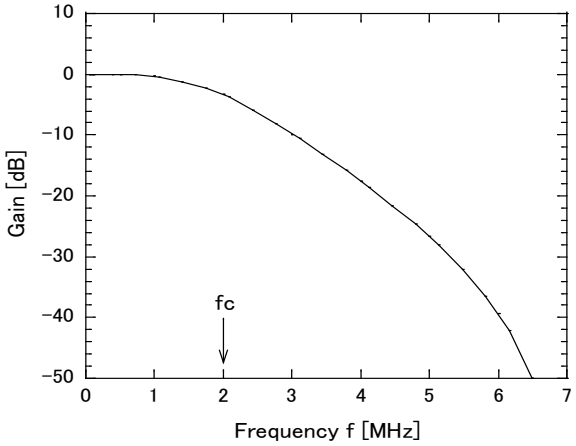


Figure 2: Theoretical frequency response of the 2nd-order SC Butterworth LPF

I=K=0.1174

G=2K+2I=0.4694

C=1+b_1+b_2=0.4694

E=1-b_2=0.7054

Here, b_1=-0.8252 and b_2=0.2946. When the smallest coefficient of I=0.1174 is replaced as a reference capacitance of 0.5 pF, each capacitance in the SC IIR LPF IC is set in proportion to the above coefficients as shown in Figure 1. Because its input signal is desirable to be maintained by a sample/hold circuit for stabilizing, this sampling circuit is also applied in the SC LPF. At this time, the transfer function is multiplied by the following zero-order hold function due to a sample-hold effect.

$$H_s(j\omega)=\frac{\sin(\omega T_s/2)}{(\omega T_s/2)} \quad (3)$$

Here, Ts represents a cycle period of sampling and switching pulses. Therefore, when the function of (2) is replaced using $z=e^{j\omega T_s}$, the magnitude of the transfer function of the second-order SC LPF considering the sample-hold effect is given by (4).

$$|H(j\omega)|=\frac{\sin(\omega T_s/2)}{(\omega T_s/2)} \cdot \frac{0.2347(1+\cos\omega T_s)}{\sqrt{1.7678-2.1368\cos(\omega T_s)+1.6513\cos(2\omega T_s)}} \quad (4)$$

The theoretical frequency response including the sample-hold effect is shown in Figure 2. The SC LPF configuration was designed referencing a SC biquad circuit with integrators in the

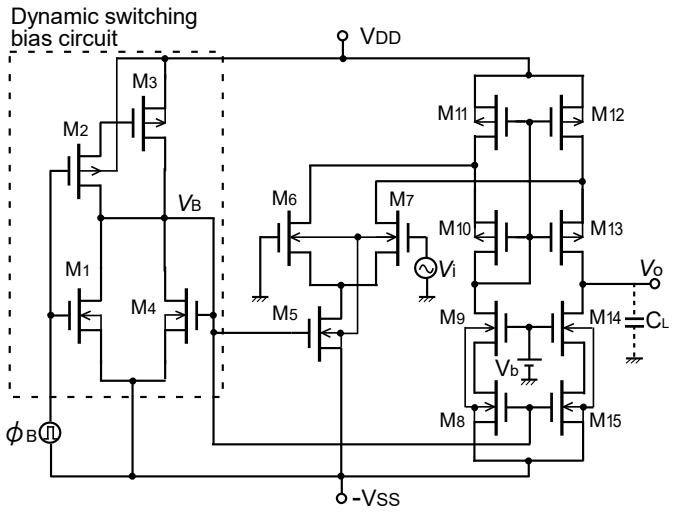


Figure 3: Configuration of the DSBFC OP Amp

reference [12]. It consists of a sample-hold circuit with a holding capacitor C_1 and a CMOS sampling switch controlled by ϕ_{SH}, CMOS switches ϕ_1, ϕ_2, capacitors A, B, C, D, E, G, and I, and two FC CMOS OP Amps with the DSB circuit (DSBFC OP Amps). A configuration of the DSBFC OP Amp enabling low power consumption, which is different from conventional ordinary OP Amps [10], is shown in Figure 3. With respect to the DSBFC OP Amps, the same CMOS channel width / length as that in the DSBFC OP Amp shown in the reference [10] was employed. The sampling switch was designed to a channel width / channel length W/L=35/2.5 (μm/μm) for each of p-MOSFET and n-MOSFET. The holding capacitor C_1 has a small capacitance of 1 pF. CMOS switches with a W/L=25/2.5 (μm/μm) are turned on and off by non-overlapping two-phase clock pulses ϕ_1 and ϕ_2, swinging from -2.5 V to 2.5 V. These sampling and CMOS switches are designed to have a balanced structure with each equal length and width of p-channel and n-channel MOSFETs (component of these CMOS switches) to delete a feed-through phenomenon, which is caused by gate clock pulses due to a capacitive coupling between gate and CMOS-switch output terminals. Major CMOS process parameters are given as a gate insulating film thickness t_{ox}=50 nm, an n-MOSFET threshold voltage V_{Tn}=0.6V, and a p-MOSFET threshold voltage V_{TP}=-0.6 V.

The operation principle of this LPF is simply described in the following. The output signal V_{o1} of OP Amp 1 is obtained as an additional output of an integrated signal of V_{in} using a negative integrator (D, G SC circuit, and OP Amp 1), an integrated signal of V_{out} using a negative integrator (D, C SC circuit, and OP Amp 1), and a signal multiplied V_{out} by E/D. The output signal V_{out} is an additional output of an integrated signal of V_{o1} using a positive integrator (A SC circuit, B, and OP Amp 2), and a signal multiplied V_{in} by I/B. V_{out} is basically fed back to an input of OP Amp 1 like this. V_{in} is also integrated twice and added after being decreased by an appropriate capacitance ratio. Due to these integration using positive / negative integrators, addition and feedback operations, the function of LPF is achieved.

The operation waveforms of the SC LPF are shown in Figure 4. In this SC LPF, charge transfer operations through the clock

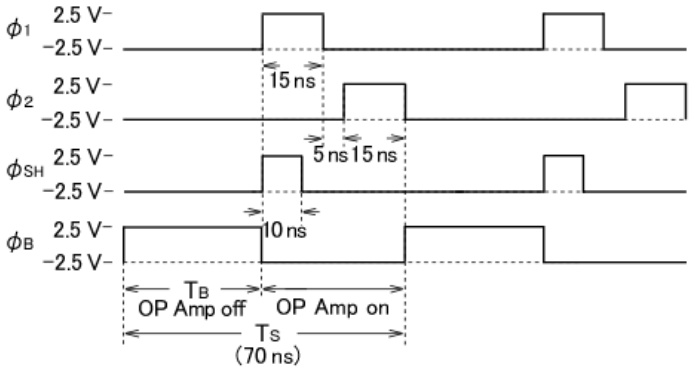


Figure 4: Operation waveforms of the 2nd-order SC LPF

pulses ϕ_1 and ϕ_2 are achieved during the on-state period of the DSBFC OP Amps. The off-state period T_B (the remaining period of the one cycle period Ts) is separately provided to realize low power dissipation for the SC LPF. An input signal is sampled during the sampling phase of ϕ_{SH} (10 ns) and the first period of clock phase ϕ_1, while its corresponding charge is stored on the holding capacitor C_1, and is transferred to an output terminal V_{out}, charging all capacitors. The voltage at the off transition of ϕ_{SH} is kept on C_1 during the remaining period of clock phase ϕ_1. During subsequent clock phase ϕ_2, each charge of two capacitors C and G is discharged and each charge of remaining capacitors is redistributed. These charge transfer operations are achieved during the on-state period of the OP Amps. During this period, the OP Amps turn on by setting a bias voltage of V_B at an appropriate level enabling M3 and M4 to operate in the saturation region, and operate normally as operational amplifiers. ϕ_B is set to low just before ϕ_1 changes to high.

Subsequently, ϕ_B becomes 2.5 V at the off-state transition of the OP Amps, at the same time ϕ_2 is switched to off. During this off period T_B, the OP Amps turn off by setting V_B at nearly -2.5 V enabling M1 to operate in a low impedance and M3 in a high impedance. Therefore, during this off period, the OP Amps do not dissipate at all. When T_B is relatively long as compared to the one cycle period Ts, the power dissipation is expected to become lower than that observed in ordinary static operation for the SC LPF using conventional OP Amps. If half GB (Gain Bandwidth product) OP Amps for the SC LPF are used in the static operation (without DSB operation), rise and fall times to stable states of filter output signals will increase to much larger than twofold, because slowly changing transition occurs at the end of transition. Therefore, an expected proper filter performance will not be able to be obtained when such OP Amps are used.

4. Simulation Results

The performance of the SC LPF was investigated by simulation using the SPICE (Simulation Program with Integrated Circuit Emphasis) program package. Operation waveforms for an input signal of 1 MHz with an amplitude of 0.3 V and an output load capacitance of 1 pF are shown in Figure 5. For the passband frequency signal, the same level output signal as the input one was obtained. The frequency response of the SC LPF is shown in Figure 6 (a) in the dynamic switching operation of the DSBFC OP Amps and (b) in the on-state (static operation) of the DSBFC

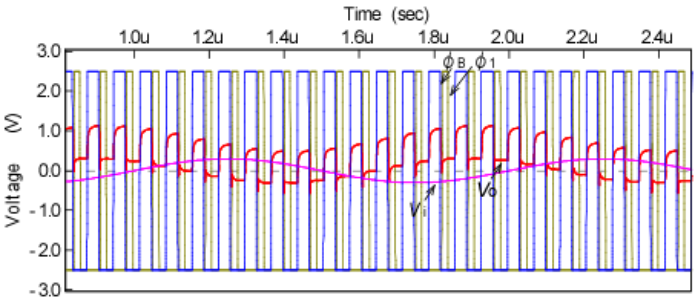
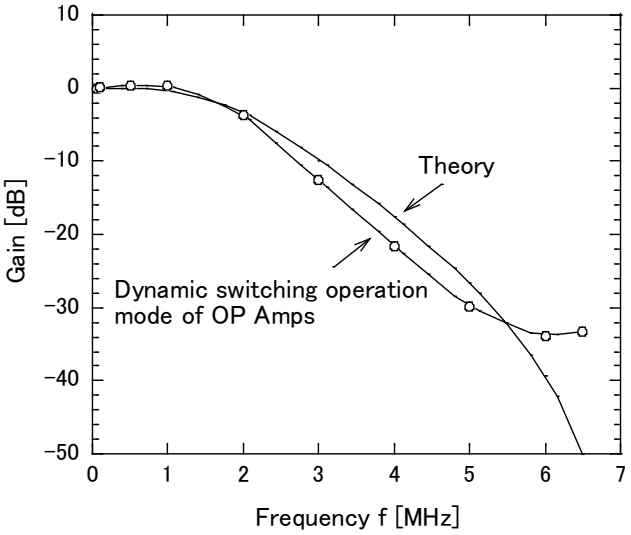
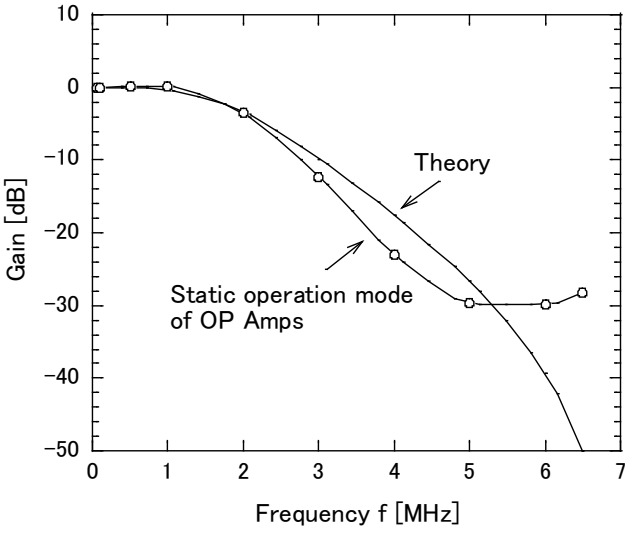


Figure 5: Simulation waveforms for the 2nd-order SC Butterworth LPF. Input signal=0.3 $V_{0\text{-}P}$, Input signal frequency $f$in=1 MHz



(a)



(b)

Figure 6: Frequency response of the 2nd-order SC Butterworth LPF. (a) Dynamic switching operation mode of the DSBFC OP Amps, T_B=35 ns, (b) Static operation mode of the DSBFC OP Amps, ϕ_B=-2.5 V

OP Amps at ϕ_B= -2.5 V. The uncertainty of gain characteristics in simulation is nearly 1 dB at the frequency of 6.5 MHz. The frequency responses in case of both modes are almost the same for the in-phase state, in which the maximum output signal is

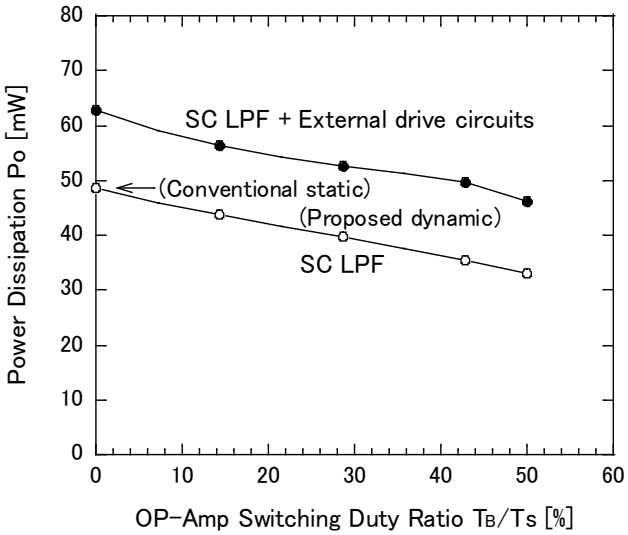


Figure 7: Power dissipation vs. OP Amp switching duty ratio in the 2nd-order SC LPF. $f$in=1 MHz. A constant pulse width of 15 ns for ϕ_1 and ϕ_2 was used

obtained depending on a phase between the input signal and the sampling pulse. The response was near the theoretical one from 100 kHz up to near 5 MHz, in the high frequency range over 6 MHz, it deteriorated due to a sampling phase effect. The gain below -33 dB was obtained at over 6 MHz within a stop-band. This was superior to that in the static operation. Therefore, it is thought that there is almost no gain deterioration caused by employing the DSB operation of the OP Amps. Though the stop-band gain in the second order LPF is not low enough as shown above, it is expected that the roll-off in the frequency response will become steeper by increasing the filter order, and so the stop-band gain will greatly decrease. Therefore, the achievement of a wide stop-band with a high attenuation will become possible.

Power dissipation vs. OP-Amp switching duty ratio with ϕ_1= ϕ_2=15 ns is shown in Figure 7. The power dissipation of the SC LPF itself excluding that of external drive circuits for ϕ_{SH}, ϕ_1, ϕ_2, and ϕ_B decreased in proportion to the off period T_B of the Op Amps as expected. In the operation mode of T_B=35 ns (=50 % switching duty ratio) and ϕ_1= ϕ_2=15 ns, the power dissipation of the SC LPF (32.9 mW) was decreased to 68 % as compared to that in the static operation of the OP Amps (48.5 mW). In the full charge transfer operation without separation of the one cycle period Ts (ϕ_1= ϕ_2=30 ns), the power dissipation of the LPF was 48.5 mW, the same as that in the static operation of the OP Amps with the above separated charge transfer mode. Most of the power dissipation of the SC LPF corresponds to total power consumption dissipated in the OP Amps themselves. The power consumption of the external drive circuits was nearly 13~14 mW. Thus, even when two DSBFC OP Amps are applied to the SC LPF, the dynamic operation of the DSBFC OP Amps enabling low power dissipation as compared to their static operation is also useful for reducing the power dissipation of SC LPF. When the SC LPF is operated at a lower dynamic switching rate, because it enables T_B/Ts to become larger than 50 %, the power dissipation of the SC LPF is expected to decrease in proportion to T_B still more. This means that the SC LPF with high-speed DSB operation OP Amps is advantageous compared to the SC LPF using static operation OP Amps with a lower GB.

Table 1: Performance comparison of SC LPFs with CMOS technology

Performance parameters	Fifth-Order SC LPF with CMOS OP Amps [13]	2nd-Order SC LPF with DSB CMOS OP Amps
Process technology	0.35 µm CMOS	2.5 µm CMOS
Power supply voltage	5 V	5 V
Sampling frequency fs	103.68 MHz (Double sampling)	14.3 MHz
Cutoff frequency fc	20 MHz	2 MHz
Power consumption	125 mW	32.9 mW (Dynamic : T_B/T_S=50 %)

When the filter order is increased, a SC LPF with DSB OP Amps uses OP Amps of the number of filter order. Therefore, the power dissipation of this SC LPF is expected to increase in proportion to its filter order. The power dissipation for the reported fifth-order SC LPF employing conventional CMOS OP Amps with a 5-V power supply and a 0.35-µm CMOS technology was 125 mW as shown in the performance comparison of Table 1 [13]. If a fifth-order SC LPF using the DSB OP Amps with the sampling frequency fs=14.3 MHz is achieved, its power dissipation will be estimated to be 82.3 mW. Considering that the power consumption of OP Amps parts in each LPF is dominant, a comparison between these values is possible. Obviously, the estimated power dissipation of a revised version of the proposed SC LPF with DSB OP Amps is much less than that of the above conventional fifth-order SC LPF.

5. Effect of Load Capacitances

Because the DSBFC OP Amp switches dynamically, its output becomes a quasi-floating state during the off-state period. In the off-state period of the OP Amp, though MOSFETs M5, M8, and M15 turn to the off-state completely, MOSFETs M11 and M12 become the on-state strongly because over the threshold voltage between each gate and source is applied. At this time MOSFETs M6, M7, M9, M10, M13, and M14 become the on-state weakly. The output terminal *Vo* of the OP Amp is set to a voltage depending on the load capacitance through the capacitive coupling between the drain and the gate of the MOSFET M13. Therefore, when a large output swing in *Vo* occurs at the off-state transition, there is a fear that the output voltage during the subsequent on-state period of the OP Amp suffers the influence of this transition. So, a dynamic offset voltage Voff (the difference of the on-state and the off-state output voltages of the OP Amp) at the off-state transition of the OP Amp vs. load capacitance C_L was tested and is shown in Figure 8. Obviously, the dynamic offset voltage depends on the load capacitance and decreases as its load capacitance becomes large because C_L compared to the drain-gate capacitance in M13 becomes large. The change of Voff against C_L of the SC LPF resembles to that of the OP Amp (Figure 9). This means that Voff of the SC LPF is mainly determined by the OP Amp's dynamic off-state transition. Gain vs. OP Amp load capacitance (of two OP Amps) for the SC LPF is shown in Figure 10. In this case, the phase between the input signal V_{in} and the sampling signal ϕ_{SH} was fixed to a constant value in each input signal to avoid the sampling phase effect. The gain became the minimum at a load capacitance of nearly 4~5 pF for the input signal frequency of 5 MHz. The reason is explained in the following. In small load capacitances, V_{off} is not only large, but also its variation is not negligible depending on the on-state output voltage. This phenomenon causes the on-state output voltage's difficulty to

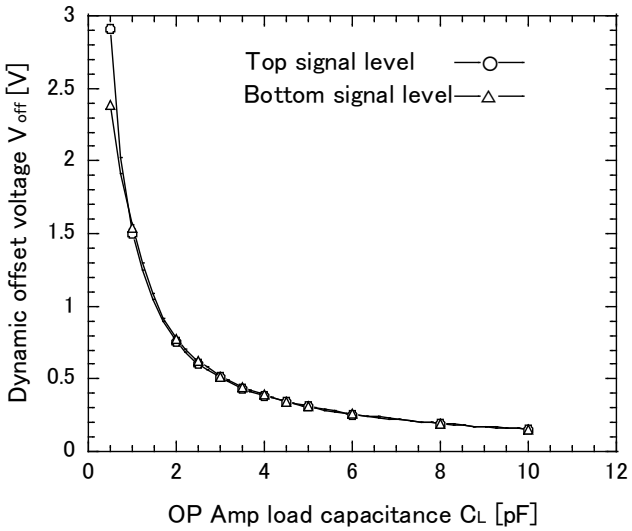


Figure 8: Dynamic offset voltage vs. load capacitance in the DSBFC OP Amp. Input signal=0.00085 $V_{0\text{-}P}$, Input signal frequency *f*in=0.5 MHz, Cutoff frequency *f*c at C_L of 1 pF =4.13 MHz

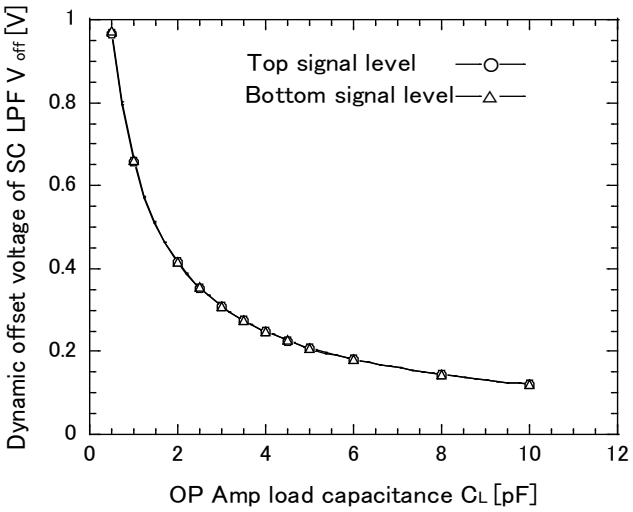


Figure 9: Dynamic offset voltage vs. OP Amp load capacitance in the SC LPF. Input signal *V*in=0.3 $V_{0\text{-}P}$, Input signal frequency *f*in=5 MHz

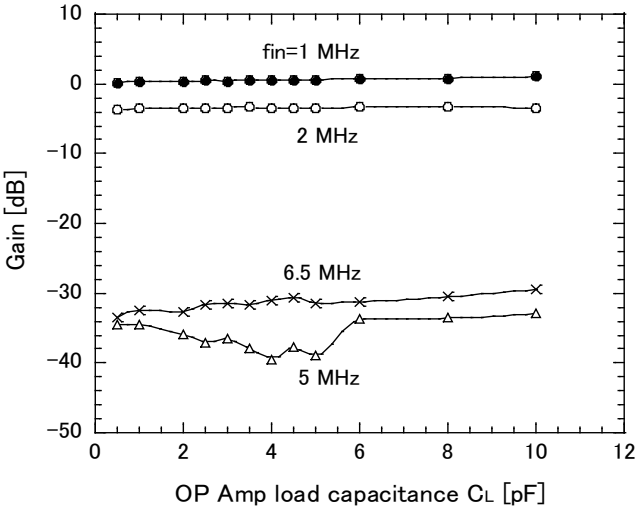


Figure 10: Gain vs. OP Amp load capacitance in the SC LPF

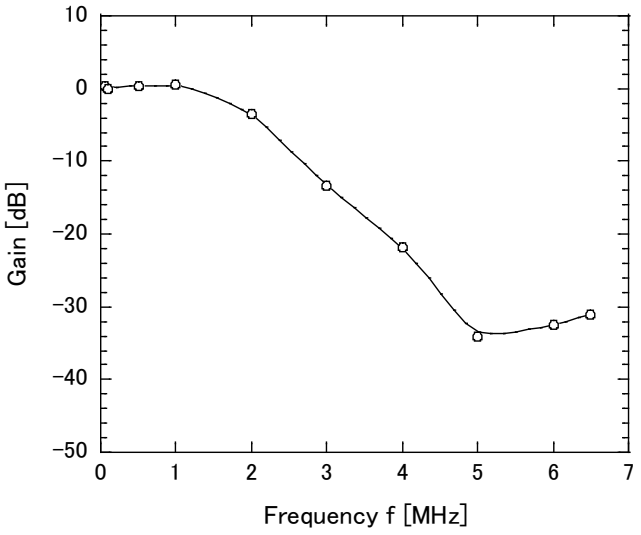


Figure 11: Frequency response of the SC LPF. OP Amp load capacitance C_L=4 pF

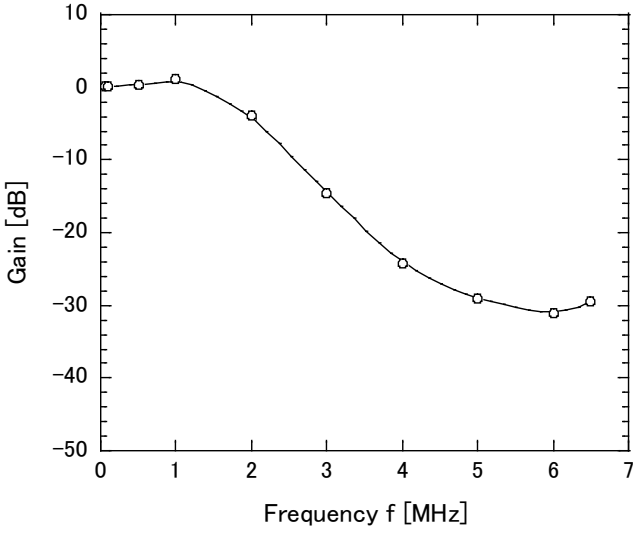


Figure 12: Frequency response of the SC LPF. OP Amp load capacitance C_L=10 pF

reach its steady state completely within the short time period of 15 ns (in ϕ_1 or ϕ_2), while permitting its output voltage to change slightly (bringing about the SC LPF gain's deterioration). In large load capacitances, it becomes also hard to reach its steady state within 15 ns (causing the gain's deterioration) due to OP Amp's bandwidth deterioration although V_{off} and its variation are very insignificant. However, at an optimum capacitance of nearly 4 pF, its gain deterioration becomes slight due to small offset transition with slight offset voltage variation and fast transition to the steady state. For other input signals of low frequencies (1, 2 MHz), the variation effect of V_{off} is negligible due to its large output signal (the change of gains is hardly seen) although its variation depending on the output signal voltage is relatively large in small load capacitances. Of course, since V_{off} and its variation for these low frequencies are negligibly small in large load capacitances, the change of gains with C_L is dominated by only OP Amp's bandwidth deterioration. For the 6.5 MHz input signal, though the gain is basically determined due to the sampling phase effect, it is slightly changing with C_L due to OP Amp's bandwidth deterioration.

Under the optimized load capacitance of 4 pF, the frequency characteristic of gain for the SC LPF was tested. As shown in Figure 11, its high frequency gain of near 5 MHz improved drastically (over 4.3 dB compared to the gain in the load capacitance C_L of 1 pF). Though its gain increased slightly (up to -31 dB) in 6.5 MHz, the amount of its gain increase is slight. On the contrary, when C_L is increased to 10 pF larger than the optimized value, the frequency response of the SC LPF deteriorated in high frequencies over 5 MHz (Figure 12). Thus we can see that the SC LPF gain characteristics can be improved by the optimization of load capacitances of DSB OP Amps. Typical characteristics are listed in Table 2.

Table 2: Typical characteristics of SC LPF with DSB OP Amps (V_{DD}=V_{SS}=2.5 V)

Performance parameters	Simulation results
Sampling and switching frequency fs	14.3 MHz
Input signal amplitude	0.3 $V_{0\text{-}P}$
Cutoff frequency fc	2 MHz
Gain at over 6 MHz within a stopband	<-31 dB (DSB mode) <-27 dB (Static mode)
Power consumption	48.5 mW (Conventional static) 32.9 mW (Proposed dynamic: T_B/Ts=50 %)

Conclusions

A switched capacitor low-pass filter employing folded-cascode CMOS OP Amps with a dynamic switching bias circuit capable of processing video signals, which enables low power consumption, operation in wide bandwidths and low power supply voltages, was proposed and its performance was evaluated. In this SC LPF, charge transfer operations through two-phase clock pulses during the on-state period of the OP Amps and non-charge transfer operation during the remaining off-state period of the OP Amps were separated. Through SPICE simulations, it was shown that the SC LPF is able to operate at a 14.3 MHz high-speed dynamic switching rate, allowing processing video signals, and a dissipated power of 68% of that observed in the static operation of the OP Amps and the full charge transfer mode without separation of the one cycle period. The power consumption in the SC LPF body except for the external drive circuits was that of OP Amps. When rearranging these results, it became clear that a lower-dissipated-power SC LPF employing DSB folded-cascode CMOS OP Amps compared to conventional SC LPFs with static operation OP Amps can be realized. The gain below -31 dB in the frequency response, which is suitable, was also obtained at over 6 MHz within a stop-band. Especially high attenuation in 5 MHz was achieved under the optimized condition of load capacitances (4 pF) of OP Amps.

Thus, the dynamic charge transfer operation during the on-state period of the OP Amps and non-charge transfer operation during their off-state period is useful for high speed operation, and reducing the power dissipation of the SC LPF. This circuit should be useful for the realization of low-power wide-band signal processing ICs including over one of multi-order low-pass, high-pass and band-pass filters. The DSB circuit achieving such operation can be applied to not only folded-cascode but telescopic, two-stage, and rail-to-rail OP Amps.

References

[1] R. Gregorian and W. E. Nicholson, "CMOS Switched-Capacitor Filters for a PCM Voice CODEC," IEEE J. Solid-State Circuits, **SC-14**(6), 970-980, Dec. 1979.

[2] R. Dessoulavy, A. Knob, F. Krummenacher, and E. A. Vittoz, "A Synchronous Switched Capacitor Filter," IEEE J. Solid-State Circuits, **SC-15**(3), 301-305, June 1980.

[3] A. Iwata, H. Kikuchi, K. Uchimura, A. Morino, and M. Nakajima, "A Single-Chip Codec with Switched-Capacitor Filters," IEEE J. Solid-State Circuits, **SC-16**(4), 315-321, Aug. 1981.

[4] J.-T. Wu, Y.-H. Chang, and K.-L. Chang, "1.2 V CMOS Switched-Capacitor Circuits," in 1996 IEEE International Solid-State Circuits Conference Digest of Technical Papers-42nd ISSCC, 388-389, 479.

[5] S. Chatterjee, Y. Tsividis, and P. Kinget, "Ultra-Low Voltage Analog Integrated Circuits," IEICE Trans. Electronics, **E89-C**(6), 673-680, June 2006.

[6] Z. Kun, W. Di, and L. Zhangfa, "A High-Performance Folded Cascode Amplifier," in 2011 International Conference on Computer and Automation Engineering-ICCAE 2011, IPCSIT, **44**, 41-44, 2012.

[7] D. B. Kasha, W. L. Lee, and A. Thomsen, "A 16-mW, 120-dB Linear Switched-Capacitor Delta-Sigma Modulator with Dynamic Biasing," IEEE J. Solid-State Circuits, **34**(7), 921-926, July 1999.

[8] H. L. Chen, Y. S. Lee, and J. S. Chiang, "Low Power Sigma Delta Modulator with Dynamic Biasing for Audio Applications," in 50th Midwest Symposium on Circuits and Systems 2007-MWSCAS 2007, 159-162.

[9] H. Wang, J. Xu, and X. Wu, "A Low Power Audio Delta-Sigma Modulator with Opamp-Shared and Opamp-Switched Techniques," in Proceedings of the International MultiConference of Engineers and Computer Scientists 2010-IMECS 2010, Ⅱ, 1296-1300.

[10] H. Wakaumi, "A Folded-Cascode OP Amplifier with a Dynamic Switching Bias Circuit," Engineering Letters, **23**(2), 92-97, June 2015.

[11] H. Wakaumi, "A Switched-Capacitor Filter with Dynamic Switching Bias OP Amplifiers," in 8th Latin American Symposium on Circuits and Systems-LASCAS 2017, 33-36, Feb. 2017.

[12] M. Takebe, J. Iwata, N. Takahashi, and H. Kunieda, Switched Capacitor Circuit, Tokyo: Gendai Kohgaku-Sha, Apr. 2005.

[13] U.-K. Moon, "CMOS High-Frequency Switched-Capacitor Filters for Telecommunication Applications," IEEE J. Solid-State Circuits, **35**(2), 212-220, Feb. 2000.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 107-118 (2017)

www.astesj.com

Special Issue on Recent Advances in Engineering Technology

ASTESJ

ISSN: 2415-6698

Cognitive Cybernetics vs. Captology

Zdenko Balaž*, Davor Predavec

Zagreb University of Applied Sciences, Electrical Engineering Department, Zagreb 10000, Croatia

ARTICLE INFO

Article history:

Received: 29 September, 2017
Accepted: 06 November, 2017
Online: 20 November, 2017

Keywords:

Captology
Cognitive Cybernetics
Intelligent Systems
Expert Systems
Learning Mechanisms
Motivation

ABSTRACT

In acronym Captology - Computers as Persuasive Technology, a persuasive component (lat. persuasibilibus - enticing) refers to the persuasive stimulation by intelligent technologies. Latter being transitive and interactive as intelligent systems, they have imposed, by their persuasivity, a 'cult of information', after which information has become a type of goods that as a utilitarian resource must be exploited quickly and efficiently. Such a widely accepted fact resulted as hype, presenting a perspective that the approach to a large amount of information and faster 'digestion' of their content will enable users to quickly get desired knowledge.

Recent investigations about persuasion processes have shown its dependence on intelligent technology factors (design, interactive computer products, web, desktop and others). Such technologies are also used to influence people's attitudes, beliefs, learning, and behaviour. Development strategies for global computer production and sales head in that direction and confirm latter statement with the promoted 3-P model: persuasive, permissive and pervasive components. Cognitive level of human integrated development is increasingly overshadowed by the contribution of artificial intelligence through its products, i.e. 'smart' creations, and by the array of shortcomings and problems that the same interactive technology brings. This paper presents a parallel between captological component of intelligent and interactive technologies on one side and illustrates examples of captological influences proved by confirmed trials within cognitive science through computer simulations of human thinking on the other side. Many studies have shown that the success of persuasion depends on the factors which have been exposed by cognitive cybernetics. Next to it, people's behavior system is transforming through the very development of society. Therefore, the influence of latter can be either positive or negative, while its extremes are already escalating in direction of a new trans-singularity and post humanism theories based on the principles of extropy.

1. Introduction

This article extends the work originally presented at 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, May 2017, [1-4]. The presented paper confirmed how human mind development is being contributed by the 'smart' creations of intelligent technologies as artificial intelligence (AI) products. Researches have been devoted to methodological directions and developing concrete tools for self-analysis and synthesis, simultaneously providing exhaustive theories about human nature being experienced as a 'new philosophy', but also as an 'anti-philosophy', defined on a new scientific course called cognitive cybernetics.

Cognitive cybernetics is the expansion of the 'classical' cybernetics, which includes scientific comprehensions of cognitive control and regulation of artificially-technical and socio-technical systems, biological organisms, organizations, processes, structures etc. The cognitive aspect of all evaluated systems elaborates and applies cognitive models, cognitive-functional and procedural system implementation, problem-solving based on natural methods by analogue cognitive computer techniques and lastly, the realization of system solutions and their organization in the context of management approaches and methodologies. This work is based

*Corresponding Author: Zdenko Balaž, Našička street 59, 10110 Zagreb, Croatia, +385 99 3111 460 Email: zbalaz@tvz.hr

on the cognitive development and implementation of ITK's (Ingenieurgesellschaft für Technische Kybernetik, Rülzheim, Germany) education modules in cooperation with National University of Vietnam, Hochschule Offenburg and Electrical Engineering Department at Zagreb University of Applied Sciences.

Cognitive Cybernetics recognize motivation as an internal state that excites, directs, and maintains the behavior of an individual. In education and training, motivation is especially important both for learning and for seeking the purpose and meaning. In academic activities or preferences, latter is described as meaningful and effort-worthy exertions, while tend to be realized and applied within importance of higher levels competences and skills. Those who are more motivated have a greater positive attitude on everything, and so on education; they persist longer in more difficult tasks, process information at a deeper level and excel in learning experiences. The classic distribution of motivation on intrinsic and extrinsic motivation in education is an incentive for activities which:

- represent the challenges of optimally assessed goals,
- can be controlled for goal achievement, which is emotionally fulfilling,
- establish the autonomy of action through the existence of feelings that can be controlled and have the effect on own development,
- encourage curiosity for the new, surprising experiences that increase the intrinsic motivation,
- include creativity and imagination where the motivation awareness is confirmed by imagination.

Approaches for defining motivation in education are:

- Behavioral approach in which understanding of motivation begins with a careful analysis of incentives and awards,
- Humanistic approach that emphasizes the understanding of the need for self-reliance through the stimulation of internal resources (sense of competence, self-esteem, and autonomy),
- The cognitivistic approach that accepts the fact peoples' activity and curiosity are inherent, and they seek information to solve relevant tasks.

An integrated approach for creation of education based on cognitive cybernetics foundation, considers:

- socio-cognitivistic focus on individual's behavior and potential interest for the impact on beliefs and expectations,
- the socio-cultural atmosphere created by participation in the community, in activities that maintain the identity of the community's interrelationships.

Therefore, motivation in the community is strengthened by the education to which it belongs and appreciates it. Specialist tutoring programs are the foundation for acquiring an expert status, where the very understanding of problem-solving is needed at that level, for which is necessary to gain the knowledge by understanding. Elementary forms of understanding are prescribed by the concept of objective mind or spirit, which guarantees the possibility of intersubjective communication at the basic level. 'Higher forms of understanding' occur when one internal difficulty or contradiction confronts with what is already known. It is then necessary to summon up and include the entire life structure, (establishing a relationship between life expression and its cognitive content), which can be found in its hermeneutical understanding, [5-8]. Research results show that understanding develops through information processing and modifies over a longer period. At a somewhat higher flexible cognitive level, it is possible to cognitively gain skills through which information is processed both intuitively and analytically. This knowledge, along with the persuasive context and information technology, closes the 'hermeneutic circle of artificial intelligence', Figure 1.

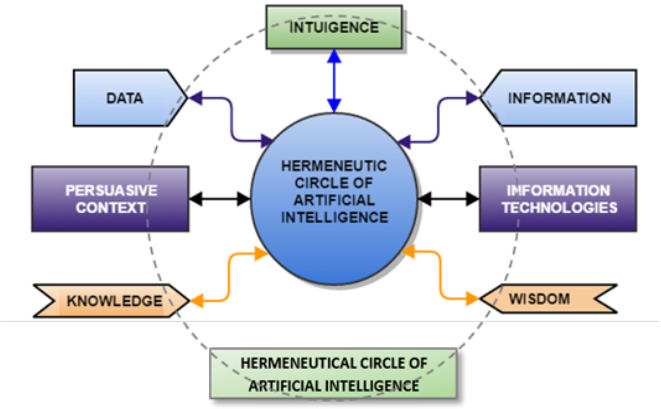


Figure 1 - Hermeneutical circle of artificial intelligence

2. Cognitive cybernetics – the future of education

Researches that deal with the captological components of intelligent and interactive technologies are confirmed by cognitive cybernetics through simulation examples of human thinking. Approaches to the study include analytical levels by evaluating the system almost in all technological branches and domains with the assumption of significant and dramatic advances in neuroscience, genomics, robotics, nanotechnology, computers and artificial intelligence with the help of other interdisciplinary studies, shown in Figure 2.

Studies and researches have further led to transhumanism extremes that advocates human body modification, including mental and psychological changes, all to control desires, mood and mental states. The everlasting fight against aging and illnesses is therefore not the only task of transhumanists, but also a try to improve the man at all levels, emphasizing simultaneous mental and physical advancements would be desirable.

If transhumanism associated with humanism is recognized as writing it as 'trans + humanism', then from the enlightened humanism comes the emphasis of a progress (its ability and desirability, not its inevitability), personal responsibility to create a better future and progress to a new society, but at the same time the awareness of accidents that can be caused by technology misuses in such society. For these reasons, it is necessary to differentiate the forms of meaning (i.e. primary and secondary), on which science assignments are based, and values that are recognized in the world, [9], addressed below in Figure 3.

Treatment by cognitive science has selected, among other things, certain factors of one's individual ability expectations to achieve desired goals and their value, respectively.

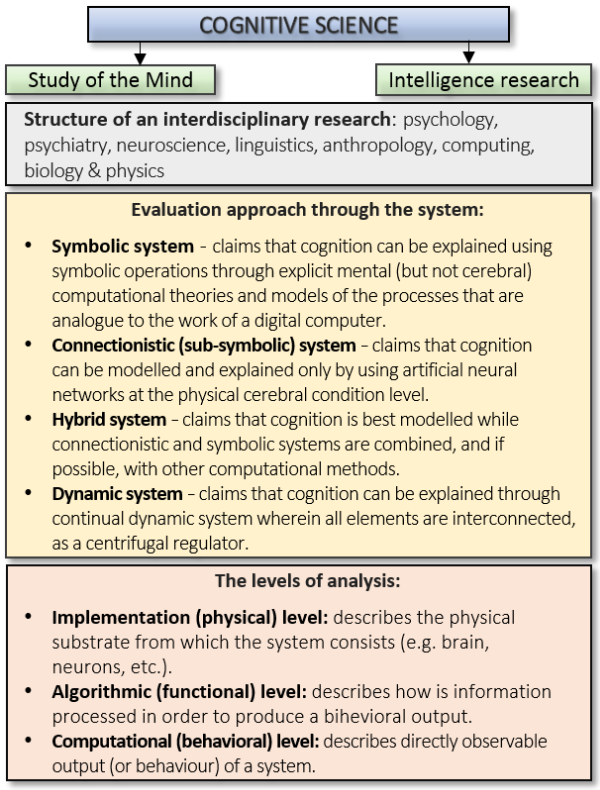


Figure 2 - Interdisciplinary study of the mind and intelligence through level analysis over the system

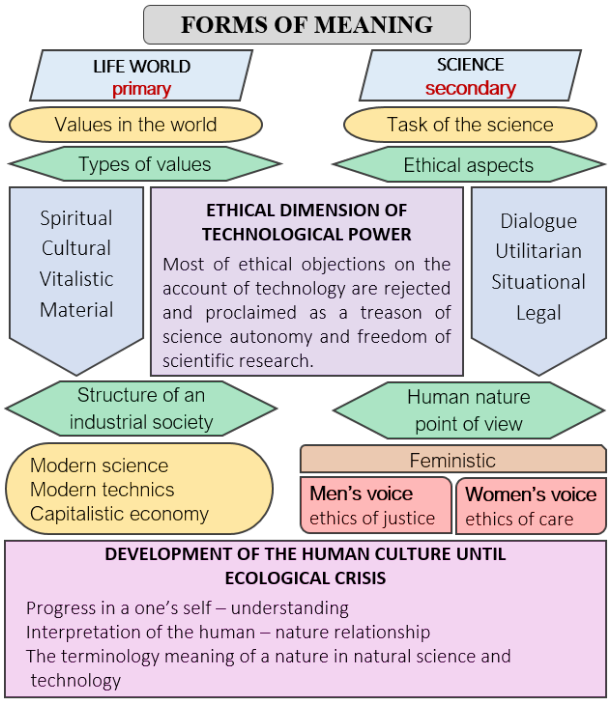


Figure 3 - Primary and secondary forms of meaning and ethical dimensions of technological power

Apart from this, it is accepted the fact that a man is a cognitive, reasonably-minded being, who accomplishes its own understanding at two levels. The first one is subjective level or the level of psychological participation, and the second is objective level or the level of intellectual participation. Hence, it is possible to conclude that hermeneutics is a universal theory of understanding in which no confusion is present because hermeneutic understanding is neither interpretation nor explanation. In this way, a differentiation between the objective understanding of the meaning on one side and subjective understanding of motives and intentions on the other side can be made. Theoretically, it is possible to attain adequate learning processes from the book, without producing any external signs of acceptance and application of such learned matter. The effects of learning efficiency, whose outcome must have been understanding, depend on many factors such as physical and biological states, prior knowledge, daily oscillations in the brain chemistry, peptide hormone levels, and ultimately – the captology. Therefore, in understanding the recognition of its either elementary or higher forms for the purpose of education system, it is important to point out the most important motivational variables, namely:

- needs,
- beliefs,
- goals as outcomes to be achieved,
- internal interest,
- emotions.

These variables are linked and processed through natural analogue systems, confirming the new concept in cognitive cybernetics – 'Brainware', [10].

2.1. Needs

Although there are many theories, all rely on the most famous one – the Maslow's hierarchy of needs that differentiates the needs of a lower order or the needs of lack - survival, security, affiliation, and self-esteem, and the needs of higher levels or the growth needs - intellectual success, aesthetic assessment, and self-realization. The latter needs can never be fully met, since when one desire is fulfilled, the other one immediately takes its place. Newer approaches to the motivation influenced by intelligent technologies focus on highlighting not only individual needs (a generated need to feel capability and competence), but also on interaction with the world, having the choice and feeling of controlling life, while at the same time on the challenge of connectedness through sustainability in the world. Such a theory, though seemingly presents the absurd paradigm because it reduces the basic human needs to autonomy, competence and connectivity, relies on the persuasion which is the second name for the captology present in cognitive cybernetics, Figure 4, [11].

2.2. Beliefs

Beliefs as a variable of cognitive cybernetic processing come down to the achievements of:

- knowledge,
- capabilities,
- causes and control,
- self - value.

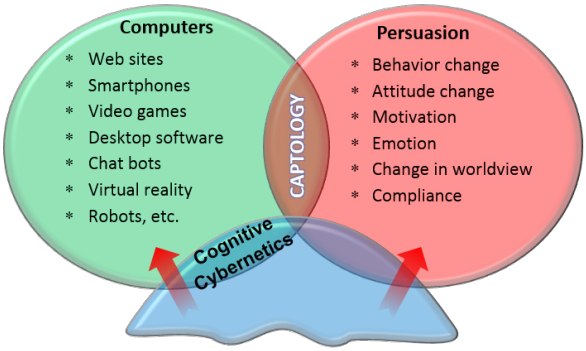


Figure 4 - The captology of cognitive cybernetics in computer persuasion, [11]

Knowledge beliefs are defined as beliefs about the structure, stability, and security of knowledge and as best ways of acquiring knowledge that influence learning strategies use.

Capability beliefs are one of the most powerful beliefs that affect motivation. There are two types of capability views: entity and incremental. The entity view refers to the belief that capability is a fixed characteristic and can not be changed. Although capabilities differ between individuals, their representation is potentially stable when acquired, but capability as a set of skills may change. By means of continuous work, learning and exercising, it is possible to increase one's knowledge and thus improve the capability.

Causes and control beliefs are based on attempt to understand the meaning of one's own and/or another peoples' behavior. This theory explains how individual explanations, justifications, and excuses about oneself and others affect motivation.

Self-value beliefs can be oriented to control or avoid failure. Orientation toward control seeks out the tendency of addressing value to achievement and the ability to set the attitude that achievement can be enhanced by targeting the goals. By mere control, one's own skills and abilities are being improved.

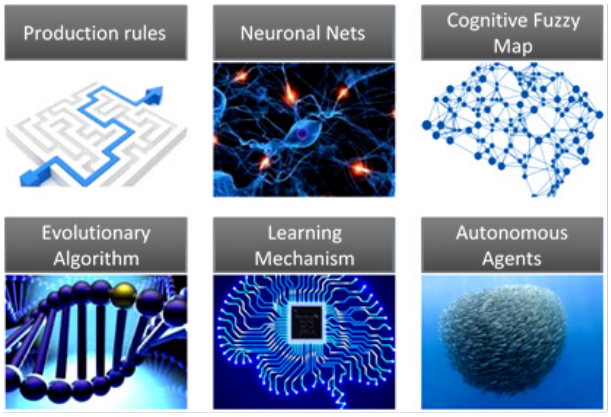


Figure 5 - Concept of cognitive solutions in environment, [11]

Orientation to avoid failure supports the entity's view of ability and sets further its goals related to performance. One can have a lack of strong ability sense and value separated from the performance, thus cognitive cybernetics serve as a simulation for technological and technical solutions (production rules, evolutionary algorithms, autonomous agents etc.), for the purpose of understanding and functioning of biological states (neural networks, learning mechanisms, cognitive fuzzy maps, etc.). In this way, mapped into education systems, strategies for avoiding failure are being worked out so they don't lead to a failure or to prevent the continuation of failure. Such strategies (shown in Figure 5) do not accept excuses or incapability, therefore the possibility to give up or accept a failure as a belief that problems arise due to low capabilities is not considered.

2.3. Goals

Goals are the outcomes or achievements to be attained and they enhance performance because they direct one's attention on the current tasks and divert from the distractors. Goals encourage devotion, for as the goal is as greater, such is the commitment. They increase perseverance, for when there is a clear goal, it is more difficult to give up. Furthermore, goals also encourage the development of new knowledge and strategies until they prove as successful ones.

Types of goals being set affect the extent of motivation that is needed to achieve them. Aims that are specific, moderately tough and likely to be achieved soon, often increase motivation and persistence. Just as goals, there are target orientations related to belief patterns about goals, also associated with achievements. Goals differ from the type of control, performing, avoiding the effort and social goals that need to be recognized for motivation.

The concept of cognitive solutions that today's civilizational environment treats as an agent systems by cognitive coverage of processes and programs, integrates human factors and machines (technologies) as inspirational cognitive goals, Figure 6.

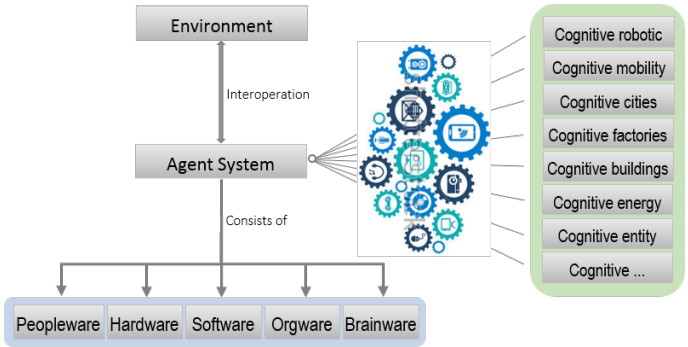


Figure 6 - Concept of environment cognitive solution set as cognitive goals, [11]

2.4. Internal interest

There are two kinds of interest: personal and situational. Personal interests relate to more persistent personal aspects, such as a long-term preference for the particular subject, while situational interests relate to short-term aspects of activity, such as a text or material that attract and retains attention. Interest in the education grows with a sense of competence, even if initially there was no interest in a certain activity. Cognitive industry, by

its evolution, is an inexhaustible source of interests, and as such through modern civilization imposing itself as a leading variable. Within cognitive cybernetics processing and understanding, interest is related to entities (synonym for the subject, article, individual, being, form of life, body, creature, organism etc.), which in ontological terms is connected with different and independent existence plus applicability, and is profiled as a cognitive entity platforms, entity ecosystems, browsers, entities, services, service portals, interoperability servers, maintenance portals, integrated server, entity application as a tool, etc., Figure 7.

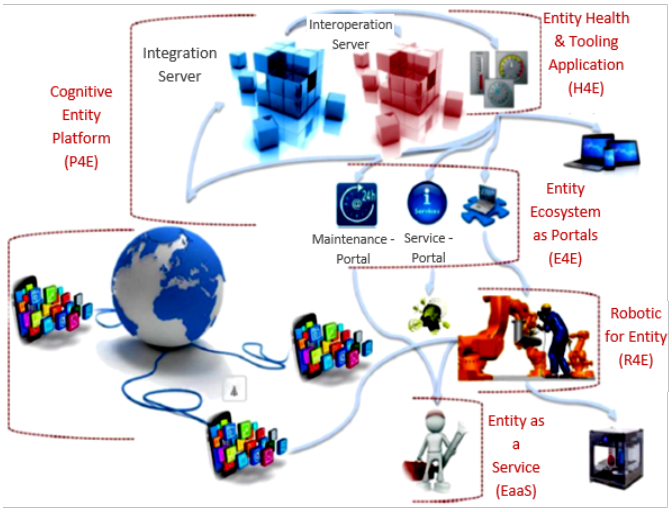


Figure 7 - Platforms of creating cognitive solutions as goal variables, [11]

2.5. Emotions

Emotions, especially pleasant ones, have very positive effects on ability beliefs (accepting an incremental view of ability), target orientation (associated with learning objectives), self-regulation, information processing (emotions are encoded with information), etc. Positive emotions encourage positive thoughts and memories, along with evaluation of education, hence such a positive effect is adequately treated for evaluation and performance of education based on the concepts of a system cognitive modelling.

In the beginning of education, emotions are very important, especially when it comes to the feelings of excitement, pleasure, fun, and curiosity, because these emotions can trigger a situational interest. It is only necessary to control them in order to prevent exaggerating effects which can lead to negative emotions, having a somewhat more complex relationship with motivation.

Additional stimulus to emotional motivation is the over-occupancy that refers to an intense form of intrinsic motivation in which an individual is in a state of profound insight, concentration and focus on the challenging activity. The experience of occupancy depends on autonomy. Intrinsic emotional motivation was often experienced in activities that were chosen willingly rather than imposed by others. Therefore, competence (skills and techniques) is so trained that it becomes automatic and is responsible for the optimal challenge (the optimal ratio of challenges and skills), creating a special emotional concentration that is not obscured with time and space.

3. Cognitive cybernetics in mystical artificial intelligence implosion

When artificial intelligence emerged in such terminology in the middle of the last century, it was the fastest growing scientific discipline. At the beginning of this century it started to be mystically missing in implosion, which can be attributed to a certain transformation. Development of artificial intelligence encompasses four periods of time, [10]:

- The period of enthusiasm and great expectations (1952 - 1969),
- The Age of Cognitive Reality (1966 - 1973),
- Knowledge-based Systems (1969-1979),
- The period of applicability (it begins in 1980s and continues as a process to date).

The fourth period is a process that goes through its five stages of transformation:

- The first stage is application of artificial intelligence in industry, from 1980s,
- The second stage is a transformation through neural networks, introduced in 1986,
- The third stage of transformation has placed artificial intelligence in the science, and is accepted since 1987,
- The fourth stage of transformation introduced super-intelligent agents, in 1990s,
- The fifth stage of transformation is the mystical implosion of artificial intelligence that is still ongoing.

Mysticism in this matter is trying to point out that implosion is a way in which artificial intelligence disappeared within itself, and at the same time in some certain connectedness, cognitive cybernetics emerges. In that mysticism lies the impossibility of terminological discernment and a clear access to technological level called 'Industry 4.0', as well as transition and emergence of the new era of 'Evolution Industry - Industry X.X.', in which cognitive cybernetics has been already implemented. Therefore, regarding such a rapid pace, it can be concluded this already 'is' the future so widely present as a subject of discussions.

Today, this future is also called the 'Society 5.0', (germ. 'Gesellschaft 5.0'). The Idea of Society 5.0. was launched from the 'Fifth Scientific and Technological Fundamental Development Plans' in Japan by the end of 2015. The idea is to bring digitalization closer to all segments of society through the creation of a so-called 'super-smart society', whose core is the social impact of technology and which is focused on cognitive cyber-products. Production processes and production itself of these cyber-products are created by developed and self-regulated autonomous cycles, partly realized through the very products.

Toward adoption of the Society 5.0 strategy, on CeBIT meeting in Hannover, 25th of January, 2017, the Japanese Prime Minister (presented Japan as the partner country) and the German chancellor agreed about opinions that although the market

prerequisites for further development brought by technology such as open borders, democratic values and a free trade, they stated: '*Millions of people do not know what to expect from the upcoming changes, and those who should change the policy and the digital sector in Europe which relies on cognitive cybernetics are often slow in key decisions*'.

3.1. Transformation approaches of cognitive cybernetics

The mystical disappearance of artificial intelligence presented above and characterized as an implosion, further shows that the emergence of cognitive cybernetics can be studied through three approaches to that transformation, as addressed below.

A. Anthropological approach

The anthropological approach of artificial intelligence into cognitive cybernetics transformation relies on the study of a human species development (life and culture), which looks at the presence and future of the human race, and links:

- nature and culture,
- past and present,
- the uniqueness of the human race through diversity of its forms.

The relationship between nature (biology) and culture defines the development of human species through cortex specialization, which enables understanding, communication, and learning. Thusly created (learned) content is called culture. As a fundamental concept of anthropology, culture is a set of behaviors that people learn and share in a given period and in a chosen natural and social environment.

The relationship between past and present although in the evolutionary sense compared to other beings follows the same natural rules, enriches people with extraordinary ability to learn and pass on the knowledge from generation to generation. This approach has accelerated the cultural evolution which has become faster than the biological evolution. In this sense, profound differences are being outlined, not only in terms of anatomy and physiology, but also in terms of the very habits so different that the essential unity of the human race is not seen anymore, and therefore the potential for innovation is harder to be recognized.

The uniqueness of human race through diversity of its forms is a consequence of the lifestyle difference not only in different times and areas, but also within the same society due to regional specialties based on the economy and different living conditions as entities. Within the same area, differences exist in the way of certain communities' lifestyle, not only immigrants' ones, but also among the sexes, ethnic, age and class groups. Responses about purposefulness inquiries given from anthropology are used to understand, create and inspire cognitive cybernetics.

B. Ontological approach

Ontological approach to the concept of cognitive cybernetics relies on strictly traditional philosophical discipline as the central branch of metaphysics that studies the being, if it is a being, as such. Modern analytical ontology is, above all, the theory of general categories such as object, attribute, event, etc. Ontology is a study of nature of the things we find in our environment, around the world, and what relationships are among them. In addition, it is related to existence, objective property of things, space, time, causality, and opportunities that result from all of this. Since cybernetics within science explains the general principles of the process management, regulation, obtaining, storing, transforming and transmitting information, both in living beings and in technical systems, therefore as a scientific discipline explores a dynamic self-regulating and self-organizing systems, while through ontologic patterns also defines system properties. In this way, accessibility is recognized in environment and by means, changes as responses within. Activity as operation is directed towards achieving the goal through takeover of the initiative. Socialization that works together with other systems and people to achieve goals also helps others in achieving theirs. Furthermore, mobility opens the possibility of acting in any environment and lastly, a learning ability based on understanding human intelligence and communicating with machines ensures exceptional availability.

Ontological approach enables expansion and connection with boundary theory areas of a system, management, information, coding, logic, formal languages and grammar, games, mathematical algorithms and programming, and robotics, which include all four existing types of cybernetics:

- Theoretical or general cybernetics,
- Applied or applicable cybernetics,
- Technical / technological cybernetics,
- Economic cybernetics.

Their most important features are the features of the system:

- a great complexity,
- stochasticity and
- auto-regulation.

Recognizing these cybernetic characteristics and features enables ontological observation through its theoretical, technical and economic views.

Theoretical point of view represents the logical structure of hypotheses, knowledge, discoveries, scientific facts, theories and laws. It involves observing, collecting, measuring and rendering data, and lastly, data experimentation, for ensuring optimal conditions under which data, information and knowledge can be obtained. The economic or organizational standpoint provides rational technology intended for further research in all essential elements.

C. Captological approach

The captological approach to cognitive cybernetics includes the cognitive character of persuasive technologies and as a superior scientific field introduces expansion limes on ethical principles. This approach to cognitive cybernetics through its purposefulness in industrial ecology, an emerging scientific interdisciplinary field, unites its development by combining natural, technical and social sciences into a unique system from

global to local level. Foundation of this concept is a cognitive captological integration based on analogy between intelligent technologies, processes in nature (biosphere) and processes in society (technosphere), Figure 8.

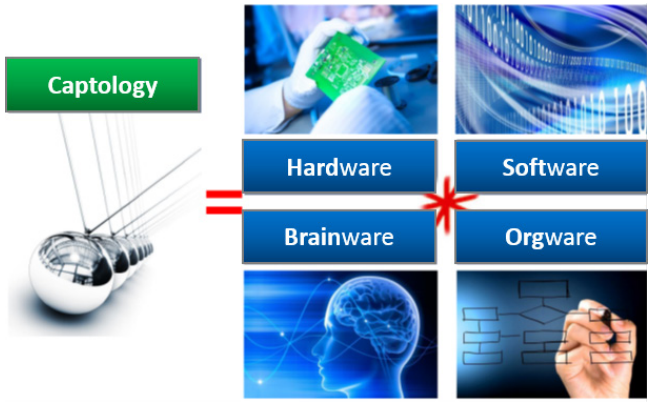


Figure 8 - Captology as a threat to innovation's access and openness

The biosphere, with its laws and principles, refers to coniferous life around the planet, summed up in earthly life as a whole in which all transformations take place. Technosphere is an artificial sphere of all kinds of technology around the planet (all technical and technological aids, creations and systems), with its own artificial time functions in which everything is immersed. Today, the ultimate development and application of intelligent technologies is experienced as a planetary intelligence system made by the Internet, mobile communications, applications and virtual structures of technological simulations. Cognitive cybernetics might enable the creation of awareness of innovations that will drive evolution of life on Earth only in a mental and spiritual direction, [12-15].

3.2. The role of human factors in cognitive cybernetics

Cognitive cybernetics has advanced to such a degree that it can do all the imaginable management and monitoring in a hardware sense. For its efficient application, education is needed to acquire special skills, knowledge and competences required for unerring behavior of decision-makers in approaches to accept innovations and consider scrupulosity. Herein come invaluable characteristics of the ability to use a 'common sense':

- Expressive abstraction and conceptual analysis for quick and successful problem solving in unexpected situations
- Unlimited flexibility.

Certain problems can be identified through persuasion perfidy in the context of treatment and understanding data and information. Research has confirmed that 'third-party tools for experiments on users from database software algorithms evaluate and rank all information accessed through the network - currently 95% for which an average educated user reaches!? Thus, it is not surprising fact that over 300 trials and about 200 signals are continually being run on a daily basis, watched and measured to help maintain 'freshness' and the speed of collection, securing, and transferring data. Although scarcely mentioned, but shows persuasion as useful, the captological arsenal is revealed as:

- Cognitive prejudices, (between more than 170 cognitive prejudices emphasize the 'unwillingness of taking risk in time deficit', and which has been led by persuasive message),
- The weapon of persuasion (this is about persuasive techniques that rely on the 'principle of scarcity' and the 'theory of authority acceptance'),
- Productivity laziness (confirmed by the fact of human mind laziness and the tendency to skip or careless reading which, due to the persuasive guidance, results in unwanted consequences),
- Decision-making fatigue, (associated with the user exhaustion by persuasive over-offerings that encourage the use of lower resistance paths).

The brain capacity for learning is enormous and created for more demanding intellectual activities than the ones commonly included into. The human brain possesses two mutually independent systems of receiving and thinking. The leading one is in the left half of the brain and it analyses, writes, reads, talks and receives logically. The main center for feelings in the right half of the brain is more complete, emotional and associative, or better said - intuitive. Selected activities of the left and right hemispheres of the brain help to its learning by the pairs model (challenges and feedbacks).

The right side of the brain is used for intuitive handling of problems based on two components:

- Simple provisional rules that exploit the proviso to hit the most important information but ignore everything else, hence allowing a quick reaction,
- Predispositions of evolved brain abilities - skills acquired by genetics and learning to transform predisposition into the ability.

As the human mind develops regarding to inexhaustible desire to try and change, not just by wanting to always follow only 'reliable' answers, the question is whether the brain can be stimulated to develop when captologies is not new but uniform, deadening situations.

3.3. Conducted trials

In addition to aforementioned persuasive addictions, the results of 'tests/exercises' can be joined from originally elaborated materials of the course Intelligent systems, conducted at the Zagreb University of Applied Sciences, Electrical Department. It is basically about mechanisms and elements of conclusion through knowledge bases and databases of an expert system. Conclusion is gained by applying acquired knowledge about distributed artificial intelligence in correlation with human intelligence and social predisposition. The task is precisely described by instructions each student receives in the envelope, and before starting, teacher reads the instructions aloud before all

students at the initial location of the exercise. The task is shared to two groups of students who, after receiving the instructions, go at two different locations. Students have drawn attention for having the opportunity to cooperate with each other, which is the goal of exercise, at last. There are five simple questions needed to be answered in manner that all students participate and create common, identical answers. The guideline text is divided into three parts (task description, approach to solving and solution guidelines). Conducting this exercise within four educational periods in three academic years, 148 students participated, and the task was successfully solved not even once.

The reason why such a simple task has not yet been solved lies in the hyper-offer of data (information). Students are mostly taught that use of information technology encourages rapid data inspection. By doing such, reading of the task is reduced to mere browsing and is very superficial, which makes easier overlook of the solving approach. Too much data drives away from deeper and more permanent entry into any single argument, fact or idea. An overwhelming number of information and characters nullify themselves. Students are apt to access the search by going to links (connections) irrelevant to problem solving, but seemingly interesting information is the very beginning of current potential termination of available concentration. This outwardly produces impatience, a lack of attention and the absence of co-operation, respectively. As in reviews and surveys of a vast data, users are prompted to quickly come and go; the same effect was achieved through the lumped text of this task. In this journey of a rapid collection and data transfer, the cognitive efficiency is lost, unfortunately. A lot of confusing information and attention-shifting parameters fragment, hinder and distort the flow of thoughts that end up as distraction. The use of mind is under influence of intelligent technology and therefore under the pressure. Part of the human mind in charge for critical thinking, analysis of reality and which should be used daily as a way of thinking in solving problems, but also in social interactions when it serves to connect with the outside world, doesn't come to the expression in this task, and there was no created understanding – mind is lost in captology for it does not follow the evaluation sequence.

A. An example of understanding when meeting other people

As an example of understanding, a study was conducted related to learning of understanding when meeting other people, already discovered by William Outhwaite, who had been looking at such situations through four categories:

1. The category of physical facts that can be noticed (how people look and what is the consequence of that). These are understandable phenomena in relevant sense, when being seen as a sign of something not physical from the past or something related to the mental state.
2. The category of mind states in others that can be identified now or can be inferred based on the previous comprehension.
3. The category of others' actions (by noticing what others do and what they consider to be doing) is described with three sub-categories:
   a) Recognizing an action by approaching from distance,
   b) Accepting of what is being said about on the same elemental level,
   c) Complex judgment of the real importance of what is being done or what is thought to be said;
4. The category of questions why certain things are being done, and what motivates people to do them.

Latter categories try to establish a difference (especially between the states of third and fourth category), or the differences between motives and the 'hermeneutic understanding'. Psychological understanding of the human mind mental states can be observed based on three conclusions about peoples' motives and their intentions, which are the following: a) visible signs, b) explicit statements, and c) knowledge of the 'situational facts', confirmed by the psychology of the situation that has been discussed since Aristotle, Figure 9. From the aforementioned examples a complexity of the understanding comprehension can be seen, which is theoretically evaluated; and people have a social process of meeting embedded in a sub-conscious algorithm operable in real-time.

B. An example of learning understanding while reading aloud

To confirm the conducted research, it is useful to consider the acquisition of understanding on research example of literacy development. Therein, it comes to conclusion that the central process of learning by reading is understanding the text and active search of meaning and purpose, as established by G. J. Whitehurst and C. J. Lonigan. The reading process or understanding what was read comes down to understanding of coding and decoding, like digital data transmission. This also presents the concept of an expert system functioning, which was followed herein for explanation and elaboration.

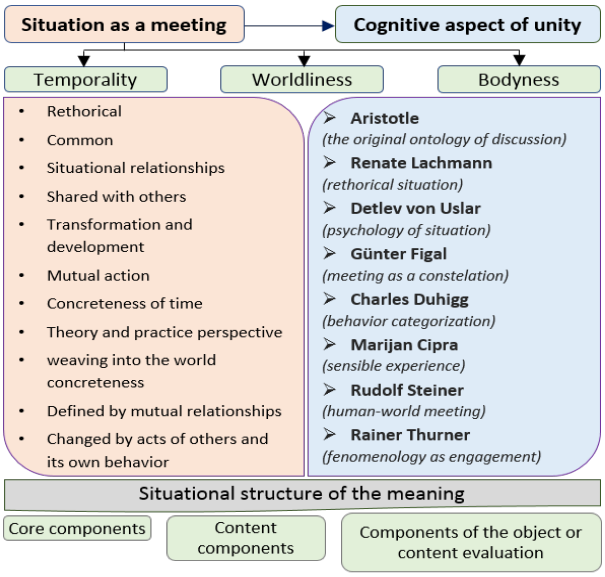


Figure 9 - Analysis of the situation psychology as a situational structure of significance

Researches on reading effects has set the groundwork for reading mechanisms evaluation, which are:

- a comparison,
- a recall from the memory,
- linking what is seen or heard, (or heard before),
- identifying essential data and rejection of redundant ones.

On these grounds, data processing while reading developed by R. J. Marzano and D. E. Paynter happens simultaneously with five concurrent processes, known as a 'Model of 5 babushkas', Figure 10, which comprises:

- Understanding the purpose of reading,
- Estimating the compliance of read data with the purpose,
- The meaning of the smallest thinking units (claims),
- Identifying the meaning of words,
- Recognizing the meaning of a read wholeness.

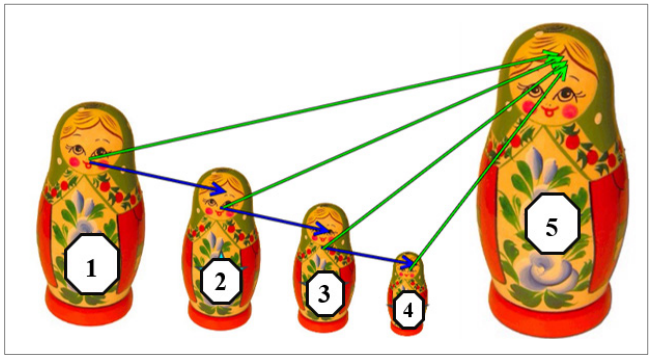


Figure 10 - Data processing while reading, (the model of 5 babushkas)

Reading, listening, watching and pronouncing are the ways of learning and each has a different contribution to learned matter. In accordance with learning mode, the left and right hemisphere along with emotional center of the brain are activated. Understanding the reading comprehension (reading aloud) is an understanding strategy that along with understanding, implies the voice sensitivity and motivation development. The above-mentioned examples, especially data processing while reading, is an important foundation for analyzing different types of reading, such as:

- mature reading,
- reading a linear text - a printed book,
- reading a linear text - e-book,
- reading by text search – a computer,
- reading hypertext – a computer,

and where reading should be distinguished from a mere search, under the influence of captology and turned into an addictive need, because it maintains a worthless dominance over digital information flow. The imposed 'Internet ethics' through different kinds of media results in persuasive addiction manifested by:

- encouraging and treating a 'thirst' for small and short-term information slots,
- curious dependence on a continuous data stream updated in the real-time; an 'updated status' which lost its purpose slightly after a few moments of release, and
- the desire to speed up the flow of information.

Aforementioned manifestations hold not any contribution to storage of learned, because nothing has been learned. Moreover, a disturbance has occurred, and no recognized brain cell nor emotional center has been activated, which today can be safely confirmed by psychometric investigations for more serious (i.e. legislative) use, still deliberately delayed, while minor trials so needed have not yet been clearly recognized and are ongoing through new researches.

C. An example of a thought material storage

In support of the researches conducted, there is also an example of storing thought material in the memory, which requires a careful and balanced information flow. An example can be visualized for the practical acceptance. It is easily possible by photography, Figure 11, to identify a chair (or chairs) that represents a thought material which can be carefully and permanently stored in a visualized mental space.

Figure 11 - The HAZU[1] revival hall, Opatička 18, Zagreb - an example with a careful and balanced storage of thought material

An example of a mess explains a fragmented, disturbed, and a ruined flow of thoughts that end up as a distraction. Hardly anyone will be able from the photos, (Figure 12 and Figure 13), to recognize the thought material at the glance, the 'seat', which should be stored in the permanent memory. Careful and balanced storing of thought material in a chaos is impossible, because from a visualized matter a particular psycho-culture is participated, lately projected in all spheres of life.

The term 'psycho-culture' comes from the culture (lat. *colere* - cultivate, nourish, *cultos* - processing, nourishing, cultivating), which represents an integrated system of attitudes, beliefs and behavior patterns characteristic for the members of society and which are not a result of biological heritage, but a social product created, transmitted and maintained through communication and

learning. The function and meaning of the culture are facilitation, extension and development of the human society.

Figure 12 - The courtyard of the house, Trešnjevka in Zagreb - an example of thought material storing represented by mess and chaos

Figure 13 - The environment of the building, Hrgovići settlement, Trešnjevka in Zagreb - an example of thought material storing represented by mess and chaos

Culture arises from the peoples' need to master the nature and to be free from animalistic life, so that man does not cultivate only external nature but also his own, thus achieving the process of humanization. Under various influences and so under the influence of intelligent technologies, culture is subjective to collective programming of the mind by which members of one community differ from others. Psycho-culture associated with Cicero's cultivation of the soul from ancient times (lat. *cultura anima*), I. Kant evolves and develops it through his 'categorical imperative' as a noble trait, but also recognizes the formation of a collective mentality in society achieved by a dominant ideology and which directs social and political processes. Each person has its own ideology, a system of values and life orientations that direct through life and determine the behavior influenced by intelligent technologies and intuigence (intuition + intelligence), [16-23]. Socio-cultural neuroscience provides better and better insights into complex interactions between socio-cultural and neuronal structures and processes, and thus within biological mechanisms that underpin socio-cultural phenomena and human behavior including all components of Kierkegaard's existentialism.

For these reasons it is useful to point out that the structure of neural network, part of the big brain cortex, represents the biological basis of our mental abilities. In adulthood, new knowledge and skills are easily adopted without changing the structure of the neural network. On the other side, educational, psychological, social and emotional environments during such development change the very structure of the neural network, especially in the brain parts crucial to adopt the most complex brain functions. This includes the affective modulation of emotional expression, the conceptualization of one's own mind, mentalization, cognitive flexibility and a work memory.

4. Continuation of research

The continuation of researches presented in the paper is based on the prepared teaching material, [24-28]. In the new education period part of the Intelligent Systems program is transformed into a new area, [29-34]. This new area will be presented at the 1st International Conference on Intelligent Human Systems Integration: Integrating People and Intelligent Systems (iHSI 2018), which will be held at the JW Marriott Marquis in Dubai, United Arab Emirates, January 7-9th, 2018.

Authors in their work have come up with recognizing:

I. Human Intelligence Systems, (HIS) - at the very beginning of the 1980s, researches in the field of artificial intelligence was focused on systems associated with human intelligence. Their goal was to incorporate the principles of human intelligence into intelligent platforms for data processing through artificial intelligent systems, which in turn resulted and actively contributed to the industry development and advancement of society, such as:
   a) advanced development and use of mechanical systems and devices in industrial production as intelligent autonomous robots,
   b) the introduction of humanoids into the social and palliative spheres of life.

II. Intelligent Human Systems, (IHS) - researches continued in the 1990s, conducted on already developed intelligent systems foundation with the intent of implementation in education, had included among others:
   a) developing and upgrading intelligent interactive technologies for use in systems with algorithms involving the principles of human thinking,
   b) scientific analysis of social activities and human intelligence through mathematical modelling, scientific brain research and cognitive sciences in general.

Such intelligent systems contain the knowledge of certain domains with sophisticated decision-making processes and the ability to explain their behavior. This feature shows the most important aspect of such systems - their ability to effectively interact with people in social environment, teaching, and assisting in availability and complex data manipulation. Furthermore, their use has resulted as a significant impact on human resources with the effect of increased demand for education, [35-40]. Knowledge (when considering Human Intelligence Systems / Intelligent

Human Systems) on the moral plane is designed to achieve what is achieved physiologically by asking the following questions:

1. How does consciousness, starting from the physical plane, finally reach the moral one?
2. How is physical balance a first step towards moral balance?
3. How is every controlled act the path towards improvement?
4. How are the rules of control also the rules of training?
5. How there is no discussion without calmness?
6. How does raising a will as a freedom and reasonable energy give man the freedom, right and task to choose between good and evil?

Although there are some suggestions today, it is more probable that in the near future answers to above questions need to be considered taking into account the parameters listed in Table 1, [41-45].

Table 1 - Parameters of the future considerations and justifications of Human Intelligence Systems / Intelligent Human Systems areas

Genetic neo-construction	Transhumanism	Ultra-Intelligence
Genetic Technology	Singularity	Super-Intelligence
	Extropy	
Biotechnics	Cyberpunk	Historical apokatastasis
Bionics		
Bioethics	Posthumanism	Human-biological planning
Biopolitics	The epoch of hypermodern	Spiritual programming
Bioconservatism	Nanotechnology	

5. Conclusion

Education for acquiring knowledge, skills and competencies will experience drastic changes in the near future because it will have to rely on cognitive cybernetics. In this paper, the results of research are presented as elaborated concepts of internationally recognized ITK institutions that launch new education programs with their partners. The Electrotechnical Department of Zagreb University of Applied Sciences is one of the partners since proven by its approach to educational standards formation and study programs alignment in polytechnic areas with the needs (existing and future) of the labor market, all in the context of globalization and technological development. Regarding to changes of higher education environment and the needs of labor market, it is necessary, as soon as possible, to harmonize them with the principles of cognitive cybernetics. This alignment must necessarily improve latter with introduction of new cognitive education modules because such activity of educational content should be a research for which human factors are at the same level of perfection and reliability as the available cognitive cyber-technology.

When information processing, problem-solving and decision-making are compared with expert system that contains human cognition and mental potential, it is possible to pull a parallel with intuition and intelligence.

The knowledge base is 'cognitively' dependent, associated with symbolic representation, and is identical to intelligence. Applying databases and knowledge bases, it would be useful to subject it to conclusion mechanism. This in biological sense of identifying with the human brain functioning (the way of thinking) suggests that purposeful use of intuition and intelligence require some sort of a skilled strategy – the intuigence, which has been confirmed by outputs of detailed researches carried out from 2008 to date, and which are presented in this paper.

When the education, psychological, social and emotional environment are linked to the presence and intimacy of intelligent systems and interactive technologies, the hypothesis of their multiple influences is confirmed; in a positive sense – through the knowledge of raising the level of integrated human mind development, and the negative one – the influence on people through neuropsychiatric disorders development. Biomedical and sociologically important concept is one of the main subjects of continuous discussion between advocates of selective theory against advocates of constructive theory of mental processing. Nowadays, the question about how to use the mind and intelligent technologies seems to be the same one. As one part of the human mind is in charge for critical thinking and reality analysis, it is used for problem solving but also in social interactions when it serves to connect with outside world. The second part of the mind is used for imagination and can only be triggered when the analytical opinion is excluded. The notion that imagination and analytical thinking are not treated as two opposing forms of thinking, but as two kinds of opinions with different functions, is the result of intelligent technologies impacts, for which human intelligence should set some directions and advices. For these reasons the question is set: does between cognitive cybernetics and captology exist a cooperation or the conflict?

References

[1] Balaž, Z., Predavec, D., "The Captology of Intelligent Systems" in 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, May 2017. http://dblp.org/db/conf/mipro/mipro2017

[2] Balaž Z., Meštrović K., "Learning and teaching from Artificial Intelligence", (article in Croatian with an abstract in English), Zagreb University of Applied Sciences; Polytechnic & Design, Vol. 2, No 1, pp 9-14; ISSN 1849-1995; Zagreb, 2014. http://polytechnicanddesign.tvz.hr/?page_id=363

[3] Balaž Z., Predavec, D., "The Brain lost in the space of persuasive no-placelessness", (article in Croatian), Public lecture and workshop conducted on 16th Brain week in Croatia, at Electrical department of Zagreb University of Applied Sciences, specialist graduate study, Zagreb, Croatia, 15th of May 2017. https://www.bib.irb.hr/867247

[4] Predavec D., "The Captology of Intelligent Systems", Master's Thesis, Zagreb University of Applied Sciences, 2017.

[5] Afrić, V., "Theoretical basics of hermeneutics", (article in Croatian language), Sociological review, Vol XIV, No. 3-4, UDK 167.7:316, pp. 191-200, 1984. https://hrcak.srce.hr/155800

[6] Heidegger, M., "Ontology (Hermeneutics and Factuality)", Bröcker-Oltmanns; Frankfurt a. M., 1988; Introduction to philosophical hermeneutics, Darmstadt 1991. http://www.iupress.indiana.edu/product_info.php?products_id=84761

[7] Wilber K., A Theory of Everything: An Integral Vision for Business, Politics, Science and Spirituality, SHAMBHALA, Boston, 2001.

[8] Lachmann R., Phantasia Memoria Rhetorica, (book in Croatian) Translated by Davor Beganović, Press: Matica Hrvatska, ISBN 953-150-150-560-8; Zagreb, 2002.

[9] Balaž, Z., "Cognitive Cybernetics - a Future Society 5.0", the paper prepared for International Symposium P&P, Zagreb, Croatia, October 2017.

[10] Haun, M., "Cognitive Computing - Steigerung des systemischen Intelligenzprofils", (article in German) ISBN 8-3-662-52951-5, DOI 10.1007/978-3-662-52952-2, Springer - Verlag Berlin, Heidelberg, 2014.

[11] Balaž, Z., Haun, M., "Cognitive Cybernetics – a future that is started" in 34th International Symposium on New Technologies, SOLARIS Hotel Niko, Šibenik, Croatia, 15th -16th May, 2017. https://www.bib.irb.hr/898570

[12] Teilhard de Chardin, J. M. P., L'Energie Humaine, (book in French), Editions du Seuil, Paris, 1962.

[13] Carr, N., The Glass Cage: How Our Computers Are Changing Us, (book in Croatian), Jesenski i Turk, Translated by Ognjen Strpić, Publisher: Znanje d.o.o..; Biblioteka „42"; Zagreb, November 2016.

[14] Detlev v. U., Psychologie und Welt, (book in German) Verlag W. Kohlhammer, Stuttgart, Berlin; Koeln, Meinz, 1972.

[15] Gardner H., Kornhaber M., Wake W., Intelligence: Multiple Perspectives, Hercourt Brace College Publishers, Fort Worth, Philadelphia, San Diego, New York, Orlando, Austin, San Antonio, Toronto, Montreal, London, Sydney, Tokyo, 1996.

[16] Balaž, Z., The Intuigence and Captological Traps of Artificial Intelligence, manuscript of the book at a review process, Zagreb, November 2016.

[17] Sternberg R. J., Beyond IQ: A Triarchic Theory of Human Intelligence, Cambridge University Press, Cambridge, 1985.

[18] Sternberg R. J., The Triarchic Mind: A New Theory of Human Intelligence, Viking, New York, 1992.

[19] Anderson M., Intelligence and Development: A Cognitive Theory, Blackwell, Oxford, 1988.

[20] Figal G., Der Sinn des Verstehens, Beiträge zur hermeneutischen Philosophie, (book in German) Philipp Reclam jun., Studgart, 1996.

[21] Adorno T. W., Ästetische Theorie, (book in German) R. Tiedemann, Gesammelte Schriften, Frankfurt. M., 1970. - 1973.

[22] Gadamer H. G., Wort und Bild – So Wahr, So Sein, (book in German), Gesammelte Werke, Tübingen, 1993.

[23] Gadamer H. G., The Legacy of Europe; MH, Edition: PARNAS, Press: Targa, Zagreb, 1997.

[24] Balaž Z., Meštrović K., Haun M., Cognitive Cybernetics - Propaedeutic of Polytechnic, the book manuscript on review, prepared issue of teaching material for Manualia Polytechnici Studiorum Zagrabiensis, Zagreb University of Applied Sciences, November 2017.

[25] Balaž Z., Predavec D., "Integration of Cognitive Cybernetics into Intelligent Human Systems?", scientific article prepared for the 1st International Conference on Intelligent Human System Integration, Dubai, 7th – 9th of January 2018.

[26] Balaž Z., "Implemented research study", course: Artificial Intelligence - Expert Systems, Electrical engineering department, Zagreb University of Applied Sciences, Specialist graduate study, 3rd semester, September - December 2013.; and course Intelligent Systems, 2nd semester, September 2015. – December 2016.

[27] Balaž, Z., "Intuigence and captology of Intelligent technologies – challenges and fears", (article in Croatian), sent on Polytechnic & Design journal, Zagreb University of Applied Sciences, October 2016.

[28] Parmy, O., "Jawbone guinea pig economy", FORBES, Vol XIV, No 3-4, UDK 167.7:316, pp. 63-68, III. 2015.

[29] Faucault, M., "The Courage of Truth", (article in Croatian) lectures from Collège de France, (1983.-1984.), ISBN 978-953-7715-61-8, translated from French: Zlatko Wurzberg, Standorf & Mizanotrop, Zagreb, 2015.

[30] Hoesle, V., "Filozofija ekološke krize", (article in Croatian) Moscow lectures, Publisher: Matica Hrvatska, Editor: Jelena Hekman, Press: Targa, ISBN 978-953-150-051-7, Zagreb, 1996.

[31] Grondin, J.: Sources of Hermeneutics, (book in Croatian) Publisher: Matica Hrvatska, Editor: Jelena Hekman, Press: Targa, ISBN 953-150-512-8, Zagreb, 1999.

[32] Jsselsteijn, W.I., Kort, Y., Midden, C., Eggen, B., Hoven, E., Persuasive Technology, ISSN 0302-9743, ISBN 3-540-34291-5, Springer - Verlag Berlin Heidelberg, LNCS 3962, 2006.

[33] Matulić, T., Metamorphosis of culture, Tertium mullenium, ISBN 9789-5324-1161-14, CIP 686343, Zagreb, 2009.

[34] Bubner, R., Izvornik, (book in Croatian) Aestetische Erfahrung, Suhrkamp Verlag, Frankfurt am Main, 1989. Translation: Estetsko iskustvo, Tihomir Engler, MH, Editor: Jelena Hekman, Biblioteka PARNAS, Press: Targa, ISBN 978-953-150-076-0, Zagreb, 1997.

[35] Carr, N., The Shallows: What the Internet Is Doing to Our Brains, (book in Croatian) Publisher: Jesenski i Turk, Translation: Ognjen Strpić, Zagreb, 2011.

[36] Kahneman, D., Thinking, Fast and Slow, (book in Croatian), Translation: Zvonko Pavić, ISBN 978-953-14-1483-8, CIP 857332, Publisher: Mozaik knjiga, Zagreb, 2013.

[37] Heisenberg, K., W., Physics and Philosophy, (book in Croatian), Translation: Stipe Kutleša, Kruzak, Press: S. Brusina – Donja Lomnica, ISBN 953-96477-3-8, Zagreb, 1997.

[38] Voltaire, Treatise on Tolerance, (book in Croatian), Translation: Bosiljka Brlečić, MH, Editor: Jelena Hekman, Biblioteka PARNAS, Press: Targa, ISBN 953-150-082-7, Zagreb, 1997.

[39] Rohr, R., The Naked Now: Learning to See as the Mystics See, (book in Croatian), Edition: Synopsis, Zagreb – Sarajevo, 2016.

[40] Petanjek, Z., "Sinaptički izazovi u djetinjstvu i adolescenciji: Izazovi i smjerovi suvremenog istraživanja mozga", (lecture in Croatian), HAZU: Symposium on Research and Brain Diseases, Zagreb, 16th March, 2016. http://www.hiim.unizg.hr/index.php/24-naslovnica/334-simpozij-hazu-2016-baw

[41] Jakovljević, M., "Mozak, duša i kultura u zdravlju i bolesti: Izazovi pred psihijatrijom 21. stoljeća", HAZU: Symposium on Research and Brain Diseases, Zagreb, 16th March, 2016. http://www.hiim.unizg.hr/index.php/24-naslovnica/334-simpozij-hazu-2016-baw

[42] Goleman, D., Emotional Intelligence, (book in Croatian), Translation: Damir Biličić, Editor: Ivanka Borovac,, Publisher: Mozaik knjiga, 2nd edition, available online, Zagreb 1997.

[43] Hood M. B., SuperSense: Why We Believe in the Unbelievable, (book in Croatian), Translation: Marko Maras, ISBN 978-953-303-175-0, Publisher: Ljevak, 2010.

[44] Paić, Ž., Posthuman Condition: The End of Human and Odds of Other History, (book in Croatian), Publisher: Litteris, ISBN 978-953-7250-41-6, Zagreb, October 2011.

[45] Paić, Ž., The Spheres of Existence: Three Studies on Kierkegaard, (book in Croatian), Publisher: Matica Hrvatska, ISBN 978-953-341-081-4, Zagreb, April 2017.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 119-129 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

Control of a three-stage medium voltage solid-state transformer

Claudio Busada[1], Hector Chiacchiarini[1], Sebastian Gomez Jorge[*1], Favio Mengatto[1], Alejandro Oliva[1], Jorge Solsona[1], German Bloch[2] and Angelica Delgadillo[2]

[1]*Instituto de Investigaciones en Ingeniería Eléctrica (IIIE), Universidad Nacional del Sur (UNS)-CONICET and Dpto. Ing. Eléctrica y de Computadoras, UNS, Av. Alem 1253, (8000) Bahía Blanca, Argentina.*

[2]*ICSA S.A. - Argentina*

ARTICLEINFO

Article history:
Received: 26 October, 2017
Accepted: 17 November, 2017
Online: 07 December, 2017

Keywords :
Solid State Transformer
Control
Power Electronics

ABSTRACT

This paper proposes the modeling and control of a Solid-State Transformer using a three-stage conversion topology. First, a rectification stage is used, where a three-phase high-voltage AC signal is converted to a DC level; this stage is then followed by a DC-DC converter, and finally an inverter is used to convert the DC into a three-phase low-voltage AC signal. The adopted topology is modeled using a simplified model for each stage, useful to design their controllers. Based on these models, the controllers are tuned to obtain a good performance to sudden load changes. This performance is tested through simulations.

1 Introduction

It is necessary to mention that this paper is an extension of work originally presented in [1]. This extension includes a detailed description of the proposed models and controllers. These models are used for tuning the controller parameters in order to obtain good performance in presence of sudden load changes.

Solid-State Transformers (SST) are emerging as a new technology capable of replacing power distribution transformers [2]. It is expected that the next generation of the power distribution transformers is based on power electronics semiconductor devices [3]. By using these semiconductors, it is possible to design an apparatus based on power converters with smaller and lighter high-frequency transformers. In this way, it is possible to obtain a smaller and lighter distribution transformer when compared with a traditional transformer of the same power rating. Roughly speaking, SSTs work as follows. In a first stage a sinusoidal signal of, typically, 50 or 60 Hz is converted to a high-frequency signal. Then, the amplitude of this signal is changed, by using a small high-frequency transformer, and finally in a third stage a new signal of the same frequency as the input signal but different amplitude is obtained.

Several topologies for the SST can be found in the literature, with their advantages and disadvantages [4]. In addition, it is possible to find some reviews dealing with the subject [5, 6]. Applications in transportation and smart grid can be found in [7]; whereas a topology based on SiC devices can be found in [8]. Also, a traditional transformer and an SST were compared in [9], and a procedure to obtain a detailed model of an SST topology can be found in [10].

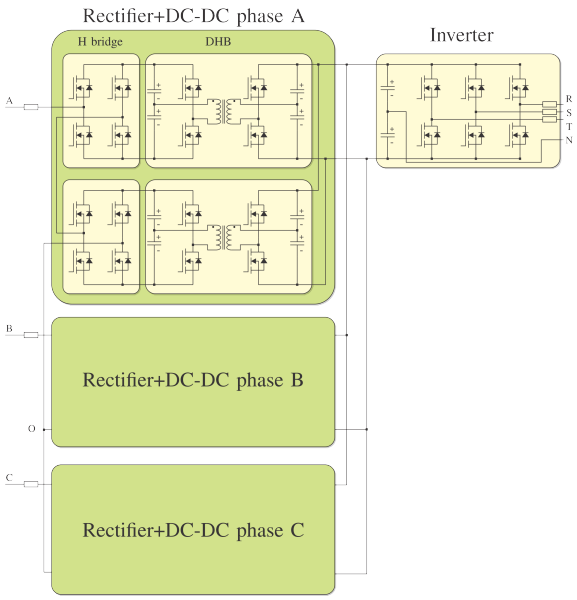


Figure 1: SST topology.

[*]Corresponding author, sebastian.gomezjorge@uns.edu.ar

A three-stage SST is considered in this paper [7]. Its topology is shown in Figure 1. The three stages are: a rectification stage, where a three-phase high voltage (HV) AC grid voltage is converted to a DC signal, a DC-DC converter stage (multiple stages in parallel) that provides isolation and performs the level adaptation, and finally, an inverter stage that converts the DC signal into a three-phase low voltage (LV) AC sinusoidal wave.

In this paper, simple models for each stage are proposed. Using these models, the required controllers are designed and their parameters are tuned to obtain a good performance in presence of sudden load changes. In order to test the performance of the proposed controllers, simulations results are introduced.

2 Proposed SST topology and model

Figure 2 shows a general diagram of the chosen topology. The SST is composed of a bidirectional multi-level HV three-phase converter, labeled Rectifier, six bidirectional isolated DC-DC converters, labeled DHB1 through DHB6, and a bidirectional LV three-phase converter with neutral, labeled Inverter. This topology has a great degree of modularity, allowing to consider the converters as decoupled systems and simplifying the design of the controllers [7].

Each converter and its simplified model is described below. Since the controllers are implemented using digital control techniques, the models are described in discrete time, assuming a sampling time T_s.

2.1 Rectifier

The rectifier is a three-phase AC-DC converter connected to the HV grid. As shown in Figure 2, the rectifier input port is modeled as a three-phase controlled voltage source $v_{rec} = [v_{reca}\ v_{recb}\ v_{recc}]^T$, coupled to the HV grid voltage $v_{hv} = [v_{hva}\ v_{hvb}\ v_{hvc}]^T$ through an inductor with current $i_{hv} = [i_{hva}\ i_{hvb}\ i_{hvc}]^T$. This coupling filter is modeled using complex vector notation (see Appendix). The zero order hold (zoh) discrete time model of the coupling filter results:

$$\vec{i}_{hv}[k+1] = \frac{T_s}{L_{rec}}\left(\bar{\vec{v}}_{hv}[k] - \vec{v}_{rec}[k]\right) + \vec{i}_{hv}[k], \qquad (1)$$

where $\bar{\vec{v}}_{hv}[k] = 0.5(\vec{v}_{hv}[k+1] + \vec{v}_{hv}[k])$ is the mean value of the HV grid voltage in the time interval $T_s k$, $T_s(k+1)$.

The output ports of the rectifier are modeled as current sources i_{iX}, with $X = 1 \ldots 6$. These currents are computed through power balance. Considering that both H bridges of each phase of the rectifier are given the same reference, it can be found that their currents are equal:

$$i_{i1}[k] = i_{i2}[k] = \frac{v_{reca}[k] i_{hva}[k]}{V_{busH1}[k] + V_{busH2}[k]}, \qquad (2)$$

$$i_{i3}[k] = i_{i4}[k] = \frac{v_{recb}[k] i_{hvb}[k]}{V_{busH3}[k] + V_{busH4}[k]}, \qquad (3)$$

$$i_{i5}[k] = i_{i6}[k] = \frac{v_{recc}[k] i_{hvc}[k]}{V_{busH5}[k] + V_{busH6}[k]}. \qquad (4)$$

Note that in this rectifier model, the independent variables (control inputs) are the three-phase components of $v_{rec}[k]$.

2.2 Dual Half Bridge (DHB)

Since there are six isolated DC buses in the rectifier, there are six DC-DC converters which are implemented using the DHB topology, which are modeled as current sources. The coupling between the output ports of the rectifier and the input port of each DHB is performed through a capacitor. The voltage across each of these capacitors is modeled through the difference equation:

$$V_{busHX}[k+1] = \frac{T_s}{C_H/2}\left(i_{iX}[k] - i_{oX}[k]\right) + V_{busHX}[k], \quad (5)$$

with $X = 1 \ldots 6$. The DHBs themselves are controlled using a phase shift strategy. Therefore, each converter can be modeled through the algebraic equation for their average power transfer [11]:

$$P_{dhbX}[k] = \frac{V_{busHX}[k]\ mV_{busL}[k]\delta_X[k]\left(\pi - |\delta_X[k]|\right)}{8\pi^2 L_d f_{sw}^{dhb}}, \quad (6)$$

where m is the DHB transformer relation, L_d is the total leakage of the transformer referred to the HV side, f_{sw}^{dhb} is the switching frequency of the DHB converter, and δ_X is the phase shift angle between the voltages generated at the HV and LV sides of the DHB under consideration. Dividing this power by $V_{busHX}[k]$, it results

$$i_{oX}[k] = \frac{mV_{busL}[k]\delta_X[k]\left(\pi - |\delta_X[k]|\right)}{8\pi^2 L_d f_{sw}^{dhb}}, \qquad (7)$$

which describes the relation between input port i_{oX} and the control action δ_X. The outputs of all the DHBs are connected in parallel, and feed the LV bus. The relation between the input signals and the output signals is obtained through power balance, as described in the following equation:

$$i_{dhbX}[k] = \frac{i_{oX}[k] V_{busHX}[k]}{V_{busL}[k]}. \qquad (8)$$

Note that in this DHB model, the independent variable (control input) for each DHB is δ_X.

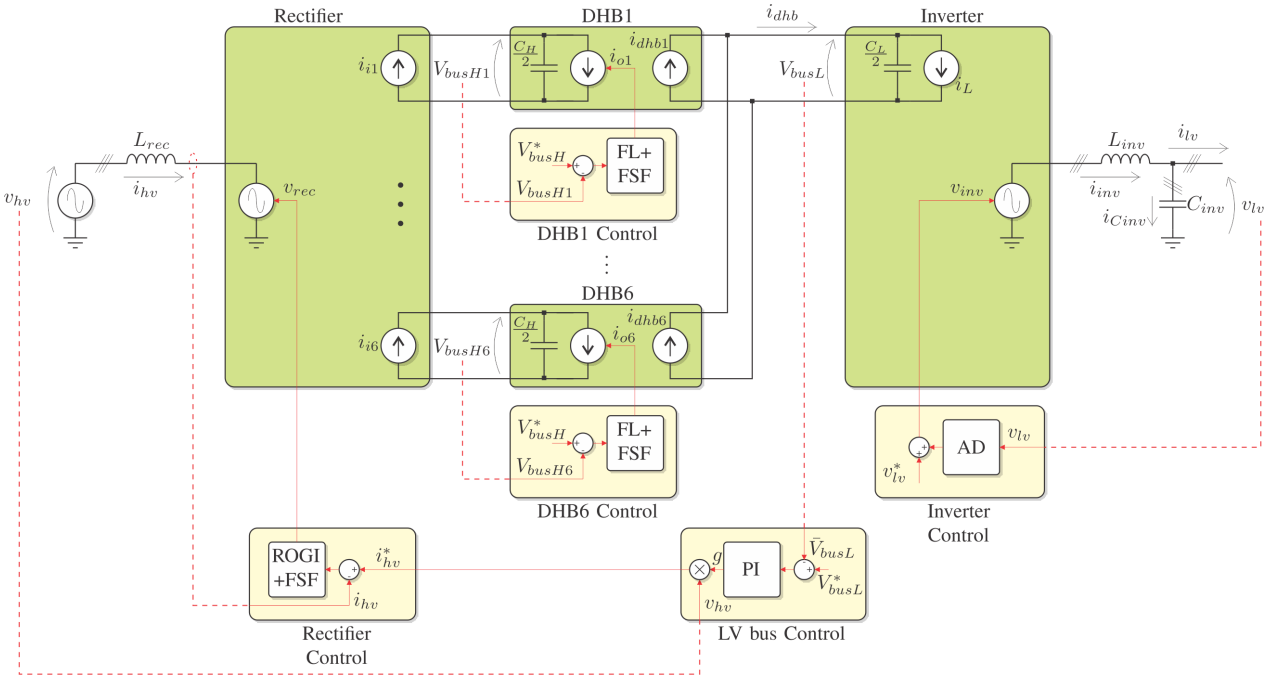


Figure 2: SST topology, model and general control scheme.

2.3 Inverter

The inverter is a three-phase DC-AC converter connected to the LV grid. The input port is modeled as a controlled current source i_L, and the output port is modeled as a three-phase controlled voltage source $v_{inv} = [v_{invr}\ v_{invs}\ v_{invt}]^T$. The load voltage and current are $v_{lv} = [v_{lvr}\ v_{lvs}\ v_{lvt}]^T$ and $i_{lv} = [i_{lvr}\ i_{lvs}\ i_{lvt}]^T$, respectively.

The coupling between the output ports of the DHBs and the input port of the inverter is performed through a capacitor. The voltage across this capacitor is modeled through the difference equation:

$$V_{busL}[k+1] = \frac{T_s}{C_L/2}\Big(i_{dhb}[k] - i_L[k]\Big) + V_{busL}[k], \qquad (9)$$

where

$$i_{dhb}[k] = \sum_{X=1}^{6} i_{dhbX}[k]. \qquad (10)$$

The coupling between the load and the output port of the inverter is performed by an LC filter. Since each leg of the inverter is controlled as a single phase converter, the zoh discretization of this filter is given for each phase of the inverter. This is denoted with subscript Y, which can be equal to r, s or t:

$$x_{invY}[k+1] = A_{inv}x_{invY}[k] + B_{inv}v_{invY}[k] + B_{inv1}i_{lvY}[k], \qquad (11)$$

where $x_{invY}[k] = [i_{invY}[k]\ v_{lvY}[k]]^T$,

$$A_{inv} = \begin{bmatrix} \cos\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) & -\sqrt{\frac{C_{inv}}{L_{inv}}}\sin\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \\ \sqrt{\frac{L_{inv}}{C_{inv}}}\sin\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) & \cos\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \end{bmatrix}, \qquad (12)$$

$$B_{inv} = \begin{bmatrix} \sqrt{\frac{C_{inv}}{L_{inv}}}\sin\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \\ 1 - \cos\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \end{bmatrix}, \qquad (13)$$

$$B_{inv1} = \begin{bmatrix} 1 - \cos\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \\ -\sqrt{\frac{L_{inv}}{C_{inv}}}\sin\left(\frac{T_s}{\sqrt{C_{inv}L_{inv}}}\right) \end{bmatrix}. \qquad (14)$$

As in the previous converters, the relation between the input signals and the output signals is obtained through power balance, as described in the following equation:

$$i_L[k] = \frac{v_{inv}[k] \bullet i_{lv}[k]}{V_{busL}[k]}, \qquad (15)$$

where the $\bullet$ operator denotes scalar product. Note that in this inverter model, the independent variables (control inputs) are the three-phase components of $v_{inv}[k]$.

3 Converter controller description

Figure 2 shows the simplified block diagrams of the proposed controllers for each converter. In this figure, dashed lines represent measured signals, and solid lines in the controllers represent control signals. In what follows, each individual controller is described.

3.1 Rectifier control

This controller is tasked to make the HV grid current i_{hv} to copy the current reference i^*_{hv}. This is

achieved through a full state feedback (FSF) control with the addition of a reduced order generalized integrator (ROGI) [12]. The ROGI is added to achieve zero steady state tracking error, since the current reference is a grid frequency positive sequence three-phase signal. Considering a one sample time processing delay, typical in digital implementations, complex vector notation, and space vector modulation (SVM), the control action is

$$v_{rec}[k] = v^*_{recSVM}[k-1], \tag{16}$$

$$v^*_{recSVM}[k] = v^*_{rec}[k] - v_{mc}[k], \tag{17}$$

$$v_{mc}[k] = 0.5\,(\max(v^*_{rec}[k]) + \min(v^*_{rec}[k])), \tag{18}$$

where max and min functions are computed between the *abc* components of $v^*_{rec}[k]$. Also,

$$\vec{v}^*_{rec}[k] = -K_{rec}\begin{bmatrix} \vec{i}_{hv}[k] - \vec{i}^*_{hv}[k] \\ \vec{v}^*_{rec}[k-1] \\ \vec{r}_1[k] \end{bmatrix}, \tag{19}$$

$$\vec{r}_1[k+1] = j(1-e^{j\omega T_s})\left(\vec{i}_{hv}[k] - \vec{i}^*_{hv}[k]\right) + e^{j\omega T_s}\vec{r}_1[k], \tag{20}$$

with K_{rec} a 1×3 complex gain vector, and $\vec{r}_1[k]$ represents the state of a ROGI tuned at fundamental grid angular frequency ω. As described in [12], gain vector K_{rec} can be found using the linear quadratic regulator method, or by pole placement using Ackerman's formula, and the matrix description of the system:

$$x_{rec}[k+1] = A_{rec}x_{rec}[k] + B_{rec}\vec{v}^*_{rec}[k], \tag{21}$$

where $x_{rec}[k] = [\vec{i}_{hv}[k]\ \vec{v}^*_{rec}[k-1]\ \vec{r}_1[k]]^T$,

$$A_{rec} = \begin{bmatrix} 1 & -\frac{T_s}{L_{rec}} & 0 \\ 0 & 0 & 0 \\ j(1-e^{j\omega T_s}) & 0 & e^{j\omega T_s} \end{bmatrix}, \tag{22}$$

$$B_{rec} = [0\ 1\ 0]^T. \tag{23}$$

The reference i^*_{hv} is the output of the LV bus controller, and it is related to the instantaneous HV grid voltage through

$$i^*_{hv}[k] = g[k]v_{hv}[k], \tag{24}$$

where g is a scalar variable signal. Therefore, depending on the sign of g and assuming v_{hv} sinusoidal with no harmonic distortion, the rectifier will source or sink a three-phase sinusoidal current to the HV grid, with unity power factor. In this paper the closed loop poles are chosen to achieve a settling time of 4.5[ms].

3.2 DHB control

The objective of the DHBs is to transfer the pulsating power of each of the H bridges of the rectifier to the LV bus, where the resulting power is non-pulsating. To achieve this objective, the controller of each DHB is designed to keep its instantaneous voltage V_{busHX} at its reference level V^*_{busH} (which is the same for all six DHBs). As shown by (7), the relation between current i_{oX} and phase-shift angle δ_X is non-linear. In order to be able to apply linear control techniques, a feedback linearization (FL) is implemented. Then, the resulting linear system is controlled through FSF.

Considering a one sample time processing delay, the control action is

$$\delta_X[k] = \delta^*_X[k-1], \tag{25}$$

where

$$\delta^*_X[k] = 0.5\pi\left(1 - \sqrt{1 - 32 f^{dhb}_{sw}\frac{|i^*_{oX}[k]|}{mV_{busL}[k]}}\right)sign(i^*_{oX}[k]), \tag{26}$$

is obtained from (7) and it is the FL equation for the system. Assuming that $V_{busL}[k] \simeq V_{busL}[k-1]$ (slow varying signal), it can be numerically verified that for a given value of $i^*_{oX}[k-1]$, evaluating (26) at $k-1$, and replacing the result in (25) and (7) yields $i_{oX}[k] \simeq i^*_{oX}[k-1]$. Therefore, the linearized system is modeled by (5) and

$$i_{oX}[k] = i^*_{oX}[k-1]. \tag{27}$$

Adding a discrete time backward Euler integrator to achieve zero steady state error, the control action for each DHB is computed as follows:

$$i^*_{oX}[k] = -K_{DHB}\begin{bmatrix} V_{busHX}[k] - V^*_{busH}[k] \\ r_0[k] \\ i_{oX}[k] \end{bmatrix}, \tag{28}$$

$$r_0[k+1] = T_s\left(V_{busHX}[k] - V^*_{busH}[k]\right) + r_0[k], \tag{29}$$

where K_{dhb} is a 1×3 gain vector, and $r_0[k]$ represents the state of the discrete integrator. Once (28) is computed, the phase shift angle δ^*_X is obtained through (26) and applied to the converter.

Gain vector K_{dhb} can be found using the linear quadratic regulator method, or by pole placement using Ackerman's formula, and the matrix description of the linearized system:

$$x_{dhb}[k+1] = A_{dhb}x_{dhb}[k] + B_{dhb}i^*_{oX}[k], \tag{30}$$

where $x_{dhb}[k] = [V_{busHX}[k]\ r_0[k]\ i_{oX}[k]]^T$,

$$A_{dhb} = \begin{bmatrix} 1 & 0 & -\frac{T_s}{C_H/2} \\ T_s & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{31}$$

$$B_{dhb} = [0\ 0\ 1]^T. \tag{32}$$

This scheme makes each DHB source power to the LV bus if its HV bus voltage is higher than V^*_{busH}, or sink power from the LV bus if its HV bus voltage is lower than V^*_{busH}. In this paper the closed loop poles are chosen to achieve a settling time of 1[ms].

3.3 LV bus control

This controller is tasked to keep the mean value of the LV bus voltage ($\bar{V}_{busL}$ in Figure 2) at the reference level V^*_{busL}. In order to do so, signal V_{busL} is filtered (filter not shown in the figure) and then a proportional integral (PI) controller is used. The output of this controller is the variable gain g. This gain is used to properly scale the measured HV grid voltage, and generate the current reference for the rectifier, defined in (24). For this reason, the dynamics of the rectifier control loop and the LV bus control loop must be decoupled. This is achieved by making the bus control loop significantly slower than the rectifier and DHB control loops.

The LV bus voltage is modeled by (9). Since there is no controlled current source directly connected to the LV bus, the design of the controller assumes that i_{dhb} is the control action. Once the control action i_{dhb} is computed, gain g is obtained through power balance:

$$g[k] = \frac{i_{dhb}[k]\bar{V}_{busL}[k]}{3V^2_{nomhv}}, \tag{33}$$

where V_{nomhv} is the nominal rms value of the HV grid voltage. As stated at the beginning of this section, this controller is slow. Therefore, both the processing delay and the dynamics of the filter used to obtain $\bar{V}_{busL}$ can be ignored without significant error. Considering this, and adding a discrete time backward Euler integrator to achieve zero steady state error, the control action is computed as follows:

$$i_{dhb}[k] = -K_{LV}\begin{bmatrix}\bar{V}_{busL}[k] - V^*_{busL}[k] \\ r_{0LV}[k]\end{bmatrix}, \tag{34}$$

$$r_{0LV}[k+1] = Ts\Big(\bar{V}_{busL}[k] - V^*_{busL}[k]\Big) + r_{0LV}[k], \tag{35}$$

where K_{LV} is a 1×2 gain vector, and $r_{0LV}[k]$ represents the state of the discrete integrator. Once (34) is computed, gain g is computed through (33), and used to generate the HV grid current reference i^*_{hv} through (24).

Gain vector K_{LV} can be found using the linear quadratic regulator method, or by pole placement using Ackerman's formula, and the matrix description of the system:

$$x_{LV}[k+1] = A_{LV}x_{LV}[k] + B_{LV}i_{dhb}[k], \tag{36}$$

where $x_{LV}[k] = [V_{busL}[k]\ r_{0LV}[k]]^T$,

$$A_{LV} = \begin{bmatrix}1 & 0 \\ T_s & 1\end{bmatrix}, \tag{37}$$

$$B_{LV} = [\frac{T_s}{C_L/2}\ 0]^T. \tag{38}$$

In this paper the closed loop poles are chosen to achieve a settling time of 100[ms].

3.4 Inverter control

This controller implements the active damping (AD) of the output LC filter resonance. The implementation of the AD requires to measure the current through C_{inv}. To avoid this measurement, an estimator through high pass filter derivation is used. This estimator only requires the measurement of v_{lv}, and only slightly increases the settling time of the control loop. The LV grid voltage reference v^*_{lv} is added to the control action of the AD. This voltage reference is a three-phase balanced sinusoidal signal, which can be generated internally, or can be obtained through a synchronization algorithm in synchronism with the HV grid voltage.

Since each leg of the inverter is controlled independently, there are three equal controllers. In what follows, they are described with subscript $Y = r$, s or t. Considering a one sample processing time, the control action for the active damping strategy is $v_{invY}[k] = v^*_{invY}[k-1]$, where

$$v^*_{invY}[k] = -K_{lv}\begin{bmatrix}\hat{i}_{CinvY}[k] \\ v_{lvY}[k] \\ v^*_{invY}[k-1]\end{bmatrix} + K^*v^*_{lvY}, \tag{39}$$

where K_{lv} is a 1×3 gain vector, K^* is a gain, and $\hat{i}_{CinvY}[k]$ is the estimated capacitor current, obtained from the zoh discretization of a high pass filter:

$$\eta_Y[k+1] = (e^{-\omega_c T_s} - 1)v_{lvY}[k] + e^{-\omega_c T_s}\eta_Y[k], \tag{40}$$

$$\hat{i}_{CinvY}[k] = C_{inv}\omega_c\Big(v_{lvY}[k] + \eta_Y[k]\Big), \tag{41}$$

with ω_c the cut off frequency of the high pass filter, and η_Y the state of the filter.

Gain vector K_{lv} is obtained by pole placement to damp the LC filter resonance using Ackerman's formula, and the matrix description of the system:

$$x_{lv}[k+1] = A_{lv}x_{lv}[k] + B_{lv}v^*_{invY}[k], \tag{42}$$

where $x_{lv}[k] = [i_{invY}[k]\ v^*_{invY}[k-1]\ v_{invY}[k]]^T$,

$$A_{lv} = \begin{bmatrix}A_{inv} & B_{inv} \\ 0\ 0 & 0\end{bmatrix}, \tag{43}$$

$$B_{lv} = [0\ 0\ 1]^T. \tag{44}$$

Finally, gain K^* is included to compensate the magnitude of the LV grid voltage reference, and is obtained evaluating:

$$TF = C_{lv}(z\,I - A^{CL}_{lv})^{-1}B_{lv}\Big|_{z=e^{j\omega T_s}}, \tag{45}$$

$$K^* = |1/TF|, \tag{46}$$

where $A^{CL}_{lv} = A_{lv} - B_{lv}K_{lv}$, I is the 3×3 identity matrix and $C_{lv} = [0\ 1\ 0]$. Note that the controllers for each phase use the same gain vector K_{lv} and gain K^*. In this paper the closed loop poles are chosen for optimal damping ($\zeta = 0.707$). This results in a settling time of approximately 2[ms].

4 Parameter selection criteria

This section gives criteria for choosing the values of the different parameters of each converter.

4.1 Rectifier coupling inductance L_{rec}

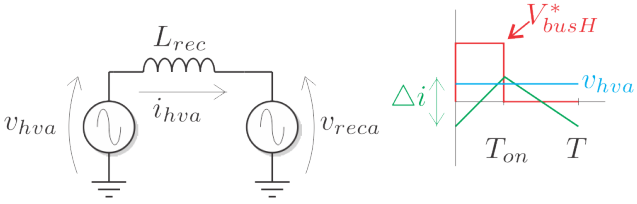


Figure 3: Rectifier phase a and switching interval variables.

The value of L_{rec} is chosen to obtain a desired current ripple $\triangle i$ when injecting zero current to the HV grid. Due to the 5-level structure of the rectifier and the use of unipolar modulation, the effective switching frequency applied to L_{rec} is $4f^{rec}_{sw}$ [1]. Also, this multilevel structure ensures that the voltage difference applied to L_{rec} will never be larger than V^*_{busH} (assuming all the HV buses are kept at that level).

Taking phase a as a reference, Figure 3 shows one switching interval. The current variation in this interval is

$$\triangle i = \frac{T_{on}}{L_{rec}}(v_{hva} - V^*_{busH}). \tag{47}$$

Since zero current injection is assumed, in this interval

$$v_{hva} \simeq V^*_{busH} d, \tag{48}$$

where $d = T_{on}/T$. Therefore, replacing this in (47),

$$\triangle i = \frac{V^*_{busH}}{L_{rec} 4f^{rec}_{sw}}(d^2 - d), \tag{49}$$

where $T = 1/(4f^{rec}_{sw})$ was used. The maximum $\triangle i$ occurs for $d = 0.5$, therefore for this condition, from (49),

$$L_{rec} = \frac{V^*_{busH}}{\triangle i 16 f^{rec}_{sw}}. \tag{50}$$

From Table 1, choosing the peak current ripple $\triangle i/2 = 0.1(\sqrt{2}I_{nomhv})$, the inductance results $L_{rec} = 190$[mH]. From an additional analysis not included in this paper, considering commercially available cores, $L_{rec} = 200$[mH] is chosen.

4.2 DHB transformer leakage inductance L_d

From (6), assuming $V_{busHX}[k] = V^*_{busH}$ and $V_{busL}[k] = V^*_{busL} = V^*_{busH}/m$ the maximum power transfer occurs for $\delta_X[k] = \pi/2$ and results

$$P_{dhbMAX} = \frac{(V^*_{busH})^2}{32 L_d f^{dhb}_{sw}}. \tag{51}$$

Under nominal operation conditions, each DHB will have to transfer a mean power $\bar{P}_{dhb} = S_{nom}/6$. Taking a safety margin of $P_{dhbMAX} = 2\bar{P}_{dhb}$ to catch any transients, and replacing this in (51), the leakage inductance results

$$L_d = \frac{3(V^*_{busH})^2}{32 S_{nom} f^{dhb}_{sw}}. \tag{52}$$

Using the necessary parameters from Table 1 to evaluate (52), it results $L_d = 8.44$[mH]. From an additional analysis not included in this paper, designing L_d considering commercially available cores, it results $L_d = 8.8$[mH] .

4.3 HV bus capacitor C_H

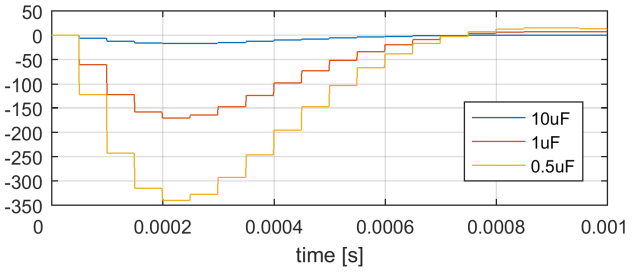


Figure 4: $\triangle V_{busH}$ [V] for a 10% HV grid voltage dip for different values of C_H.

During the event of dips or swells in the HV grid voltage v_{hv}, there will be a transient behavior in each of the HV buses. This can lead to unacceptable under/over voltages in V_{busHX}. The HV buses voltage variations during these transients are determined by the value of C_H, the settling time of each DHB control loop, and the magnitude of the dips or swells. Since the evaluation of these transients involves the DHB control loop, the process is iterative. For a given value of C_H, gain K_{dhb} must be computed. Then, the performance is evaluated through a simple simulation applying the desired dip or swell magnitude. Finally, the process is repeated until the desired performance is attained.

The evaluation is performed considering a step change in the magnitude of v_{hv}. In the following, the worst case scenario for V_{busH1} is analyzed, which is similar for the remaining buses.

The maximum instantaneous power variation in phase a occurs when the rectifier is sinking nominal current, and at the instantaneous peak of $v_{hva}[k]$ its magnitude changes $\triangle v_{hv}$ volts. From (1), assuming that previous to the step $\vec{v}_{hv} \simeq \vec{v}_{rec}$, the magnitude of the change in the HV grid current is

$$\triangle i_{hv} \simeq \frac{T_s}{L_{rec}} \triangle v_{hv}. \tag{53}$$

From (2), assuming that previous to the step $V_{busH1} = V_{busH1} = V^*_{busH}$, the variation in current i_{i1} is

$$\triangle i_i = \frac{\sqrt{2} V_{nomhv} \triangle i_{hv}}{2 V^*_{busH}} = \frac{T_s}{L_{rec}} \frac{\sqrt{2} V_{nomhv} \triangle v_{hv}}{2 V^*_{busH}}, \tag{54}$$

where (53) was used for the last equality. Now the effect of $\triangle v_{hv}$ in V_{busH1}, defined as $\triangle V_{busH}$, is evaluated simulating the closed loop response of the linearized DHB (30)-(32):

$$\triangle x_{dhb}[k+1] = A^{cl}_{dhb}\triangle x_{dhb}[k] + B^{i_i}_{dhb}\triangle i_i, \quad (55)$$

$$\triangle V_{busH}[k] = [1\ 0\ 0]\triangle x_{dhb}[k], \quad (56)$$

where $B^{i_i}_{dhb} = [\frac{T_s}{C_H/2}\ 0\ 0]^T$ and $A^{cl}_{dhb} = A_{dhb} - B_{dhb}K_{dhb}$. Figure 4 shows simulation results for this system for $C_H = 10$ [μF], 1[μF] and 0.5[μF] when a 10% dip is simulated ($\triangle v_{hv} = -0.1V_{nomhv}$) for the DHB control designed with a settling time of 1[ms]. From the results $C_H = 1$ [μF] is chosen, since it results in a 2.5% voltage variation.

4.4 LV bus capacitor C_L

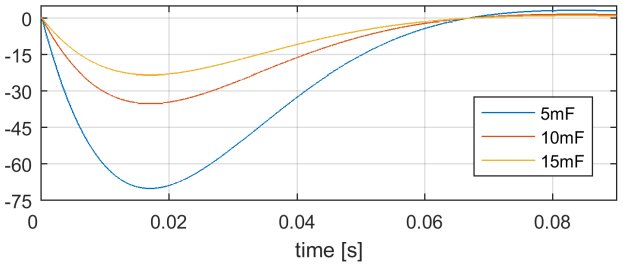


Figure 5: $\triangle V_{busL}$ [V] for a nominal load sudden connection and different values of C_L.

The value of capacitor C_L is chosen so that in the event of a sudden nominal load connection, the LV bus voltage does not go below the minimum value required for normal operation. The minimum LV bus voltage required for operation is

$$V_{busLMIN} = 2\sqrt{2}V_{nomlv} = 622V, \quad (57)$$

where the value is obtained from Table 1. Therefore, a conservative design criterion is to keep $V_{busL} > 700V$ when a nominal load is connected. This is equivalent to obtain a LV bus voltage variation $\triangle V_{busL} < 100V$.

To evaluate the LV bus voltage variation for different values of C_L, gain vector K_{LV} is computed for each given value, and then a simulation is performed. From (36)-(38), the following system is simulated:

$$\triangle x_{LV}[k+1] = A^{cl}_{LV}\triangle x_{LV}[k] + B^{cl}_{LV}\triangle i_L, \quad (58)$$

$$\triangle V_{busL}[k] = [1\ 0]\triangle x_{LV}[k], \quad (59)$$

where $A^{CL}_{LV} = A_{LV} - B_{LV}K_{LV}$, $B^{cl}_{LV} = -B_{LV}$ and $\triangle i_L = S_{nom}/V^*_{busL}$. Figure 5 shows simulation results for this system for $C_L = 5$ [mF], 10[mF] and 15[mF] when a nominal load is suddenly connected and the LV bus voltage control loop is designed with a settling time of 100[ms]. From the results all capacitor values meet the requirements, however $C_L = 10$ [mF] is chosen because of reduced ripple under unbalanced load conditions.

4.5 Inverter filter $L_{inv}C_{inv}$

The typical output impedance of a standard transformer is 2%-5% of its base impedance. Therefore, L_{inv} is chosen so that its impedance at angular frequency ω is 2% of the base impedance:

$$L_{inv} = 0.02\frac{3V^2_{nomlv}}{\omega S_{nom}} = 462\ [\mu H]. \quad (60)$$

On the other hand, capacitor C_{inv} limits the bandwidth of the output. Here, to obtain a fast transient response, the cut-off angular frequency of the output filter is chosen

$$\omega_{cut-off} = 20\omega. \quad (61)$$

Since the resonance frequency of the filter is approximately equal to the cut-off frequency,

$$\omega_{res} \simeq \omega_{cut-off} = \frac{1}{\sqrt{L_{inv}C_{inv}}}, \quad (62)$$

then

$$C_{inv} = \frac{1}{(20\omega)^2 L_{inv}} = 55\ [\mu F]. \quad (63)$$

5 Operation analysis of the control system and simulation results

This section presents simulation results of the proposed SST when a load is suddenly connected and disconnected. Additionally, results for sudden nonlinear load connection are included. The results are obtained using the switching models of all converters. The system parameter summary is shown in Table 1.

Table 1: System parameter summary

RECTIFIER AND LV BUS		
Param.	Value	Description
S_{nom}	20[kVA]	SST nominal power
V_{nomhv}	7621[Vrms]	HV nom. phase voltage
I_{nomhv}	0.875[Arms]	HV nom. phase current
f	50[Hz]	HV grid frequency
L_{rec}	200[mH]	HV filter inductance
V^*_{busL}	800[V]	LV bus reference
C_L	10[mF]	LV bus capacitor
f^{rec}_{sw}	8[kHz]	Switching frequency
DHB		
C_H	1[μF]	HV bus capacitor
C^{dhb}_L	56[μF]	LV DHB capacitor
L_d	8.8[mH]	DHB leakage ind.
V^*_{busH}	6000[V]	HV bus reference
m	7.5	Transformer relation
f^{dhb}_{sw}	20[kHz]	Switching frequency
INVERTER		
L_{inv}	461.2[μH]	LC filter ind.
C_{inv}	55[μF]	LC filter cap.
V_{nomlv}	220[Vrms]	LV phase voltage
f^{inv}_{sw}	20[kHz]	Switching frequency

The controllers were designed so that the rectifier control loop has a settling time of 4.5[ms], the DHB control loop has a settling time of 1[ms], and the LV bus voltage control loop has a settling time of 100[ms]. The settling time of the inverter AD loop is defined by the cutoff frequency of its LC output filter plus the estimation of the current through C_{inv}. This settling time results approximately 2[ms].

5.1 Sudden load connection

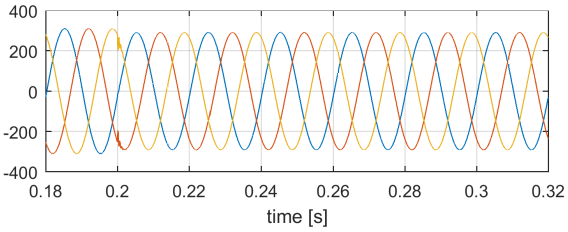


Figure 6: v_{lv} [V] sudden load connection.

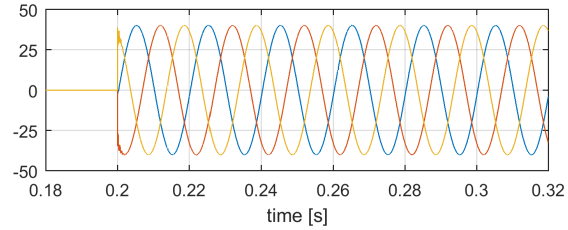


Figure 7: i_{lv} [A] sudden load connection.

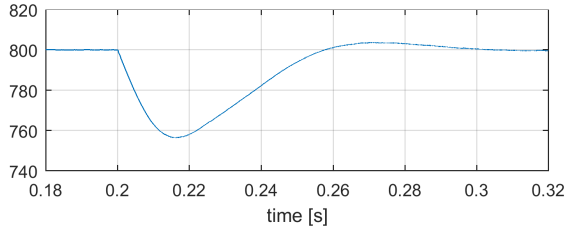


Figure 8: V_{busL} [V] sudden load connection.

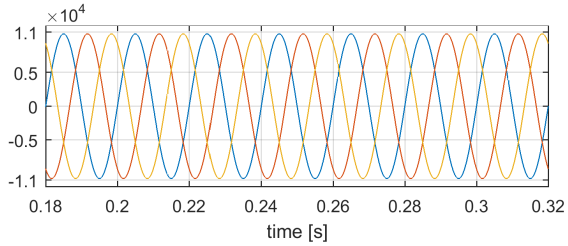


Figure 9: v_{hv} [V] sudden load connection.

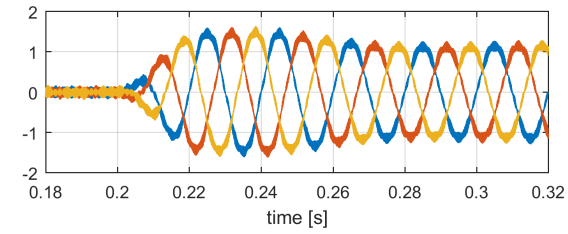


Figure 10: i_{hv} [A] sudden load connection.

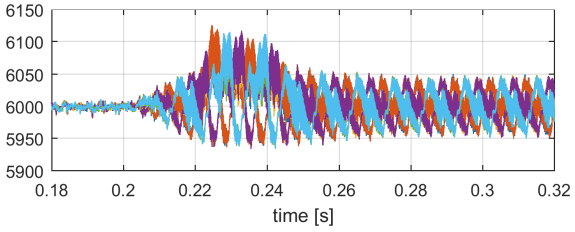


Figure 11: $V_{busH1} - V_{busH6}$ [V] sudden load connection.

The simulation results shown in Figures 6-11 start at $t = 0.18$[s] with the SST in steady state, with no load connected on the LV side. This means that the HV and LV buses start at their reference voltage levels and that a balanced three-phase sinusoidal voltage v_{lv} is generated without taking significant current i_{hv} from the HV grid.

At $t = 0.2$[s] a resistive nominal load is connected at the inverter output. The following sequence of events occurs (refer to Figure 2 for the definition of the variables):

- Figures 6 and 7 show v_{lv} and i_{lv}, respectively. As expected, when the load is connected v_{lv} has a short transient.
- Through power balance, current source i_L takes current from capacitor C_L, reducing voltage V_{busL} as shown in Figure 8. As can be seen, its settling time is approximately 100[ms], as designed.
- The LV bus voltage control loop detects the reduction of $\bar{V}_{busL}$, increasing in turn the value of g, resulting in $g > 0$. As a result, current reference i^*_{hv} magnitude increases.
- Commanded by its current controller, the rectifier sinks active power from the grid, with unity power factor. Figures 9 and 10 show v_{hv} and i_{hv}, respectively. Here the settling time of the current is tied to the settling time of g, defined by the LV bus control loop.
- Through power balance, the current source outputs of the rectifier charges capacitors C_H of the HV buses, increasing their voltage levels. The voltages of the HV buses is shown in Figure 11.
- The control loop of each DHB detects the increase in their respective HV bus voltage V_{busHX}. As a result, it commands each DHB to sink current i_{oX} in order to decrease the instantaneous value V_{busHX} to its reference value V^*_{busH} once again. This results in the short transient increase in the mean value of voltages V_{busHX} seen in Figure 11. Through power balance, each output current i_{dhbX} sources current to the LV bus, charging C_L to V^*_{busL} again, as shown in Figure 8.

By the end of this sequence, in steady state, the system is delivering power to the load and taking current with unity power factor from the HV grid.

5.2 Sudden load disconnection

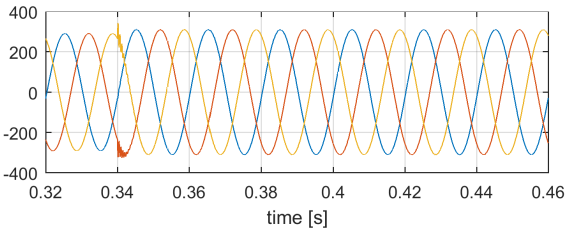


Figure 12: v_{lv} [V] sudden load disconnection.

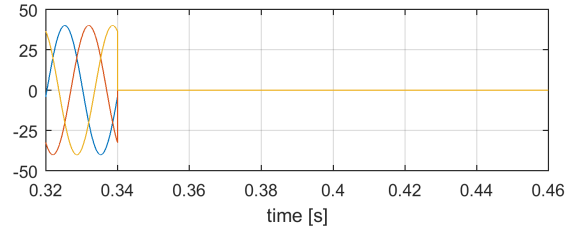


Figure 13: i_{lv} [A] sudden load disconnection.

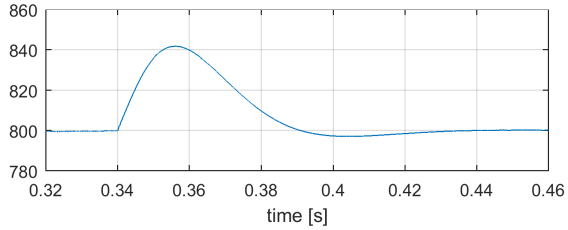


Figure 14: V_{busL} [V] sudden load disconnection.

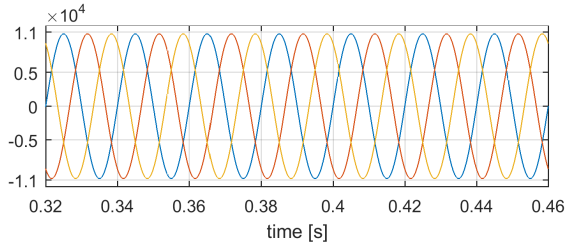


Figure 15: v_{hv} [V] sudden load disconnection.

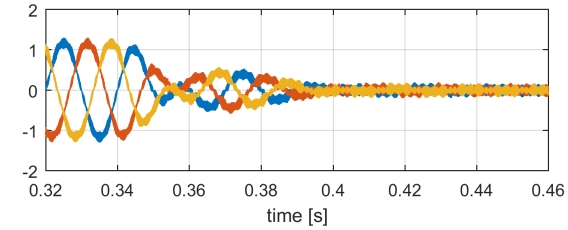


Figure 16: i_{hv} [A] sudden load disconnection.

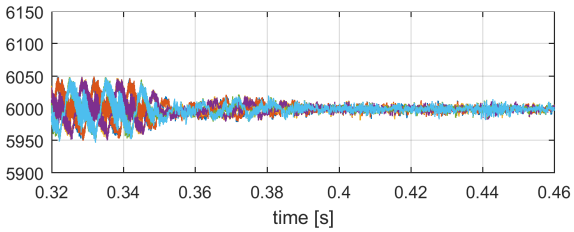


Figure 17: $V_{busH1} - V_{busH6}$ [V] sudden load disconnection.

Starting from the previous condition, if the load is now disconnected, the following sequence of events will occur:

- Figures 12 and 13 show v_{lv} and i_{lv}, respectively. As expected, when the load is disconnected v_{lv} has a short transient.
- Through power balance, current source i_L stops sinking current from capacitor C_L, which charges because the DHBs are still transferring power from the HV grid. This is seen in Figure 14.
- LV bus voltage control loop detects the voltage increase in $\bar{V}_{busL}$, decreasing in turn the value of g, resulting in $g < 0$. As a result, current reference i^*_{hv} magnitude decreases.
- Commanded by its current controller, the rectifier goes from sinking to supplying active power to the grid, with unity power factor. Figures 15 and 16 show v_{hv} and i_{hv}, respectively.
- Through power balance, the current source outputs of the rectifier discharge capacitors C_H of the HV buses, decreasing their voltage levels, as seen in Figure 17.
- The control loop of each DHB detects the decrease in their respective HV bus voltage V_{busHX}. As a result, it commands each DHB to source current i_{oX} in order to increase the instantaneous value V_{busHX} to its reference value V^*_{busH} once again. This results in the short transient decrease in the mean value of voltages V_{busHX} seen in Figure 17. Through power balance, each output current i_{dhbX} sinks current from the LV bus, discharging C_L to V^*_{busL} again, as shown in Figure 14.

By the end of this sequence, in steady state, the system has its buses at their reference values, and the inverter generates v_{lv} without sinking current i_{hv} from the HV grid.

5.3 Sudden non-linear load connection

The simulation of section 5.1 is repeated for the sudden connection of a balanced non-linear load. For each phase of the inverter, the load is composed of a single phase rectifier, feeding a 1[mH] inductor in series an RC load where R and C are in parallel (with R=19.5[Ω] and C=1[μF]). The results of this simulation are shown in Figures 18-23. As can be seen in these results, the proposed topology works well under highly non-linear load conditions.

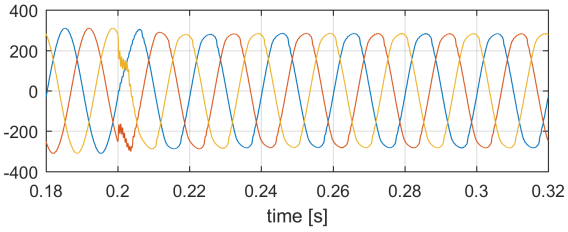


Figure 18: v_{lv} [V] sudden non-linear load connection.

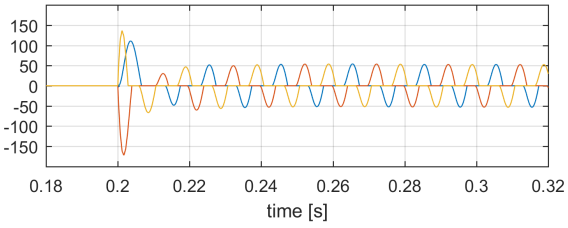


Figure 19: i_{lv} [A] sudden non-linear load connection.

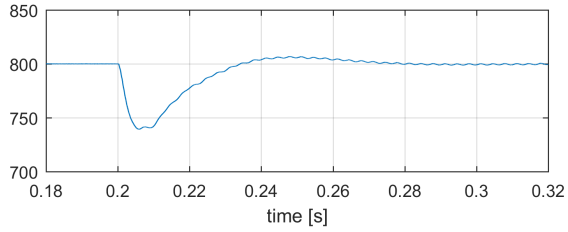


Figure 20: V_{busL} [V] sudden non-linear load connection.

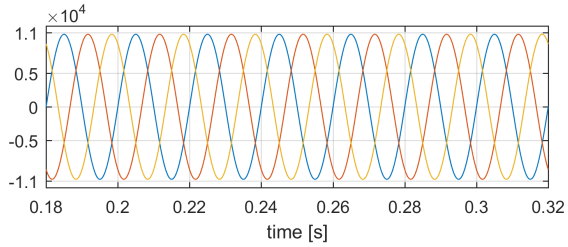


Figure 21: v_{hv} [V] sudden non-linear load connection.

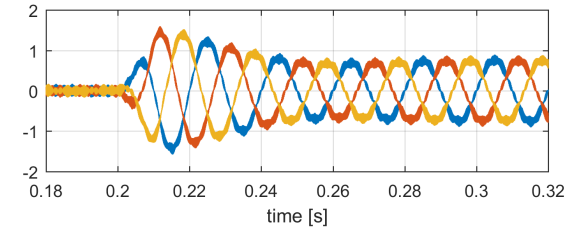


Figure 22: i_{hv} [A] sudden non-linear load connection.

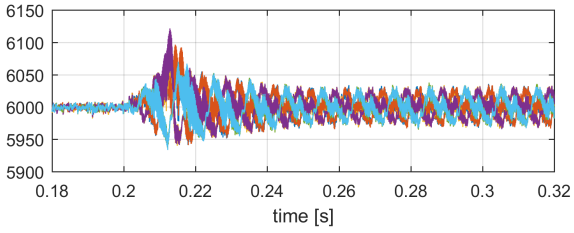


Figure 23: $V_{busH1} - V_{busH6}$ [V] sudden non-linear load connection.

5.4 Sudden non-linear unbalanced load connection

The simulation of the previous section is repeated, using the same non-linear load, but connecting these loads to only two of the three phases. The results of this simulation are shown in 24-29. As can be seen in these results, the proposed topology also works well for unbalanced non-linear loads.

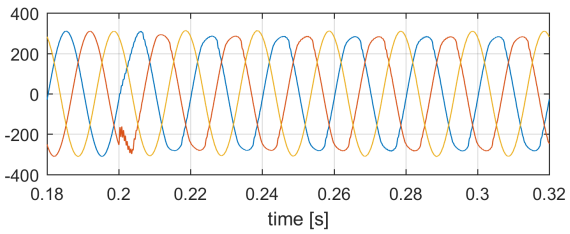


Figure 24: v_{lv} [V] sudden non-linear unbalanced load connection.

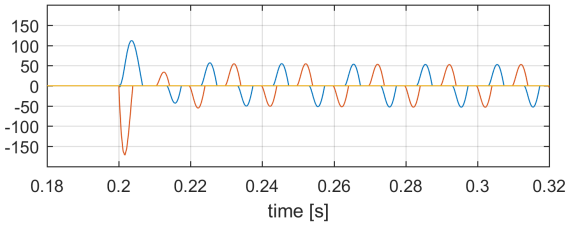


Figure 25: i_{lv} [A] sudden non-linear unbalanced load connection.

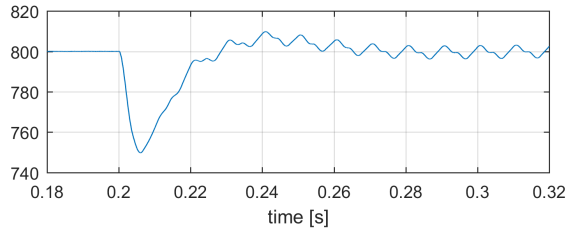


Figure 26: V_{busL} [V] sudden non-linear unbalanced load connection.

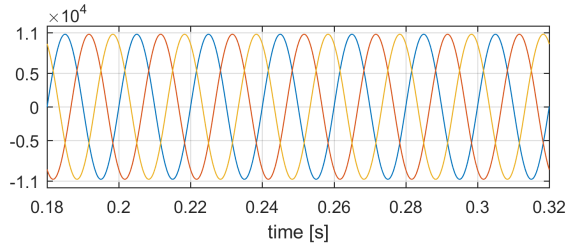


Figure 27: v_{hv} [V] sudden non-linear unbalanced load connection.

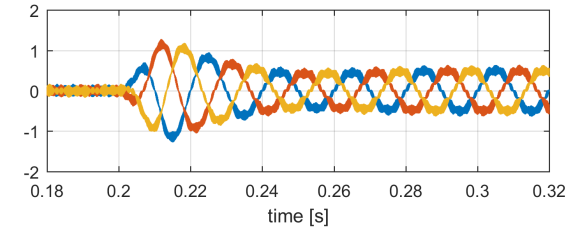


Figure 28: i_{hv} [A] sudden non-linear unbalanced load connection.

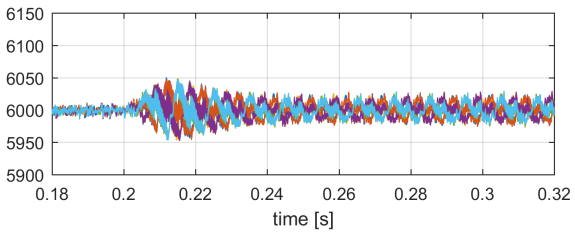


Figure 29: $V_{busH1} - V_{busH6}$ [V] sudden non-linear unbalanced load connection.

6 Conclusions

In this paper a model and a control strategy for an SST topology is proposed. The analyzed SST is designed using three separate stages: a rectifier, a DC-DC converter and an inverter. The proposed design approach is to model each stage using simple models. These simple models help to design and tune the controllers for each stage. The main focus of the tuning procedure is to obtain good performance to sudden load connections. Criteria to choose the main parameters of the SST are also given.

To validate the proposed method, simulation results are presented. The results show that the proposed topology and control strategies perform as expected from design conditions. Moreover, the staged design of the SST allows to decouple the HV side from the LV side in regard to load disturbances, which is an additional feature of the SST when compared to traditional transformers. Simulations were performed for both linear and non-linear loads. Moreover, a non-linear unbalanced load was also tested, with good performance results.

Acknowledgment The authors are grateful to CONICET, UNIVERSIDAD NACIONAL DEL SUR and ANPCyT for their institutional and economic support

Appendix

Given a three-phase signal $f = [f_a\ f_b\ f_c]^T$, its $\alpha\beta$ components are obtained through Clarke's transform:

$$\begin{bmatrix} f_\alpha \\ f_\beta \end{bmatrix} = \frac{2}{3}\begin{bmatrix} 1 & -0.5 & -0.5 \\ 0 & \sqrt{3}/2 & -\sqrt{(3)}/2 \end{bmatrix}\begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix}. \quad (64)$$

Then, the complex vector $\vec{f}$ is defined as

$$\vec{f} = f_\alpha + j f_\beta, \quad (65)$$

where $j = \sqrt{-1}$.

References

[1] C. Busada, H. Chiacchiarini, S. G. Jorge, F. Mengatto, A. Oliva, J. Solsona, G. Bloch, and A. Delgadillo, "Modeling and control of a medium voltage three-phase solid-state transformer," in *2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, April 2017, pp. 556–561. DOI: 10.1109/CPE.2017.7915232

[2] J. Van der Merwe and H. d. T. Mouton, "The solid-state transformer concept: A new era in power distribution," in *AFRICON 2009*, 2009. DOI: 10.1109/AFRCON.2009.5308264

[3] G. T. Heydt, "The next generation of power distribution systems," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 225–235, 2010. DOI: 10.1109/TSG.2010.2080328

[4] S. Falcones, X. Mao, and R. Ayyanar, "Topology comparison for solid state transformer implementation," in *IEEE PES General Meeting*. IEEE, 2010, pp. 1–8. DOI: 10.1109/PES.2010.5590086

[5] X. She, R. Burgos, G. Wang, F. Wang, and A. Q. Huang, "Review of solid state transformer in the distribution system: From components to field application," in *2012 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2012, pp. 4077–4084. DOI: 10.1109/ECCE.2012.6342269

[6] X. She, A. Q. Huang, and R. Burgos, "Review of solid-state transformer technologies and their application in power distribution systems," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 1, no. 3, pp. 186–198, 2013. DOI: 10.1109/JESTPE.2013.2277917

[7] J. W. Kolar and G. Ortiz, "Solid-state-transformers: key components of future traction and smart grid systems," in *Proc. of the International Power Electronics Conference (IPEC), Hiroshima, Japan*, 2014.

[8] F. Wang, A. Huang, X. Ni *et al.*, "A 3.6 kv high performance solid state transformer based on 13kv sic mosfet," in *2014 IEEE 5th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*. IEEE, 2014, pp. 1–8. DOI: 10.1109/PEDG.2014.6878693

[9] J. E. Huber and J. W. Kolar, "Volume/weight/cost comparison of a 1mva 10 kv/400 v solid-state against a conventional low-frequency distribution transformer," in *2014 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2014, pp. 4545–4552. DOI: 10.1109/ECCE.2014.6954023

[10] Z. Yu, R. Ayyanar, and I. Husain, "A detailed analytical model of a solid state transformer," in *2015 IEEE Energy Conversion Congress and Exposition (ECCE)*. IEEE, 2015, pp. 723–729. DOI: 10.1109/ECCE.2015.7309761

[11] H. Fan and H. Li, "High-frequency transformer isolated bidirectional dc-dc converter modules with high efficiency over wide load range for 20 kva solid-state transformer," *IEEE Transactions on Power Electronics*, vol. 26, no. 12, pp. 3599–3608, Dec 2011. DOI: 10.1109/TPEL.2011.2160652

[12] C. A. Busada, S. G. Jorge, A. E. Leon, and J. A. Solsona, "Current controller based on reduced order generalized integrators for distributed generation systems," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 7, pp. 2898–2909, July 2012. DOI: 10.1109/TIE.2011.2167892

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 130-135 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Leaf-shaped solar cell antenna for Energy Harvesting and RF Transmission in ku-band

Chokri Baccouch*, Hedi Sakli, Dhaou Bouchouicha, Taoufik Aguili

National Engineering School of Tunis, Tunis El Manar University, B.P. 37, Le Belvédère, 1002 SYS'COM Laboratory, Tunis, Tunisia.

ARTICLE INFO

Article history:
Received: 31 August, 2017
Accepted: 13 November, 2017
Online: 10 December, 2017

Keywords:
Antenna
Solar cell
Communication systems

ABSTRACT

In this paper, a hybrid system solar cell antenna was studied. Our contribution lies in the combination of both optical and radio-frequency signals in a solar cell antenna. Hence, we propose a new method to merge antenna and solar cell. The leaf- shaped patch solar cell antenna was dedicated at a time to the energy harvesting and the RF transmission. A performance analysis of the solar cell antenna was conducted using Advanced Design System (ADS) software. Simulation results showed a resonance at a frequency of 15.77 GHz with an effective return loss of -27.62dB and a gain of 5.77dBi.

1. Introduction

Solar energy was considered as one of the most renewable energy resources that could fill the shortage of energy needs. In fact, human beings used the solar energy to produce electricity. Although the photovoltaic (PV) cell was first employed to produce it, its efficiency remained limited [1]. Solar energy has become an attractive alternative for powering autonomous communication systems. Many communication systems, such as standalone systems like environmental monitoring system, vehicular communication and satellite systems [2] require electric energy for their operation and use. Solar powered communication systems have received considerable attention due to their ability to operate without the necessity of being connected to an electricity grid. This has become a significant challenge when it comes to powering communication systems in remote places where the electricity grid is not available. In order to address this challenge, the use of photovoltaic in communication systems has recently been the subject of many researches. The photovoltaic systems of power generation when combined with communication systems can provide compact and reliable autonomous communication systems, which can be used for many applications.

Traditionally, telecommunication antennas and solar cells do not mix. They must operate independently so as not to interfere. They compete for the available space on mobile and standalone systems which are generally limited in size. Furthermore they may be bulky and expensive and they limit the capabilities of product designs. In space applications for instance, both solar panels and communication system are major contributors to the overall size and weight of the satellites and combining these two systems could save real estate and cost. In urban areas, wireless communication infrastructures could be integrated into solar panels. Finally the integration of antennas and solar cell panels can be essential for the transportability of emergency autonomous communication stations [3-4] and [5-6].

In this paper, we describe the solar cell antenna in the second section; we detail the proposed approach as well as the effects of optical transparency on the antenna parameters. In the third section we explain the proposed solar cell antenna structure by giving a model of the mesh patch with optimized geometry. The simulations results of the solar cell antenna are presented. A solar cell antenna array has also been studied in this section by giving their parameters. The solar cell antenna equivalent circuit as well as the RF / DC decoupling circuit have been studied in order to eliminate the influence of the RF signal on the DC signal and to show that this hybrid system will be both for the recovery of a solar cell, energy and RF transmission. Finally, we conclude the paper.

2. Solar cell antenna

2.1. Approach

Several attempts to integrate solar cells with patch antennas have been studied, but they did not pay attention to the antenna patch when it is placed above a solar cell. To remedy the drawbacks, especially of limited surface area, encountered during the integration of patch antenna with a photovoltaic cell, we

* Chokri Baccouch, National Engineering School of Tunis, Tunis El Manar University, B.P. 37, Le Belvédère, 1002 SYS'COM Laboratory, Tunis, Tunisia,

propose here new types of antennas based on optically transparent photovoltaic cells with mesh patches.

Optically transparent mesh patch antennas are antennas that have a certain level of optical transparency. These transparent antennas, for example, are potentially suitable for integration with solar cells for small satellites. Traditional patch antennas used on small satellites compete with solar cells for the surface. However, a mesh patch antenna can be placed directly on the solar cells and solve the surface limitation problem. For such integration, a high optical transparency of the patch antenna is required from the solar cells point of view, since solar cells require sufficient sunlight to generate adequate electrical power. On the other hand, the antenna must have at least acceptable electrical properties at the same time so that it can radiate correctly and efficiently [7-8] and [9-11].

The hybrid system studied in this paper was devoted at a time to the energy harvesting and the RF transmission for satellite communication. This structure was made of a photovoltaic cell in which the front grid was designed to have a miniature antenna suited to the band transmission and to minimize the power loss of the cell devoted to energy conversion. The received electrical energy was used to operate the complete system.

2.2. Optical transparency effects on the antenna parameters

The openings in the conductor make it possible to pass the light while the mesh can still be conceived as an efficient radiator. Mesh patch antennas, compared to their solid counterparts, have a higher surface resistance due to much lower conductor coverage. Consequently, it is normal, when the optical transparency increases, to have a low performance of the antenna.

The antenna optical transparency is given most often by the ratio of the mesh patch surface to the patch total area (eq.1).

$$T = \left[1 - \frac{A_{conductor}}{A_{patch}}\right].100\% \qquad (1)$$

It has been demonstrated that, when the line width is fixed, we notice that during the patch transparency increase the frequency of resonance decreases. Since t the patch physical length is the main determinant of the resonance frequency as noted by Clasen [7], the increase in transparency introduces the possibility of minimizing the antenna design. Using the mesh patch increases the fringing fields; therefore, increase the antenna electrical length [12]. Also from Figure 1 one is able to see that increasing the transparency decreases the bandwidth. Bandwidth is defined by the following equation:

$$BW = \frac{f_h - f_l}{f_r} \qquad (2)$$

With f_h the highest frequency at which the voltage standing wave ratio (VSWR) is below 2, f_l the lowest frequency at which the VSWR is below 2, and f_r is the frequency of the minimum VSWR.

Also as seen in Figures.2, 3 and 4, the increase in transparency decreases the antenna efficiency, gain and directivity.

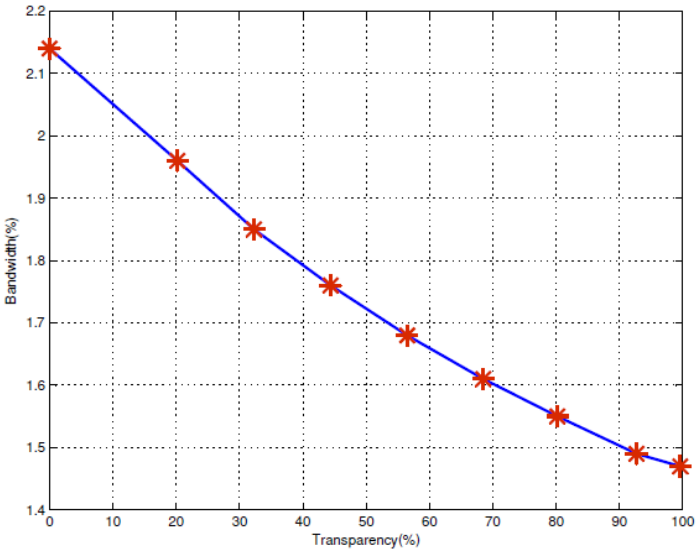


Figure 1: Bandwidth in terms of transparency

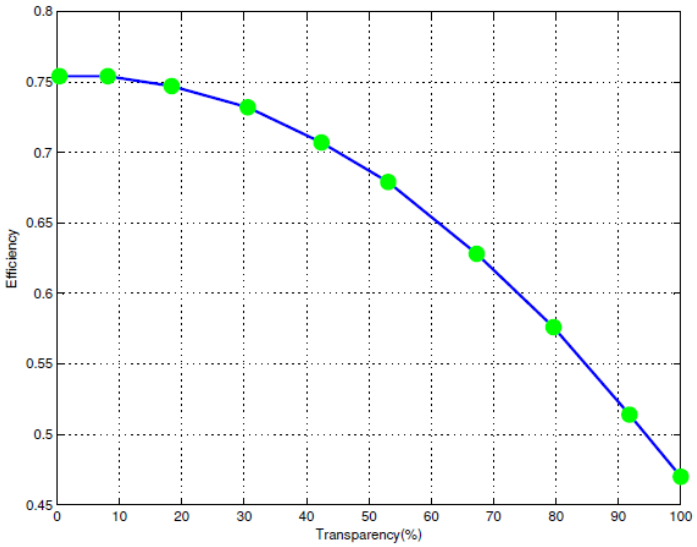


Figure 2: Efficiency in terms of transparency

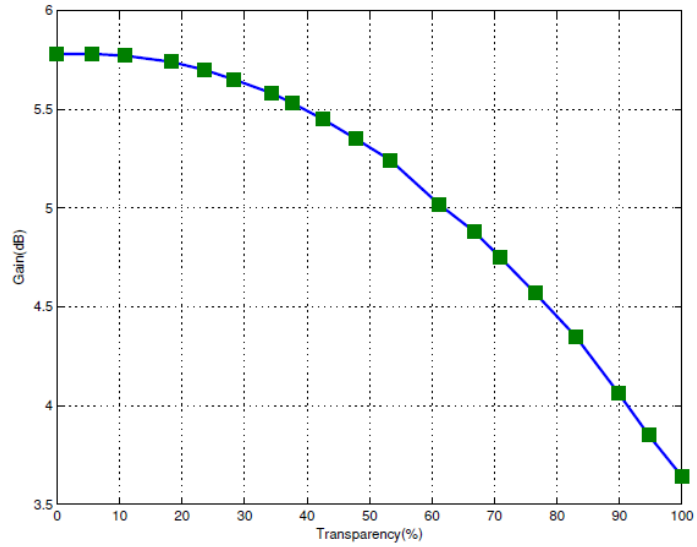


Figure 3: Gain in terms of transparency

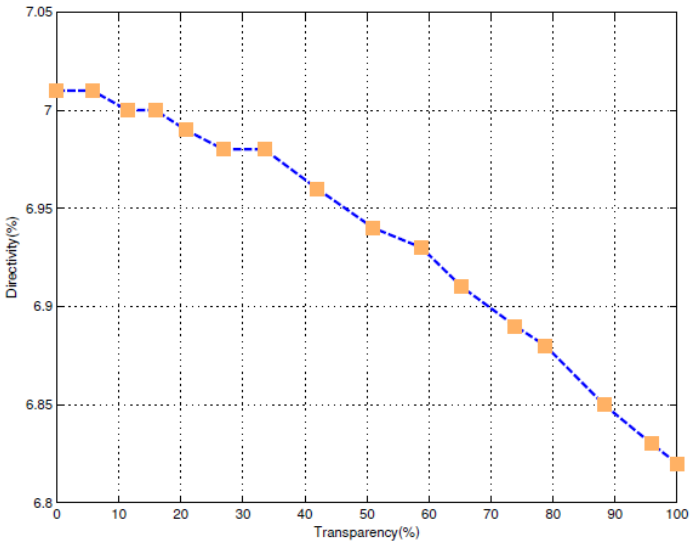


Figure 4: Directivity in terms of transparency

2.3. Effect of the line width

The study of optically transparent antennas is based essentially on the optimization of the mesh patch dimensions. In this paper, antennas with an optical transparency of 70%, 80% and 90% are generated. Each transparency depended of the line width and the number of mesh lines. Simulation results present the efficiency of the mesh antenna as a function of the line width for each transparency is plotted in figure 5. We interpret that for a given transparency, efficiency of a mesh antenna can be improved by refining the line width. From fabrication standpoint of view, the line width realization of 0.1 mm or more is simple. The efficiency and transparency are adequate for this hybrid system of solar cell antenna for satellite and terrestrial communications. From the simulation results (see Figure 5), we conclude that with a line width of 0.1 mm, it is possible to achieve more than 60% efficiency with an antenna of 90% transparency, 63% efficiency with an antenna of 80% transparency and 67% efficiency with an antenna of 70% transparency.

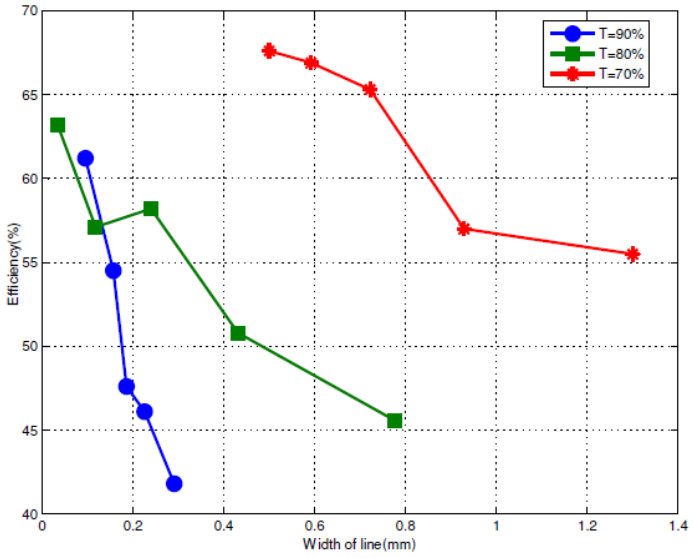


Figure 5: Efficiency in terms of width of line

Efficiency is inversely proportional with optical transparency, this is true because traditional patch antennas are more efficient but when they are integrated with solar cells they screen these and can reduce the collected electrical power. Our method allows us to propose a hybrid solar cell antenna system for energy harvesting and RF transmission for satellite and terrestrial communications.

3. Simulations results and discussions

3.1. Structure of solar cell antenna

We propose a solar cell antenna with mesh patch. A mathematical model that we have already studied to minimize the power losses of the solar cell antenna and improve the conversion efficiency as a solar cell [13-14]. Optimization of the maximum electrical power collected as a function of finger width was determined. The optimal width W_f has been used for the mesh patch design or front face collection grid.

Improving the performance of a solar cell depends not only on the materials and structure but also on the design of the metal grid front face.

With this mesh patch structure two types of waves can exist. The optical waves that will be absorbed by the silicon (semiconductor) and the RF waves that will be collected by the mesh patch metal whose width of fingers W_f is optimized. The structure designed was a solar cell antenna printed on a multi-layered substrate Figure 6.

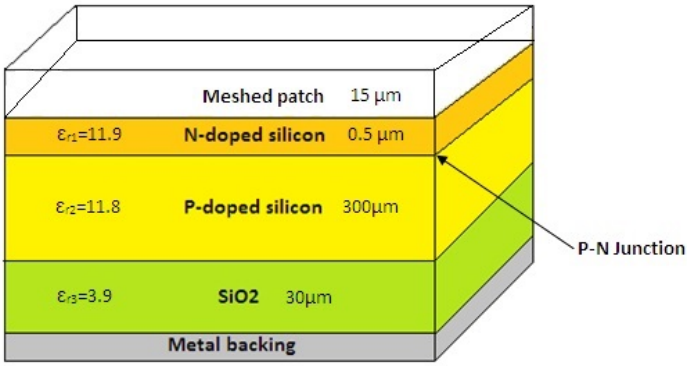


Figure 6: Multi-layered substrate of solar cell antenna

The silicon on an insulating layer SiO_2 confers to the components that are realized, a higher operating frequency, an ability to operate at low voltage and low power consumption and an insensitivity to the effects of ionizing radiation. When using a thin-film substrate we observe that the simulations are complex and time consuming. The meshed patch of solar cell antenna proposed is given in Figure 7. The dimensions of the proposed patch of solar cell antenna are given in the following table (Table 1).

Table 1: Dimensions of solar cell antenna

Parameters	Values
Width of busbars	W_b=0.5mm
Width of finger	W_f=60 µm
Length of feedline	L= 9.1 mm
Width of feedline	W=0.235 mm

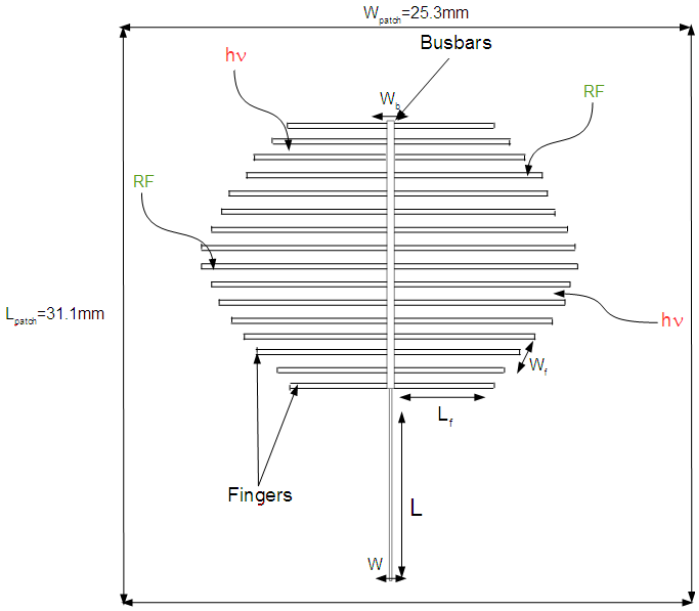


Figure 7: Meshed patch of solar cell antenna proposed

This antenna was excited by a microstrip line of characteristic impedance equal to 50 Ω. In this paper, the transparency of a patch is defined by the percentage of see-though area on the patch as shown below (see equation 3):

$$T = \left[1 - \frac{A_{conductor}}{A_{patch}}\right].100\% = \left[1 - \frac{W_b L_b + WL + \sum_{1}^{32} W_f L_f}{W_{patch} L_{patch}}\right].100\% \quad (3)$$

where W_f is the mesh line width, L_f length of each mesh line, W_b is the busbars line width, L_b the busbars length, W is the InsetFeed width, L is the InsetFeed length, W_{patch} is the patch width and L_{patch} is the patch length (see Figure7).

3.2. Antenna Parameters

As an antenna, the electromagnetic performances of the solar cell antenna were simulated by using the ADS software. The simulated S11 parameter was presented in Figure 8. An antenna should be perfect radiator, rather than perfect absorber. The amount of radiated power returned back through the port can be calculated for finding return loss at that resonating frequency. For the resonant frequencies the return loss should be less than -10dB i.e. S11<-10 dB. Simulation results show that the designed antenna can be used as a frequency antenna with an effective return loss of -27.62dB at 15.77GHz. To the normal scale $S11=10-2.762, ie only 0.17% of the emitted power will be reflected as well as the impedance Z_a=49.82 Ω which confirms the adaptation of the antenna. The adaptation of our antenna is substantiated by the return loss low value S11. Similarly, its operation can be assessed by the VSWR curve which is another performance measurement parameter at the input. It indicates the signal reflection due to a mismatch between the impedance of the antenna and that of the source. The reflected signal leads to the appearance of the standing wave voltage in the feed line which, in turn, destroys the reliability of the antenna. The ideal value of VSWR is 1 which means that 100% of the delivered power is accepted (no reflection). But, practically speaking, it is tolerable up to 2 [15]. The VSWR obtained for this antenna equal to 1.0036.

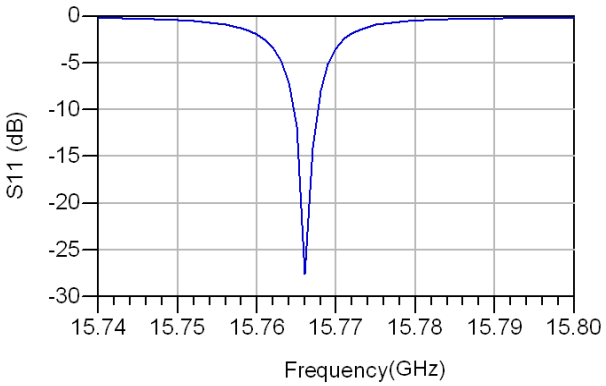


Figure 8: Reflection coefficient S11

The radiation of our antenna is gathered in the main lobes which have a wide beam width at -3 dB and whose amplitude is much higher than that of the side lobes. This reflects a very directive radiation by offering a total directivity equal to 7.26 dB at 15.77 GHz. The radiation pattern of this antenna at a frequency of 15.77GHz is shown in Figure 9. The polarization of the radiated field was linear.

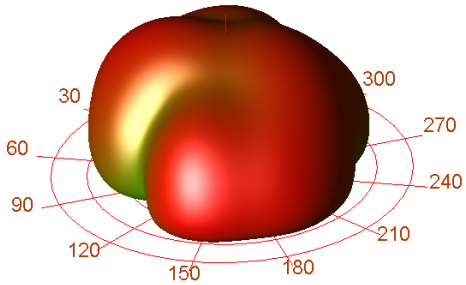


Figure.9: Radiation pattern

When designing an antenna, the gain must be taken into consideration as it is an important metric. The good values of S11 and VSWR are not enough to confirm a good radiation. Antenna gain describes how much power is transmitted in the direction of peak radiation to that of an isotropic source. The gain of this antenna was 5.77dBi.

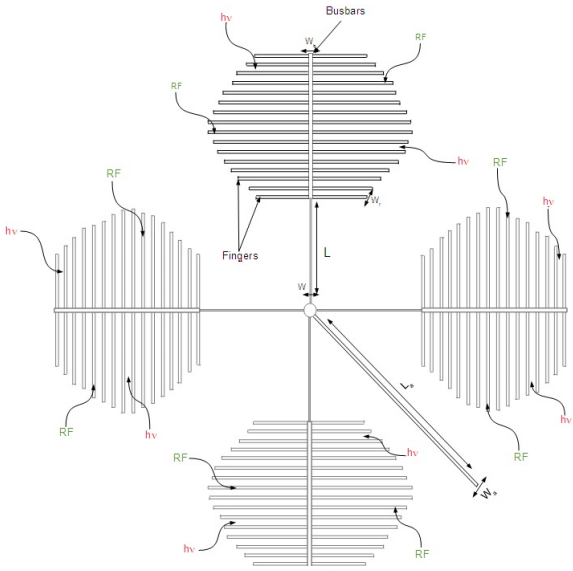


Figure 10: Array of antennas

3.3. Array of antennas

In order to increase the gain, we proposed an array of antennas shown in Figure 10. Simulation results of S11 parameter and radiation pattern are given, respectively, in Figures 11 and 12.

This array of antennas was excited by a microstrip line of La=16.82 mm in length and W_a=0.228 mm in width and of characteristic impedance equal to 50 Ω. S11 parameter shown in Figure 11 of a dual resonant frequency at 15.72GHz and 15.77 GHz with an effective return loss respectively of -20.18 dB and -22.81dB.

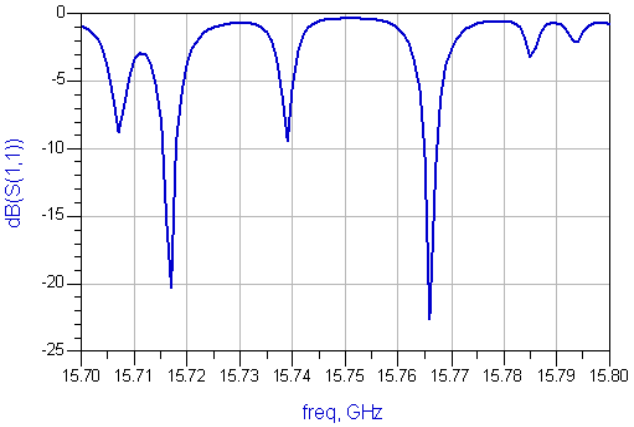


Figure 11: Reflection parameter S11 of array antennas

The radiation pattern of this array of antennas is shown in figure 12. It is gathered in the main four lobes. The gain at the resonance frequency of 15.77GHz is 9.76 dB and the directivity is 11.24 dB.

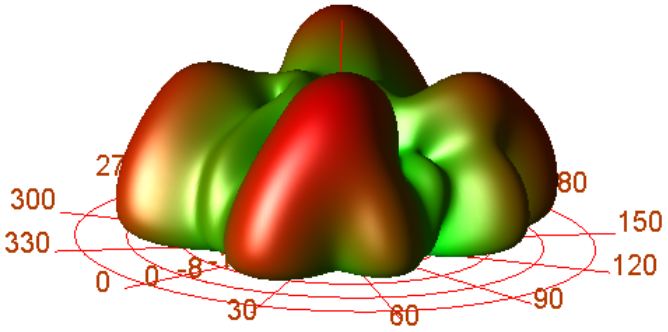


Figure 12: Radiation pattern of array antennas

The designed array antennas are showing excellent gain and directivity at the resonating frequencies with high radiation efficiency.

3.4. Equivalent circuit of solar cell antenna

Solar cell equivalent circuit is specially simulated in Advanced Design System for measuring power (P), current (I) and voltage (V). Microstrip antenna equivalent circuit is separately simulated and the same output is measured. Finally both microstrip and solar cell circuits are combined and output is measured. While comparing both outputs, combination of microstrip and solar cell equivalent circuit produces high voltage (V), current (I) and power (P). The simulation of SOLAN in ADS provides the following values of current (I) =872.8mA Voltage (V) =8.728mV and Power (P) =7.6177984mW were measured as shown in Figure 13.

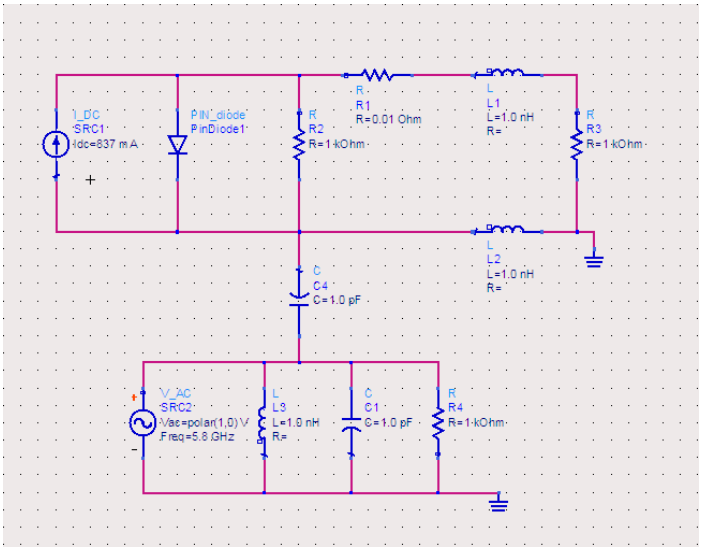


Figure 13: Equivalent circuit of solar cell antenna

3.5. Decoupling RF/DC in solar cell antenna proposed

Since the solar cell has a DC circuit, the direct current path must be decoupled from the RF signal path in such a way that the DC load has no influence on the RF properties of the antenna. The decoupling can be realized by means of concentrated reactive elements and distributed elements, respectively. In this case, we propose an inductive and capacitive RF/DC decoupling circuit shown in Figure 14.

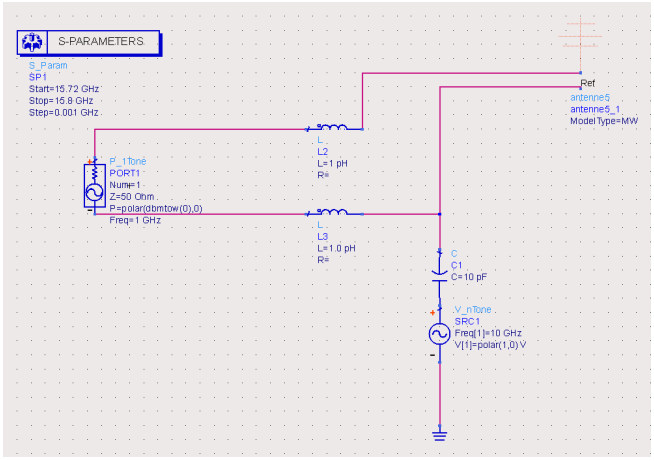


Figure 14: RF/DC decoupling circuit

The decoupling circuit was simulated (see Figure 15) and values are compared to the S11 parameter shown in Figure 8. The simulation of the reflection parameter S11 of decoupling circuit provides the same result of resonant frequency of the proposed solar cell antenna in Layout with a slight increase of return loss.

After RF/DC decoupling, this solar cell antenna can meet the purpose for which it is being studied energy harvesting for the relevant telecommunication system power supply and RF

transmission. Based on the simulations results obtained, the solar cell antenna proposed here can be used for ku band satellite communication. Figure 16 explains well the operation of this solar cell antenna.

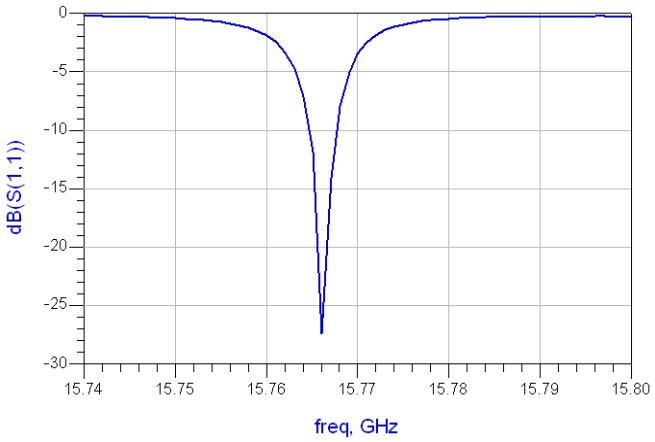


Figure 15: Simulation of RF/DC decoupling circuit

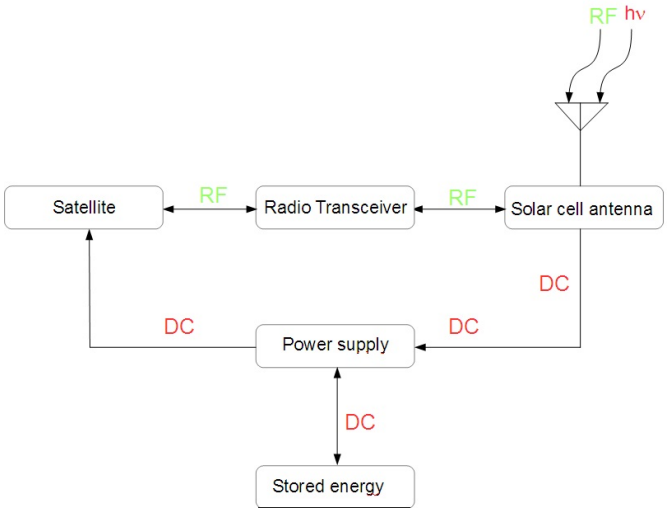


Figure 16: Self-contained wireless system with solar cell antenna

4. Conclusion

In this paper, we presented a solar cell antenna with leaf-shaped patch. This optically transparent antenna is particularly well adapted for energy harvesting and for the RF transmission in Ku band with a large gain. Good performances are obtained for this solar cell antenna, as a PV cell with its collected electrical power and as an antenna with its gain 5.77dBi and 9.76 dBi of array of antennas, directivity and efficiency. We conclude that the hybridized method where the patch element of a patch antenna has been replaced by a solar cell is more efficient than a simple integration where the patch antenna was placed directly on the solar cell is what which limited the available area of the cell and consequently a loss of electrical energy collected.

References

[1] Maharaja M., Kalaiselvan C, "Integration of Antenna and Solar Cell for Satellite and Terrestrial Communication". International Journal of Scientific and Research Publications, **3**(5), 2013.

[2] Danesh, M. and J. R. Long, "An autonomous wireless sensor node incorporating a solar cell antenna for energy harvesting", IEEE Trans. Microw. Theory Tech, **59** (12), 3546--3555, 2011. DOI: https://10.1109/TMTT.2011.2171043.

[3] Shynu, S. V., M. J. R. Ons,P. McEvoy,M. J. Ammann,S. J. McCormack, B. Norton, "Integration of microstrip patch antenna with polycrystalline silicon solar cell". IEEE Trans. Antennas Propag, **57** (12), 3969--3972, 2009. DOI: https://10.1109/TAP.2009.2026438.

[4] A. Suresh Kumar, S. Sundaravadivelu, "Performance analysis of solar cell antenna with hybrid mesh and aght-8 material". Scholarly Journal of Scientific Research and Essay (SJSRE), **3**(4), 51-55, 2014. DOI: http://dx.doi.org/10.1109/ghtc.2011.14.

[5] Yurduseven, O., Smith, D., Pearsall, N., Forbes, I, "A solar cell stacked slot-loases suspended microstrip patch antenna with multiband resonance characteristics for WLAN and WIMAX SYSTEMS", Progress In Electromagnetics Research, 2013, 142, 321--332. DOI: http://dx.doi.org/10.2528/pier13081502.

[6] Ons, M.J.R., Shynu, S.V., Ammann, M.J., McCormack, S., Norton, B, "Investigation on Proximity-Coupled Microstrip Integrated PV Antenna", IEEE Antennas and Propagation,EuCAP 2007, 3, 1--3. DOI: http://dx.doi.org/10.1049/ic.2007.1374.

[7] G. Clasen and R. Langley, "Meshed patch antennas," IEEE Transactions on Antennas and Propagation, vol. 52, pp. 1412-1416, 2004.

[8] C. Baccouch, D. Bouchouicha, H. Sakli, T. Aguili, "Patch Antenna on a Solar Cell for Satellite Communications". International Journal on Communications Antenna and Propagation, **6**(6), December 2016. DOI: https://doi.org/10.15866/irecap.v6i6.9807.

[9] C. Baccouch, H. Sakli, D. Bouchouicha, T. Aguili, "Patch antenna based on a photovoltaic solar cell grid collection". 2016 Progress in Electromagnetic Research Symposium (PIERS). DOI: https:// 10.1109/PIERS.2016.7734278.

[10] C. Baccouch, H. Sakli, D. Bouchouicha, T. Aguili, "Patch Antenna based on a Photovoltaic Cell with a Dual resonance Frequency". ADVANCED ELECTROMAGNETICS, VOL. 5, NO. 3, November 2016. DOI: https://10.7716/aem.v5i3.425.

[11] Turpin, T. W. and R. Baktur, "Meshed patch antennas integrated on solar cells. IEEE Antennas Wireless Propag. Lett, 2009, Vol. 8, 693– 696. DOI: https://10.1109/LAWP.2009.2025522.

[12] C. A. Balanis, Antenna Theory Analysis and Design. Hoboken, NJ: John Wiley and Sons, 2005.

[13] C. Baccouch, D. Bouchouicha, H. Sakli and T. Aguili, "Optimization of the Collecting Grid Front Side of a Photovoltaic Cell Dedicated to the RF Transmission", 2nd International Conference on Automation, Control, Engineering and Computer Science ACECS, 22- 24 March 2015 – Sousse, Tunisia.

[14] C. Baccouch, D. Bouchouicha, H. Sakli, T. Aguili, "Patch Antenna on a Solar Cell for Satellite Communications". International Journal on Communications Antenna and Propagation, 6(6), December 2016. DOI: https://10.15866/irecap.v6i6.9807.

[15] M. Titaouine, "Analysis of Antennas Micro ruban by the model of the cavity, the model of the transmission line and the method of moments", Magister thesis University Ferhat Abbas, Sétif, 1998.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 136-144 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Distributing the computation in combinatorial optimization experiments over the cloud

Mario Brcic*, Nikica Hlupic, Nenad Katanic

University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, 10000, Croatia

ARTICLEINFO

Article history:
Received: 26 October, 2017
Accepted: 12 November, 2017
Online: 10 December, 2017

Keywords:
Combinatorial optimization
Computational experiments
Cloud computing

ABSTRACT

Combinatorial optimization is an area of great importance since many of the real-world problems have discrete parameters which are part of the objective function to be optimized. Development of combinatorial optimization algorithms is guided by the empirical study of the candidate ideas and their performance over a wide range of settings or scenarios to infer general conclusions. Number of scenarios can be overwhelming, especially when modeling uncertainty in some of the problem's parameters. Since the process is also iterative and many ideas and hypotheses may be tested, execution time of each experiment has an important role in the efficiency and successfulness. Structure of such experiments allows for significant execution time improvement by distributing the computation. We focus on the cloud computing as a cost-efficient solution in these circumstances. In this paper we present a system for validating and comparing stochastic combinatorial optimization algorithms. The system also deals with selection of the optimal settings for computational nodes and number of nodes in terms of performance-cost tradeoff. We present applications of the system on a new class of project scheduling problem. We show that we can optimize the selection over cloud service providers as one of the settings and, according to the model, it resulted in a substantial cost-savings while meeting the deadline.

1. Introduction

This paper is an extension of work originally presented in conference MIPRO 2017 [1].

Combinatorial optimization (CO) is a research field with many important real-world applications. Scheduling [2], auctions [3], and vehicle routing [4] are just a few notable examples. Combinatorial optimization is a subfield of mathematical optimization. It deals with problems where optimal selection needs to be done from a discrete feasible set. Exhaustive search evaluates all possible solutions before selecting the best one which is infeasible for realistic problem sizes. There are special classes of CO problems that can be solved with polynomial-time algorithms such as shortest paths, flows, spanning trees, and matching. However, many interesting problems are NP-complete and for these problems, unless P=NP, there are no computationally efficient solving algorithms. For such problems, different search or metaheuristic algorithms are created in order to get as good as possible performance in a realistic amount of time. Design of such algorithms is an intrinsically empirical process, guided by the experiments while the ranking of different design choices, hyperparameter values and algorithms depends on the results from experimental runs, often performed on benchmark test sets.

Each experiment consists of experiment units which denote the smallest indivisible executable unit. Experiment units are in the focus of this paper as they tend to be independent during the execution, which enables a high degree of parallelization. In deterministic CO problems all the parameters are deterministic. For them, each sampled instance of CO problem comprises an experimental unit. In stochastic and robust CO problems, some of the parameters are uncertain or unknown. In that case, each combination of sampled problem instance with its sampled parameter scenarios makes one experimental unit. For that reason, the number of experimental units in stochastic and robust problems can grow exponentially in the number of uncertain parameters. Additionally, the aforementioned set of experimental units is increased in Cartesian product with other experimental factors, as shown in Figure 1. Such factors include hyperparameter values, used algorithm, and algorithm design choices.

*Corresponding Author: Mario Brcic, Unska 3, Zagreb 10000, Croatia, +38516129951, Email: mario.brcic@fer.hr

https://dx.doi.org/10.25046/aj020617

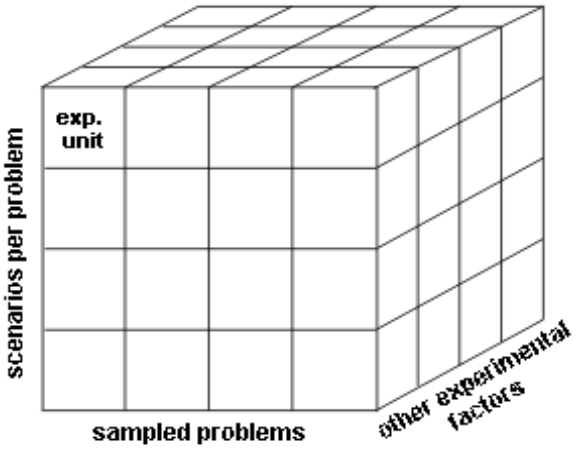


Figure 1. Exponential growth of number of experimental units with added factors which constitute independent dimensions

The experiment requires execution of all experimental units and empirical data must be recorded over those runs. This can be expensive operation, depending on the number of experimental units and the complexity of the underlying problem in each unit. However, as already mentioned, experimental units tend to be computationally independent and the process of experiment execution is inherently parallelizable over the units. In the current state-of-the-art combinatorial optimization, experimentation is mostly done on single computer. This was the case in stochastic variant of vehicle routing problem in [5] and project scheduling in [6]. There were applications where the problem was of such complexity that computer cluster had to be employed in order to make the algorithm practically usable. This was the case for fleet optimization in [7].

In this paper we present an architecture of a system for distributing extensive computational experiments over the cloud that takes advantage of the independence between experimental units to achieve inexpensive scalability in the cloud. This architecture can be implemented as a standalone system or it can use some of the frameworks and systems described in section 2 if the amount of reuse and alignment of the blueprint with the implemented functionalities in those systems can be achieved. The architecture was designed in a way to reduce the amount of communication to a minimum level while still enabling efficient load balancing. We explain the distributed design of data storage for the experimental results. The distributed design reduces the communication overhead which is a problematic aspect for high-performance computing (HPC) in the cloud. Storing rich experimental data is important for further analysis and also important as a mean of scientific scrutiny by enabling efficient sampling-based reproducibility. An optimization model is presented that describes the best choice of settings for running specific experiments. The model uses data from pilot runs which execute sampled subset of experiment units. Finally, the proposed architecture was applied on a specific problem of developing and validating algorithms for a new type of complex stochastic combinatorial optimization problem in project scheduling.

The paper is organized as follows: in section 2 we present the idea of high-performance computing in the cloud as an alternative to in-house grids. Section 3 lists the related work. Section 4 presents the general architecture of our system for distributed experimental runs. In section 5, a distributed design of the storage is described. Applications of the architecture on a real research problem are described in section 6. Section 7 drafts future research ideas and section 8 offers concluding remarks.

2. High-performance computing in the cloud

Distributed computing is unavoidable in high-performance computing (HPC) where job is divided between many available processors in order to significantly reduce the runtime with currently available hardware. Traditionally, dedicated in-house grids (super-computers) are used. They are difficult to setup, maintain and operate [8]. In this paper we shall deal with cloud computing as a flexible and cheap resource alternative that can be rented on demand instead of being owned the whole time – for a lower overall cost and a relatively small performance penalty. Studies have been conducted on the matter of using the cloud for high performance computing. In [9] and [10] authors concluded that there is a limit on a number of used computational nodes where coupled applications are competitive with in-house grids. Beyond that limit, overheads become overwhelming performance detractors. In [11] virtualization, latency and system noise are singled out as the biggest issues in comparison to dedicated supercomputers. They found that research groups with limited access to supercomputer resources and with varying demand might find cloud computing to be a beneficial choice. Latency is problematic due to used commodity equipment in most of the cloud infrastructure and network virtualization. Virtualization introduces performance penalties through network virtualization and other virtualization overheads while accessing physical resources. In [11], network virtualization was found to be the primary bottleneck of the cloud that increases latency, reduces bandwidth and interferes with processes. System noise affects the performance due to multi-tenancy which introduces resource sharing with virtual machines deployed on the same physical hardware. Service providers can manipulate the degree of multi-tenancy, which enables greater profit by overallocating resources to the users. The problems with HPC in the cloud were reiterated in [12] where authors have put the focus on necessary reductions in communication overhead and virtualization. For the former, they propose the implementation of better load balancing, and using bare-metal containers for the latter. Comparative study in [13] confirmed raw performance superiority of in-house grids to Amazon's Elastic Compute Cloud (EC2) cluster. However, waiting time in queue on HPC clusters plays a significant role when taking turnaround time into account . In such circumstances EC2 cluster could produce better turnaround times. The cost-effectiveness of running HPC application was observed as dependable on raw performance and scalability. Cloud computing enables utilizing available monetary resources to rent practically as many as possible identical processing instances. This identicality of processing instances is desirable in running experiments as it sets all runs in the identical environment. This keeps most of the variance in measurements related to designed experimental factors. The effect of system noise on experiment results can be reduced using effective randomization in the job balancing.

3. Related work

In the last decade, with advent of Big Data, usage of cloud computing became all-pervasive. Related to scientific

applications, authors in [14] presented a reproducible genome sequencing task that was run in the cloud for a small cost with a near linear scalability. In [15], a new artificial intelligence algorithm for complex control tasks has been created. It features linear speedups with over a thousand workers on a public cloud service to cut down on total execution time in comparison to other algorithms. Authors in [16] have created distributed architecture for deep neural networks in the cloud for sensor fusion and inference based on the data from multitude of end devices. The communication cost was reduced by a factor of over 20x compared to the alternative. All of the aforementioned applications intentionally achieved low communication requirements between subtasks, hence avoiding or mitigating problems related to the communication overhead in cloud computing. Distributed execution engines have been created to simplify parallelization of computationally intensive tasks. Ray [17] is a python-based example of such system which enables computations.

Existing tools that provide support in design and comparison of optimization algorithms are listed in this paragraph. Comparing Continuous Optimizers (COCO) [18] is a platform for continuous optimization, hence it does not cover combinatorial optimization. Nevertheless, it has many of the features needed in a tool for our needs. It has a library of standard benchmark problems on which optimizers can be compared. Experiments can utilize Shared Memory Parallelism (SMP), but grid computing is not utilized. For that reason, big-scale execution in the cloud is not standard feature of that platform. ParadisEO [19] is a white-box C++ framework for reusable design of parallel and distributed metaheuristics. It has features and components helpful for creating new algorithms. The intent of this framework is to simplify the design of topologies within a single running system, i.e. a single optimizer that can be distributed. However, it does not specify the efficient way of executing distributed experiments. Java Evolutionary Computation Toolkit (ECJ) [20] is an option similar to ParadisEO. Multi-Objective Evolutionary Algorithm (MOEA) [21] Java-based Framework deals with multi-objective optimization by combining the features of COCO, ParadisEO, and ECJ. We have pointed out that COCO covers only continuous optimizers, while other tools enable easier and faster algorithm design. The latter is achieved through reusability of common algorithmic components when the optimization problem and algorithm design have favorable features. These tools do not specify guidelines for distributing extensive computational experiments over the cloud.

Performance of executing the experiment in the cloud is an important issue. Predictions can be used in scheduling as well as in finding optimal settings of computational nodes. The focus of [22] is on comparing public cloud providers using measurements on specific applications. These measurements can inform the processes of provider selection and performance prediction. Performance prediction using machine learning for improving the quality of system management decisions has been investigated in [23]. Authors in [24] used machine learning to predict the execution time of computational fluid dynamics applications in the cloud. These predictions were used in scheduling algorithms. A system for efficient performance prediction for large-scale analytics on EC2 cloud has been created in [25]. It utilizes optimal experimental design in order to minimize the resources in building the model. The system is used to find the optimal configuration in number of instances. Our architecture uses simple statistical model for performance prediction in order to calculate cost-optimal instance-type and number of necessary instances in order to satisfy desired probabilistic level of satisfying the deadline. In our case, settings also include the cloud provider, hence combining the intents of aforementioned works: performance prediction and optimization of node-selection that takes into account the cloud provider as well.

4. Architecture

The intended use-case is inherently parallelizable task. This architecture is designed in a way to use those favorable features of the task in order to achieve low communication between the computational nodes. In that way, latency is not an issue and there is only occasional communication where bandwidth plays the main role. Communication between the nodes is necessary for creating the computational nodes, sending instructions for job chunks (that is, batches) to them, and during migration of the final results. Instructions for job chunks contain small amount of information. Chunks are sized in a way to keep the nodes occupied for some time. There is a tradeoff between achieving good job balancing and reducing communication overheads in relation to the amount of computation done on the node. Computational nodes keep all generated and logged experimental data locally in their part of distributed database. The results, raw or processed, can be pooled periodically, upon finalization of the assigned experiment chunk or at the end of node's part in the experiment. Aforementioned features make the problem suitable for cloud deployment as we can avoid communication-related detrimental effects on performance.

Our pipeline architecture includes five general stages in the experimental process:

1. Creation of experimental data, that is the data about experiment units in initial seed database.

2. Creation of node images populated with all the necessary data and optimized code for executing the experiment units, logging and storing the results.

3. Optimizing the settings of execution environment. Pilot runs are used on a small subset of experiment units for different settings of the nodes, including a service provider to inform the optimization procedure.

4. Executing the experiment on pre-calculated number of identical nodes with settings selected in the previous stage.

5. Collecting experimental data from computational nodes, conducting analysis and getting the results.

In the rest of this section, we shall describe each of the stages in more details. All steps are enumerated according to their corresponding figures.

4.1. Creation of the experimental data

The data needed for running an experiment is created in the first phase, which is depicted in Figure 2. This means creating all the data that sufficiently describes experiment units so that they can be created and executed.

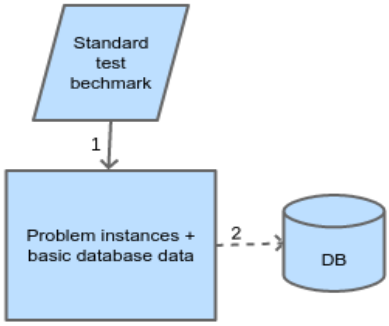


Figure 2. Preparing the seed database

The first step in this stage is optional. It uses standard test set as a source for creating problem instances. For many well-researched CO problems, such sets are shared within the community of the researchers to enable consistent ranking of algorithms. In case of new CO problems, maybe a standard test benchmark related to the problem can be found and used for creation of problem instances. Also, problem-instance generators are available for some problems and custom test set can be created. In this step, sampling and transformations of problem instances can be done to create our own set of problem instances that are going into an experiment.

Actually used test set is created at this point. The "seed" database is created and populated with the metadata necessary for experiment execution. Also, it is filled with all the data about experiment units needed for execution and result logging. Part of the data can be in the form of external files if that is more appropriate, but they must be linked from a database. In order to enable distributed execution, a database (and external data) is distributed over the nodes in such a way that each node has its independent, unsynchronized version of a database. Each such database is initiated from a singular seed database. Horizontal fragmentation of writeable relations is employed as a mean of data distribution over the nodes. Details of a database are given in section 5.

4.2. Creation of node images

One or more virtual machine (VM) images need to be created at this stage (shown in Figure 3). The exact number depends on the requirements in a phase of pilot runs. Each image contains a copy of a seed database and all the necessary code for running the experiment chunks on the node. Executable code is tuned to the intended hardware. After that, image is migrated to the cloud service from where it can be easily deployed for creation of the computational nodes.

4.3. Optimizing the execution environment

It is hard to know exactly in advance what settings of the computational node are efficient. For that reason we need an optimization phase of an experiment with regards to the settings of the experiment run. This is a stochastic combinatorial optimization problem as well and it can be stated as a problem of minimizing the monetary cost of renting cloud instances under a constraint of

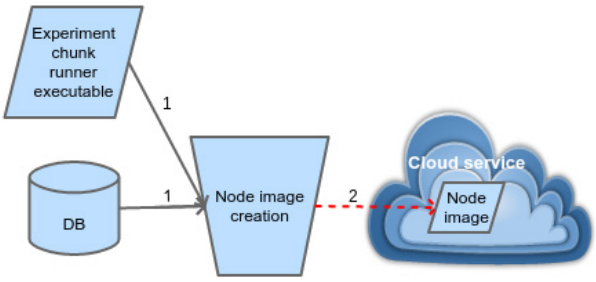


Figure 3. Creating the image for virtual machines in computational nodes

achieving desired probability of finishing before the selected deadline. The problem can be formulated and expressed as:

$$\min \mathbf{E} \sum_{x \in S} C_x(n_x) \quad (1)$$

with the constraints:

$$\sum_{x \in S} \delta_x = 1, \quad (2)$$

$$\mathbf{P}\{T_x(n_x) \leq D\} \geq \delta_x \cdot p, \ \forall x \in S, \quad (3)$$

$$\delta_x \in \{0,1\}, \ \forall x \in S, \quad (4)$$

$$n_x \in \mathbf{N}_0, \ \forall x \in S. \quad (5)$$

In the above formulation **E** is the expectation operator and **P** is a probability measure of the set. The above problem (1)-(5) has several parameters with values known prior to optimization: D is a desired deadline, S is a set of node-type options, and p is a desired probability of achieving the deadline. Also, C_x is a random cost of running the experiment on n_x instances of node-type x. The cost is random as it depends on the utilization durations of n_x nodes. T_x is a random experiment finish time when running it on n_x instances of node-type x. Decision variables are n_x and δ_x where n_x represents a number of instances of type x (5), and δ_x represents exclusive choice between the node-types (4). Randomness in this problem originates from the noise in execution due to hardware reasons and uncertainty in computational requirements of each experiment unit. That randomness is reflected in C_x and T_x. These functions can be created using statistical analysis or machine learning on the data from pilot runs. The objective function (1) is the expected value of total cost of the experiment. It sums the costs across all possible node-types but only one of those costs is going to be non-zero. The first constraint (2) ensures that only one node-type is selected. A set of chance-constraints (3) ensures that for selected node-type x a number n_x of nodes is selected so as to achieve a desired level of safety. All zero-valued δ_x make these constraints trivially satisfiable for all non-selected node-types x. For the selected node-type, a constraint enforces that a probability of finishing before the deadline must not be lower than the prescribed p.

We propose to make a parameter search over a set of node settings by running a simple and small experiment (using only a subset of experiment units) on each setting in S. A set of options should be small in order to reduce the cost of doing the pilot runs. That set can be composed by a careful pre-selection based on available data from previous general analyses such as [22] and based on analysis of the code for experiment execution as in [26].

The settings with the most impact are the selected cloud-service provider and cloud-instance type. The latter can be further customized across some set of subsettings, but usually the types are predefined less flexibly. The most important subsettings pertain to hardware components: CPU, disk size, RAM, etc. Pilot runs are identical to executing the experiment described in the following subsection. The difference from the latter is in the scale; in pilot runs just one experimental node per different setting is used , while many experimental nodes with identical settings are used during the experimental run. Collected data are analyzed by a procedure described in subsection 4.5. Based on the collected data and optimization model, economically most efficient node setting is chosen for the full experiment. Also, for that choice we get an estimate of the total experiment run time and cost by taking the summary statistics of T_x and C_x. Based on a desired due date, we infer the necessary number of computational instances N. N is the value of only non-zero n_x in optimal solution of the the optimization model.

4.4. Executing the experiment

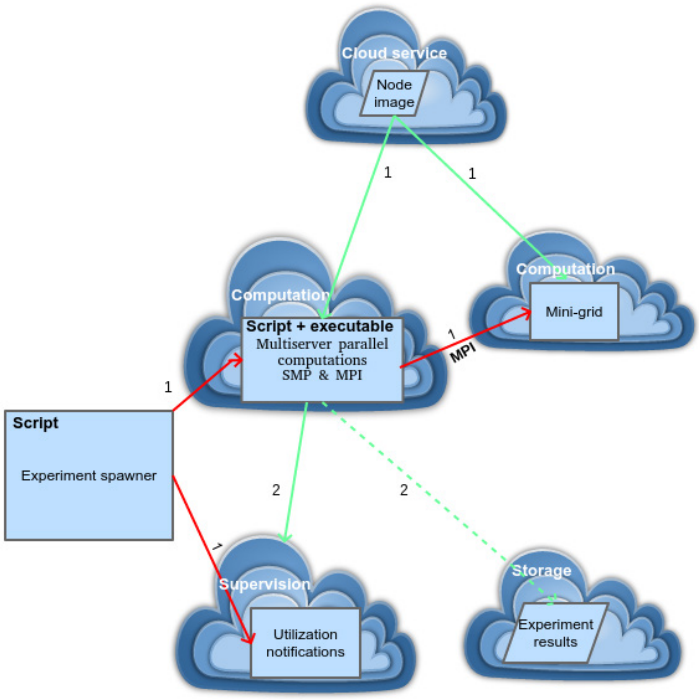


Figure 4. Execution of the experiment

The experiment is executed by distributing chunks across computational nodes (Figure 4). Experiment spawner is a script that balances the load between instantiated computational nodes. The experiment is partitioned into disjoint experimental chunks. Each chunk is a set of experiment units to be executed on a single computational instance.

1. Experiment spawner creates N cloud nodes and sends over network the parameters that define their workload chunks. These parameters constitute a small amount of information. Each experimental node runs a chunk runner which processes its assigned workload. Each runner, depending on the type of a problem, can instantiate additional computational nodes to form a mini-grid using Message Passing Interface (MPI). That is done if some of the algorithms require such execution architecture by design. SMP can be switched on by the parameters sent from the experiment spawner. Each node sets triggers for utilization alarms at the performance supervisor. Triggers improve the efficiency of a system by notifying the subscribers of different events. This information can be used for better, more prompt load balancing and it can reduce the renting costs.

2. At the end of processing a chunk, a computational node triggers supervisor's alarm when it can get another chunk to execute. When the node finishes with processing, it migrates its results to a permanent cloud storage and it gets terminated.

4.5. Result analysis

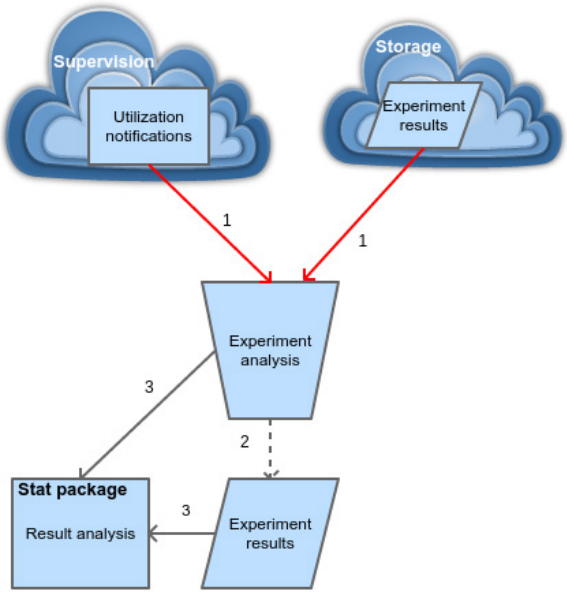


Figure 5. Analysis of experimental results

Users are notified by supervisor's alarm about the finalized jobs. A procedure is subscribed to notifications and reacts to them as shown in Figure 5. At first, the data is migrated from a computational node to some storage, locally or in the cloud. If the node has no more jobs to process, it is terminated to reduce costs. Analysis is done iteratively over the partial data as they pool to the storage. At the end of the process, with all the experimental results available, concluding results are created.

5. The database

The main objective of our system is to produce detailed experimental data as fast as possible. We have decided to store all the data into a database as it simplifies manipulations with overwhelming amounts of data.

It is assumed that running computational experiments is expensive both temporally and monetary. Storage, on the other hand, is much cheaper on both accounts. For that reason, as much data as possible should be stored for future analyses to avoid experiment re-runs.

Practicing scrutiny is important in science in order to prune mistakes and misconduct. Replication studies simply repeat experiments to check if the results match. In computer science, this

can be done by sampled replications that repeat small sampled set of experiment units. In each repetition, the same settings for pseudo-random number generators (PRNG), algorithms and problem instances should be used to create near-identical conditions as with the initial experiment. The results of replication study must match the stored values for everything except the execution time which includes incontrollable system noise. This is more efficient than replications of physical experiments where the exact conditions cannot be repeated and results should only statistically match. In the latter case, a greater number, if not all, of experimental units needs to be repeated in order to make conclusions. Also, executing experiment units in computational experiments is often cheaper than for physical experiments. The probability of non-matching results gets exponentially smaller with the size of randomly generated sample for replication.

5.1. Design

As described in section 2, communication overheads are serious performance detractors to scalability in the cloud in comparison to the in-house grids. Grids have tight interconnection and synchronization. For that reason, we have decided to minimize communication frequency between the cloud instances. All instances have their own database for storing results that springs from the initial seed database which is copied to all instances during creation of the node image as described in subsection 4.2. Seed database holds metadata and identification/replication data. The former defines all necessary structures to store the experiment data. The latter are the basic data necessary for identification of experiment units. Such data include the shared information for all instances of CO problems, used optimization algorithms, PRNG types and uncertainty scenarios in the case of stochastic or robust CO. Hence, all the data that define and describe experiment units (Figure 1) are present in a seed database.

Each node is created with a separate copy of a seed database. These copies make up a distributed database. The writeable relations, which record the experimental run data, are horizontally fragmented. Horizontal fragmentation keeps table schemas and distributes table rows across the nodes, as depicted in Figure 6.

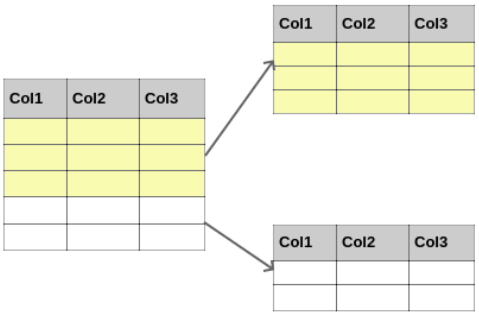


Figure 6. An illustrative example of horizontal fragmentation

During the experiment execution, each node saves two types of data: identification/replication data, and performance data. The former describe the settings of an executed experiment unit. This consists of the settings of the used optimization algorithm, PRNGs and other components that hold the key to replicating the execution of the experimental unit during replication studies. Performance data track decisions made during the execution and measurements of their effects. Quantitatively they make up the majority of generated and stored data, and the final conclusions are based on them.

Distributed data uses the identification scheme that combines together node-specific identification/replication information with shared information that originated from a seed database. That scheme makes data aggregation from all chunks unambiguous.

6. Application

The primary motivation for creating previously described architecture was a practical problem. As a result of different branch of research, we have worked with the new type of stochastic project scheduling problem, Cost-based Flexible Stochastic Resource Constrained Project Scheduling Problem (CBF-SRCPSP), defined in [27]. This problem deals with proactive-reactive project scheduling which makes the synchronization between project collaborators easier. It uses an agreed upon baseline schedule that stores the time-agreements for starting times of different project activities. Deviations of real start times from these agreements are penalized for inflexible, change-sensitive activities. The additional element in this model is that the changes to the baseline schedule can be profitable if they can be undertaken sufficiently far into the future. This extension to the features of baseline schedule makes the solution space much more complex due to the aforementioned proactive rescheduling operations. Searching the solution space becomes more time-consuming, though it pays off. We were lead empirically through the design and creation of optimization algorithms for CBF-SRCPSP.

The experimental set was created based on the standard Project Scheduling Problem Library (PSPLIB) set of instances for deterministic resource constrained project scheduling problems [28]. Cluster sampling was used to select sets of template instances from J30, J60, and J120 problem sets. The latter consist of projects that have 30, 60, and 120 project activities, respectively. The templates were expanded to fit CBF-SRCPSP model by modeling the stochasticity in activity durations with discretized beta distributions with a combination of selected and randomly generated distribution parameters. Additional parameters for each activity were randomly generated, according to the selected triangular distributions. After this procedure we ended up with 300 instances of each size which sprang up from template instances extended to fit the new model. The templates were instantiated according to the two experimental factors regarding to the project deadline: tightness of the deadline and bonus for early-finishing the project. For each instance, we generated 1000 activity duration scenarios from the discretized beta distributions. Activity durations were the only source of uncertainty. All of this data: project instances and uncertainty scenarios per each project were fed into a seed database. The used database was sqlite3 as it fit the needs of our experiment. Its simplicity trades-off well with its shortcomings in comparison to the more elaborate database management systems.

The seed database was also populated with the metadata about used PRNGs and CO algorithms. We have used two PRNGs: Mersenne twister with careful parameterization [29] and Threefry [30]. Several search algorithms were developed during the algorithm design. We had at our disposal implementations of optimization algorithms from the CO literature that were already

available in the used simulation library [26]. The final experiment included one benchmark CO algorithm from the literature and two selected newly-developed algorithms that were hypothesized to significantly outperform the benchmark. However, the analysis necessary for proving the hypothesis necessitated sufficient number of samples. The newly-developed algorithms were computationally expensive due to a search in more complex solution space than the benchmark.

The node images used Linux with gcc and necessary libraries. New algorithms were developed upon the C++ simulation library from [26]. This means that the most computationally intense parts, namely simulation-based experimentation and database logging were coded in C++ for better performance. The mini-grids were not used, since all the algorithms used only shared memory parallelization with two threads using OpenMP to speed up the search. This choice was based on the experiment in [26] regarding the parallel efficiency of the used simulation library. That also determined the number of computational cores per instance.

The total experiment was run on two occasions. The second experiment was done due to improvements to the developed algorithms and the results from the second study were used in the final experiment report. Here we shall cover both experiment runs as they utilized different experimental choices. In both cases, the raw results were downloaded locally using the python script. Statistical language R was used for results analysis. RSQLite R package was used to query the databases for the relevant data and to gather them together from all the sources.

6.1. The first experimental run

During the first experiment run we have selected Amazon Web Services (AWS) [31] as a service provider. Therefore, we did not use service provider as the experimental factor in pilot runs during the selection of the ideal instance type. We have used EC2 for computations, and have opted, after pilot runs, for c3.large instances with 2 dedicated physical cores of Intel Xeon E5-2680 v2 (Ivy Bridge) processors and 3.75 GiB of RAM. Simple Storage Service (S3) was used to store experimental results. The combination of CloudWatch (CW) and Simple Notification Service (SNS) was used for supervision and utilization notifications. Experiment spawner was coded in python, using the boto API [32] for accessing AWS and paramiko module [33] for controlling SSH2 connections.

Experiment consisted of 1.6 million experiment units. Experiment chunking was done across the problem instances. The chunk runner accepts parameters that describe chunk boundaries. These parameters are just several bytes in size and it is all the information needed for initiating the experiment execution on a computational instance. The total computational workload of the experiment was estimated based on running small sample of experiment units during the pilot run. It was estimated that the total workload is three and a half months of computational labor on the available hardware. The deadline was set to 7 days which we wanted to achieve with 90% probability. We calculated the necessary number of computational instances to satisfy this requirement. This resulted in using up to 26 cloud instances and reducing the total duration to 6 days. The database for each instance was migrated to S3 at the end of chunk execution and it awaited further analysis there. The cost of the first experiment was 445$.

6.2. The second experimental run

The second experiment was run on improved algorithms. The number of the most computationally expensive units was significantly increased, resulting in the total of 2.1 million experimental units. The optimization stage was used to select the node-type, also taking into account the service provider. We have used prior knowledge of the characteristics of our experiment runner - low working memory consumption (under 500MB) and high CPU utilization - to narrow down a range of instances. We tested the instances with 2 processors and as close as possible to 2GiB of memory. Market research was used to select the small set of service providers: Online Virtual Hosting (OVH), AWS (due to the use in the first experiment), and Linode. The selected instance-types are listed in Table 1.

Table 1. Members of alternative set *S* for optimization of execution environment

	OVH	**Amazon EC2**	**Linode**
instance-type	2vCores@ 2.4GHz 8GB RAM	c3.large 2vCores@ 2.8GHz 3.75GB RAM	2vCores@ 2.5GHz 4GB RAM
price	13.49$ (monthly)	0.105 $(hourly)	20$ (monthly)

The identical pilot run was used on all instance-types. Seven different types of experiment units were sampled into the pilot's unit set. The price of pilot runs in the optimization stage was 1.1$. The measurements were used to model execution durations for each combination of instance-type x and unit type u as Gaussian random variables $d_{x,u}$. Then, the total experiment duration for each instance-type is Gaussian random variable where w_u is the number of units of type u in our experiment. We approximated T_x by assuming that the total work modeled by d_x is simply equally divided among n_x computational instances.

It was estimated with the probability of 90% for the fastest option that the total workload is just above three years. The ranking and the necessary number of instances were calculated in order to satisfy the selected deadline of 30 days with the probability of 90%. The deadline was set to 30 days to take advantage of the monthly pricing. Figure 7 shows the optimal expected total costs for each option in Table 1. The results in Figure 7 were anonymized due to the legal concerns. Based on the available data, we have selected the C3 type instance as the most efficient with the estimated cost just below 500$. The experiment was run and it finished after 23 days of executing.

7. Future work

Possible future research ideas include improvement to the robustness of the total workload estimator. In our application, different simplifying assumptions were made and the error of their approximation effect should be investigated.

In order to reduce costs of the research community, general data about computational experiments (duration and prices) could be shared online with the public. This data can be useful for

creating promising and efficient alternative sets for optimizing the execution environment.

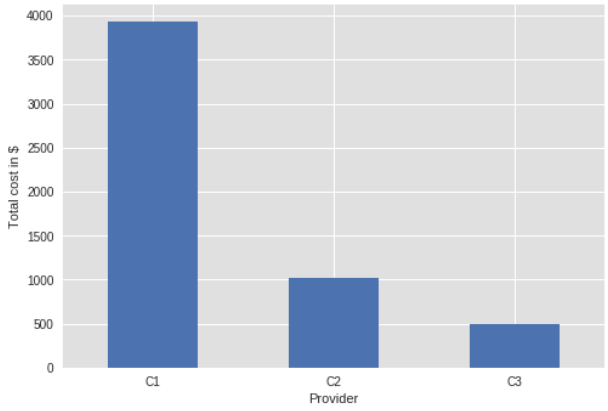


Figure 7 Estimates of total expected cost for anonymized permutation of options from Table 1

An automated tool for the composition of alternative sets could be utilized based on the experiment requirements. There is a potential in using pilot run on just one carefully selected instance type and then utilizing machine learning to find the best option and the number of necessary instances N for the desired target duration and the probability of achieving it. More advanced notions of risk and chance constraints can be used to account for the uncertainty in estimations and its economic effect. That can inform budget planning and the project management that undertakes the experiment.

Many parts of the architecture presented in section 4 are abstracted from the details of the particular experiments and can be reused in different settings. Component-based framework for general experimenting can be created. In that way, reuse of existing components can be increased and a code generator for repetitive parts can speed up the development, especially if the user does most of the manual process through an intuitive graphical user interface. Cloud costs can be reduced further by better node tracking and possibly using the cheaper spot instances for non-critical computations, especially during the algorithm design, prototyping and various pilot runs. The performance penalty in case of using such instances should be investigated in order to drive recommendations for configurations that utilize them.

8. Conclusion

Cloud computing is still not a simple and clear choice for high performance computing due to the issues pointed out in [11]: communication overhead, virtualization and system noise. The efficiency of cloud depends greatly on the specifics of the problem that we are trying to solve.

We have presented a system for distributing combinatorial optimization experiments over the cloud. Doing computational experiments for validation and guiding the design of CO algorithms has a specific property that it can be parallelized across experimental units that tend to be independent. This allows for low coupling between the simultaneous tasks and circumvents the issues related to the communication overhead.

Our system records rich data about the experimental runs in order to reduce the need for re-runs of experiments which may be computationally and monetarily expensive. In order to keep communication overhead to the minimum, distributed database with horizontal fragmentation was used. Each cloud node populates only the data related to its assigned disjoint experiment chunk. The unambiguity of the data across the system is, hence, kept without additional effort. When the tested algorithms use distributed computations in mini-grids, they should keep the grid size within the limits of recent studies, such as given in [11], to get the best performance benefits. It is expected that additional tuning and advances in cloud computing technology will increase the limits found by these studies.

We described two successful applications of our proposed system on the newly developed algorithms for complex stochastic combinatorial optimization problem, CBF-SRCPSP. Initial estimated sequential duration of several months to several years was reduced to under a month (the first experiment under a week) by distributing the execution. The optimization stage of our architecture finds the best settings for the execution environment. This enables the selection of the best instance-type across different cloud-service providers. In our case, that significantly reduced the cost of running the experiment.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] M. Brcic and N. Hlupic, "Cloud-distributed computational experiments for combinatorial optimization," in 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017, pp. 197–201., http://doi.org/ 10.23919/MIPRO.2017.7973417

[2] S. Westphal, "Scheduling the German Basketball League," Interfaces, vol. 44, no. 5, pp. 498–508, Oct. 2014.

[3] Y. Sheffi, "Combinatorial Auctions in the Procurement of Transportation Services," Interfaces, vol. 34, no. 4, pp. 245–252, Aug. 2004.

[4] G. Kant, M. Jacks, and C. Aantjes, "Coca-Cola Enterprises Optimizes Vehicle Routes for Efficient Product Delivery," Interfaces, vol. 38, no. 1, pp. 40–50, Feb. 2008.

[5] J. C. Goodson, J. W. Ohlmann, and B. W. Thomas, "Rollout Policies for Dynamic Solutions to the Multivehicle Routing Problem with Stochastic Demand and Duration Limits," Oper Res, vol. 61, no. 1, pp. 138–154, Jan. 2013., http://dx.doi.org/10.1287/opre.1120.1127

[6] P. Lamas and E. Demeulemeester, "A purely proactive scheduling procedure for the resource-constrained project scheduling problem with stochastic activity durations," J. Sched., vol. 19, no. 4, pp. 409–428, Aug. 2016., http://doi.org/ 10.1007/s10951-015-0423-3

[7] H. P. Simão, A. George, W. B. Powell, T. Gifford, J. Nienow, and J. Day, "Approximate Dynamic Programming Captures Fleet Operations for Schneider National," Interfaces, vol. 40, no. 5, pp. 342–352, Jul. 2010., http://doi.org/ 10.1287/inte.1100.0510

[8] C. Vecchiola, S. Pandey, and R. Buyya, "High-Performance Cloud Computing: A View of Scientific Applications," in 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009, pp. 4–16., http://doi.org/ 10.1109/I-SPAN.2009.150

[9] C. Evangelinos and C. N. Hill, "Cloud Computing for parallel Scientific HPC Applications: Feasibility of Running Coupled Atmosphere-Ocean Climate Models on Amazon's EC2," in In The 1st Workshop on Cloud Computing and its Applications (CCA), 2008.

[10] D. Kondo, B. Javadi, P. Malecot, F. Cappello, and D. P. Anderson, "Cost-benefit analysis of Cloud Computing versus desktop grids," in 2009 IEEE International Symposium on Parallel Distributed Processing, 2009, pp. 1–12., http://doi.org/ 10.1109/IPDPS.2009.5160911

[11] A. Gupta et al., "The Who, What, Why, and How of High Performance Computing in the Cloud," in 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, 2013, vol. 1, pp. 306–314., http://doi.org/ 10.1109/CloudCom.2013.47

[12] D. Tomić, Z. Car, and D. Ogrizović, "Running HPC applications on many million cores Cloud," in 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017, pp. 209–214., http://doi.org/ 10.23919/MIPRO.2017.7973420

[13] A. Marathe et al., "A Comparative Study of High-performance Computing on the Cloud," in Proceedings of the 22Nd International Symposium on High-performance Parallel and Distributed Computing, New York, NY, USA, 2013, pp. 239–250., http://doi.acm.org/10.1145/2462902.2462919

[14] S. S. Shringarpure, A. Carroll, F. M. D. L. Vega, and C. D. Bustamante, "Inexpensive and Highly Reproducible Cloud-Based Variant Calling of 2,535 Human Genomes," PLOS ONE, vol. 10, no. 6, p. e0129277, Jun. 2015., http://doi.org/ 10.1371/journal.pone.0129277

[15] "[1703.03864] Evolution Strategies as a Scalable Alternative to Reinforcement Learning." [Online]. Available: https://arxiv.org/abs/1703.03864. [Accessed: 10-Sep-2017].

[16] S. Teerapittayanon, B. McDanel, and H. T. Kung, "Distributed Deep Neural Networks Over the Cloud, the Edge and End Devices," in 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), 2017, pp. 328–339., http://doi.org/10.1109/ICDCS.2017.226

[17] "Ray — Ray 0.2.0 documentation." [Online]. Available: http://ray.readthedocs.io/en/latest/. [Accessed: 11-Sep-2017].

[18] N. Hansen, A. Auger, O. Mersmann, T. Tusar, and D. Brockhoff, "COCO: A Platform for Comparing Continuous Optimizers in a Black-Box Setting," ArXiv160308785 Cs Stat, Mar. 2016.

[19] S. Cahon, N. Melab, and E.-G. Talbi, "ParadisEO: A Framework for the Reusable Design of Parallel and Distributed Metaheuristics," J. Heuristics, vol. 10, no. 3, pp. 357–380, May 2004., http://doi.org/ 10.1023/B:HEUR.0000026900.92269.ec

[20] D. R. White, "Software review: the ECJ toolkit," Genet. Program. Evolvable Mach., vol. 13, no. 1, pp. 65–67, Mar. 2012., http://doi.org/ 10.1007/s10710-011-9148-z

[21] "MOEA Framework, a Java library for multiobjective evolutionary algorithms." [Online]. Available: http://moeaframework.org/index.html. [Accessed: 05-Feb-2017].

[22] A. Li, X. Yang, S. Kandula, and M. Zhang, "CloudCmp: Comparing Public Cloud Providers," in Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, New York, NY, USA, 2010, pp. 1–14., http://doi.acm.org/10.1145/1879141.1879143

[23] A. S. Ganapathi, "Predicting and Optimizing System Utilization and Performance via Statistical Machine Learning," University of California at Berkeley, Berkeley, CA, USA, 2009.

[24] D. N. Hieu, T. T. Minh, T. V. Quang, B. X. Giang, and T. V. Hoai, "A Machine Learning-Based Approach for Predicting the Execution Time of CFD Applications on Cloud Computing Environment," in Future Data and Security Engineering, 2016, pp. 40–52., http://doi.org/ 10.1007/978-3-319-48057-2_3

[25] S. Venkataraman, Z. Yang, M. Franklin, B. Recht, and I. Stoica, "Ernest: Efficient Performance Prediction for Large-scale Advanced Analytics," in Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation, Berkeley, CA, USA, 2016, pp. 363–378.

[26] M. Brčić and N. Hlupić, "Simulation library for Resource Constrained Project Scheduling with uncertain activity durations," in 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014, pp. 1041–1046., http://doi.org/ 10.1109/MIPRO.2014.6859722

[27] M. Brčić, D. Kalpić, and M. Katić, "Proactive Reactive Scheduling in Resource Constrained Projects with Flexibility and Quality Robustness Requirements," Comb. Optim., pp. 112–124, Aug. 2014., http://doi.org/ 10.1007/978-3-319-09174-7_10

[28] R. Kolisch and A. Sprecher, "PSPLIB - A project scheduling problem library: OR Software - ORSEP Operations Research Software Exchange Program," Eur. J. Oper. Res., vol. 96, no. 1, pp. 205–216, Jan. 1997.

[29] M. Matsumoto and T. Nishimura, "Dynamic Creation of Pseudorandom Number Generators," presented at the Proceedings of the Third International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, 1998, pp. 56–69.

[30] J. K. Salmon, M. A. Moraes, R. O. Dror, and D. E. Shaw, "Parallel random numbers: As easy as 1, 2, 3," in High Performance Computing, Networking, Storage and Analysis (SC), 2011 International Conference for, 2011, pp. 1–12., http://doi.org/ 10.1145/2063384.2063405

[31] "Amazon Web Services (AWS) - Cloud Computing Services," Amazon Web Services, Inc. [Online]. Available: //aws.amazon.com/. [Accessed: 14-Dec-2014].

[32] "boto GitHub repository," GitHub. [Online]. Available: https://github.com/boto/boto. [Accessed: 29-Sep-2014].

[33] "Welcome to Paramiko! — Paramiko documentation." [Online]. Available: http://www.paramiko.org/. [Accessed: 25-Oct-2017].

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 145-150 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

A Relational Database Model and Tools for Environmental Sound Recognition

Yuksel Arslan*, Abdussamet Tanış, Huseyin Canbolat

Electrical and Electronics Engineering, Ankara Yıldırım Beyazıt University, 06010, Turkey

ARTICLE INFO

Article history:
Received: 29 October, 2017
Accepted: 17 November,2017
Online: 10 December, 2017

Keywords:
Database Design
Environmental Sound Recognition
Machine Learning

ABSTRACT

Environmental sound recognition (ESR) has become a hot topic in recent years. ESR is mainly based on machine learning (ML) and ML algorithms require first a training database. This database must comprise the sounds to be recognized and other related sounds. An ESR system needs the database during training, testing and in the production stage. In this paper, we present the design and pilot establishment of a database which will assists all researchers who want to establish an ESR system. This database employs relational database model which is not used for this task before. We explain in this paper design and implementation details of the database, data collection and load process. Besides we explain the tools and developed graphical user interface for a desktop application and for the WEB.

1. Introduction

This paper is an extension of work presented in 25th Signal Processing and Communications Applications Conference (SIU), 2017 [1]. The database design and implementation described in that paper was mainly for impulsive sound detection and the database was for a hazardous sound recognition application. Here the extended database is for all kinds of environmental sounds and it is aimed for all kinds of ESR applications.

Historically non-speech sound recognition has not received as much attention as automatic speech recognition (ASR). ASR has well established algorithms and databases while research has begun much earlier. Automatic ESR (AESR) is getting attraction since last two decades. We can list the following applications of AESR: In military, forensic and law enforcement domain there are studies on gunshot detection systems. In [2], a gunshot detection system is proposed. In [3], the gunshot blast is used to identify the caliber of the gun. In [4] and [5] ESR is used for robot navigation. ESR can be used for home monitoring. It can be used to assist elderly people living in their home alone [6], [7]. In [8], it is used for home automation. In [9] and [10], ESR is used for recognition of animal sounds. In the surveillance area, it is used for surveillance of road [11], public transport [12], elevator [13] and office corridor [14].designations. ESR system design is started with the training phase. How a sound database is used in training phase of an ESR system is explained in Figure 1.

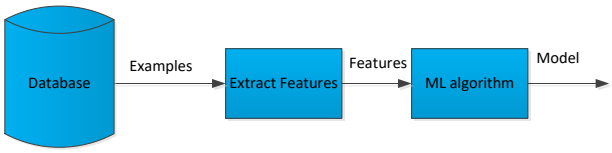


Figure 1: ESR system in training phase

During the development stage of an ESR system the desired ML algorithm must be trained with the sounds to be recognized. The database provides the sound clips to be recognized and other sound clips which are negative examples. After training, a model is developed and this model is used for testing. In Figure 1, the database provides positive and negative examples; features are inputs to the ML algorithm, ML algorithm using these features produce the model.

In Figure 2, testing phase of an ESR system is shown. In testing phase database provides the positive and negative examples, model produces the predictions about the examples provided and at last predictions are compared with the truth provided by the database and a performance rate is reached. According to this performance rate, ML algorithm or the feature set or other parameters of the ML algorithm may be needed to change. Then the whole training and testing phase start again. This process continues until an acceptable performance rate is reached.

After model creation and testing, this model is used in the production phase (Figure 3). In production phase, sounds come

*Yuksel Arslan, Ankara, Cankaya Ilkbahar Mah. Guneypark Kumeevleri A35/35 Turkey, Email: yuksel_arslan@yahoo.com

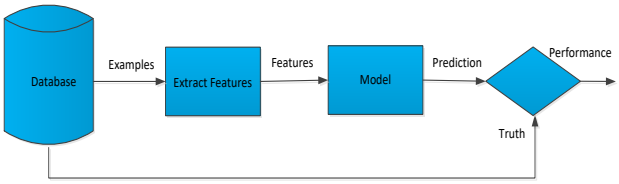


Figure 2: ESR system in testing phase

from the environment, model makes the predictions about the classes of sounds and we actually may not know the real truth. In production stage we may update sounds in database, update the model even we may change the ML algorithm. In each case we need the database.

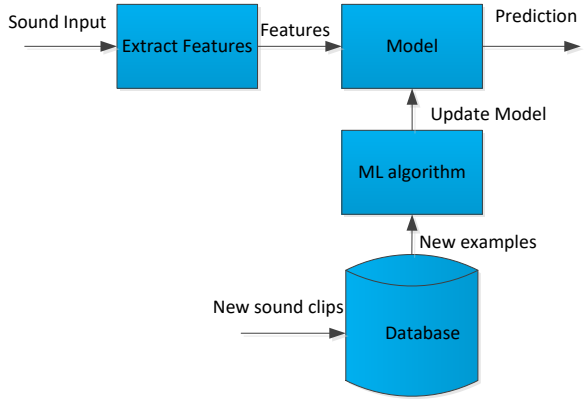


Figure 3: ESR system in production phase

This paper is organized as follows: In section II we will make a literature review of ESR databases. In section III a new relational database model for ESR systems will be explained. At the last, section IV, the contribution of this work and other planned activities will be explained.

2. Literature Review

In this section we will review the mostly referenced databases in ESR related papers. The structure of the databases will be explored; pros and cons of the structure will be argued if applicable.

One of the databases that have been mostly reverenced is Real World Computing Partnership's (RWCP) non-speech database [15]. Using a standard single microphone, acoustic signals of about 100 types of sound sources were measured in an anechoic room as the dry source. By using the impulse response of three kinds of rooms, convolution with the 48 kHz sampling signal was carried out to reconstruct the sound signal in respective rooms. The database structure (structure of directories) and navigation to sounds is explained and provided by an HTML page as shown in Figure 4 for dry sources. As an ESR developer we must examine the structure of HTML file and find the desired sound clips. Maintenance of the data and usage is not so easy with this kind of structure.

Detection and Classification of Acoustic Scenes and Events (DCASE) is an official IEEE Audio and Acousti Acoustic Signal Processing challenge. For challenge a database is prepared and it is also a resource for ESR researchers. DCASE 2016 challenge consists of four tasks [16]. The goal of Task-2 is to detect sound events for example "bird singing", "car passing by" that are present within an audio segment. To be prepared for the challenge two datasets, train and development are given. Train is used to create the model and development is to test the model.

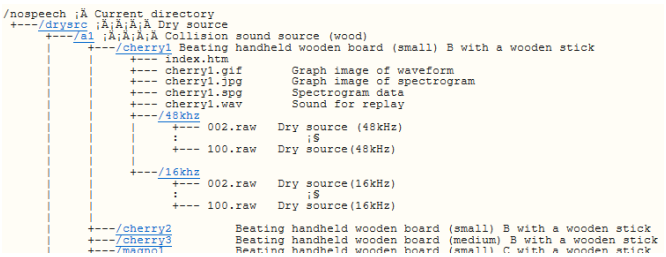


Figure 4: RWCP dry source sound clips structure

For training all sound files are in one directory and a readme file explains the details, such as sampling frequency, quantization bit depth, etc. For development dataset there is an annotation text file for each sound clip. Each .txt file contains information about the onset, offset, and event class for each event in the scene, separated by a tab.

DCASE 2017 Task-2 dataset contains ".yaml" files for annotations [17]. ".yaml" files (Figure 5) can be read by a Matlab command, so it is easy to work with these structured files.

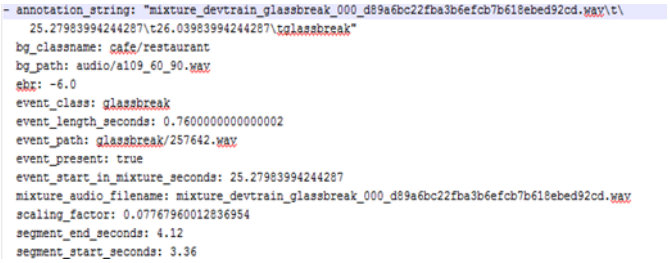


Figure 5 DCASE 2017 task-2 development dataset "yaml" file for glass break

Dataset for Environmental Sound Classification (ESC) contains two databases. There are 10 classes and each class has 40 clips in ESC-10 dataset. ESC-10 is subset of ESC-50 which contains 50 classes and each class has 40 clips. There is a readme file in which there is a line for each clip explaining the details of the clip for each dataset [18].

In [19], a dataset of annotated urban sound clips are recorded and taxonomy for urban sound sources is proposed. The dataset contains 10 sound classes with 18.5 hours of annotated sound event occurrences. The dataset contains a CSV file (Figure 6) explaining the details of each recording.

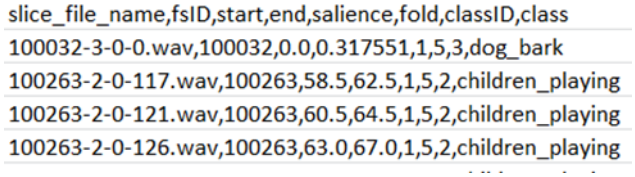


Figure 6: CSV file explaining Urbansound database

Another ESR database is http://www.desra.org. In [20], the aim of this database, details of the design and the sources used are

explained. It is designed as multi-purpose database. The database contains variety of sounds from different events with thorough background information for each sound file. It is accessible from the Internet (The database is not fully functional at the moment). The database was designed considering for admin tasks and general user level tasks. Web front end provides the functionality for user level tasks.

http://www.auditorylab.org is [21] another database. This database was constructed by Carnegie Mellon University to examine the human ability to use sounds to understand what events are happening in the environment. All the sounds in the database are recorded in a controlled way. The laboratory and the recording media used are technically detailed. In the construction of database sounds are grouped by the event that makes the sound. Sounds of events like impact, rolling can be downloaded from this database.

3. Database Design

The databases reviewed in Section 2 are prepared for just resources for the development of ESR applications. Many of them provide the data and the files for correctly handling the data. In this context they are valuable resources for all ESR researchers. Desra [20], provides extra tools such as a web graphical user interface for searching and testing.

In the development process of an ESR system, some main functions training, testing and production explained in Section I. During this development process, many sound clips are used; new sound clips may be added or deleted. We deal with many features extracted from these sound clips; we create models using different ML algorithms. Then we compare the models; try to find best features and best ML algorithms. This loop continues. Our first goal in this database design is to help researchers as much as possible to ease the burden of handling data. The second goal is to help maintenance of the data. Our data is sound clips, features, algorithms and models. Addition, deletion, searching, annotation, backup and restore can be thought as the maintenance task.

3.1 Taxonomy of Environmental Sounds

We need a taxonomy to be able to store, search and retrieve the data from the database. During literature review we see two kinds of taxonomy. In the first taxonomy, the sounds can be grouped by the event that makes the sound. In Figure 7, grouping of sounds defined in [21] is seen.

title	author	type	modified
LEGAL NOTICE	Auditory Lab User	Page	2009-06-24 21:16
Impact Events	Auditory Lab User	Folder	2009-06-25 10:15
Deformation Events	Auditory Lab User	Folder	2009-06-25 10:16
Rolling Events	Auditory Lab User	Folder	2009-06-25 10:17
Air Events	Auditory Lab User	Folder	2009-06-25 10:17
Liquid Events	Auditory Lab User	Folder	2009-06-25 10:18
Experiment Stimuli	Auditory Lab User	Folder	2009-06-25 10:19

Figure 7: Grouping of sounds based on sound producing events [21]

Another hierarchical grouping is seen in Figure 8. This is also classified in the first taxonomy. This is the grouping defined in [22] based on sound producing events and the listeners' description of the sound. Second taxonomy is based on the sound source [19][23]. The subset of the taxonomy defined in [23] is taken into consideration for urban sounds given in [19].

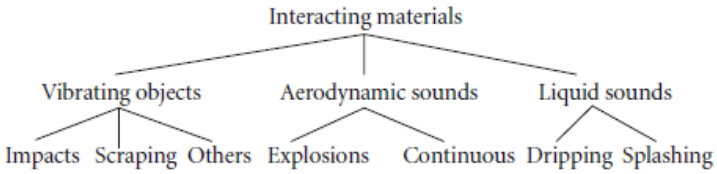


Figure 8: Grouping of sounds defined in [22]

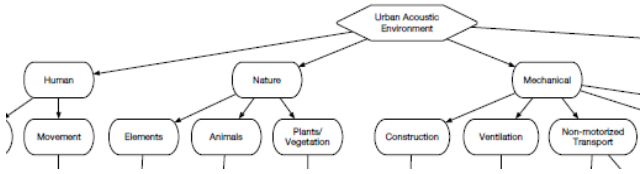


Figure 9 A part of the urban sound taxonomy defined in [19]

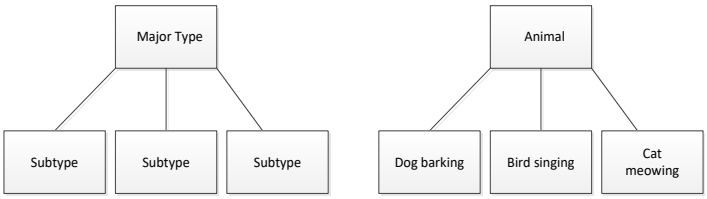


Figure 10: Environmental sounds grouping in the database

3.2 Non Functional Requirements of the Database

- There will be sound clips in the database. These sound clips will contain environmental sounds. Each sound clip can have more than one environmental sound. These environmental sound clips can belong to different major environmental sound types. Origin of these sound clips, such as own recording, from an internet site or from another database, should be entered to database. Recording details should be entered. Size, file type of sound clip and the path where the sound clips are recorded should be entered to the database.
- There should be major types, such as human, nature, animal, etc. There should be subtypes, such as dog bark, gunshot etc. and these must belong to major type. Start and end sample index or start/end time of these clips should be known.
- There will be environmental sound clips in the database. These clips should be extracted from sound clips defined in first paragraph. Each of these sound clips should have a subtype and it should be known from which sound clips it

is extracted. Extraction method should be entered. If it is extracted by an algorithm the algorithm name otherwise as "manual" should be entered to database.

- Recording details should be entered to the database such as, sampling rate, quantization bit, channel size, etc.
- Background clips should be entered to the database. Each background clip should have a type. Each background clip should also have file type, file size, recording detail and the path where it is recorded.

3.3 Functional Requirements

- There should be scripts which will take algorithms as arguments and extract the environmental sounds from sound clips. These extracted clips should be entered to the database as explained in 3.2.
- There should be scripts which will embed environmental sound into noise clips at desired Signal to Noise Ratio (SNR) level.
- There should be scripts which will extract features from environmental sound clips and store to the database.
- The scripts, their help files, paths should be stored in the database.

3.4 Graphical User Interface (GUI) Requirements

- Administrators can use the GUI for meta data entrance, deletion and update such as major types, subtypes, recording details.
- Administrators can use the GUI for data load manually or using the data loader script.
- Researchers can use the GUI for searching and downloading the desired environmental sounds.
- Researchers can load their features and models to the database.

3.5 Database Implementation

The implemented database will include some data which can be stored by way of data types found in the standard database software and also sound files. These sound files will not be stored to database instead they will be stored in the operating system file structure and the path to this file is just recorded in the database.

Microsoft SQL 2008 Database Server is used to create tables. The tables and the relationships between them are shown in Figure 11. The scripts providing functionality are coded in Matlab 2011a. The files are stored in "mat" type when required.

For functional requirements the following scripts are coded using Matlab.

- Environmental Sound Detector: This script is an interface between algorithms that extracts environmental sounds from the sound clips. Different algorithms can be used here for extraction of the environmental sounds. The algorithms must conform to this script interface definition. This script takes the algorithm name that will be started and the environmental sound types which will be searched are given as arguments.
- Environmental Sound Embedder: This script takes the type of the environmental sound, noise type, SNR level and at last the number of required record count to be created. The script merges the environmental sound clip with the noise clip at the desired SNR level and records it to the database table.
- Feature Extractor: This script acts as an interface between feature extractors. It takes the path of the feature extractor from the features table, feature name and the environmental sound type from command line. After extracting features, it is saved as a mat file.
- Data Loader: To load data from other databases this script is used to interface with data loading scripts.

The GUI is developed using Microsoft Visual Studio with C# language. GUI has admin utilities and end user tools. In Figure 12(a), it is seen the part of GUI providing admin operations. These admin operations are additions and deletions of major types, subtypes and sampling information. Besides the GUI provides the administrators load data one by one or using a script to load as a batch.

The GUI provides the users some tools. The tools are for searching the database and testing their algorithms on the sound clips given by the database. Users can search the database to see the sound clips with desired type.

End users can search the database with noise clip type, sound clip type and with SNR interval then select a clip and copy it to their own computer. After finding the start index of the embedded environment sound clip, by writing the start index of this environmental sound to the edit box on the GUI and by clicking the test button, they can test their algorithm correctness. The WEB GUI for end user tools is seen in Figure 12(b).

After database and scripts implementation we loaded the data from Urbansound [19] database. Now many operations on the database can be done either using GUI or an SQL command line.

4. Conclusion

The lack of common database for environmental sound recognition is an important obstacle in front of the researchers. The development of and ESR system is a tough process during which someone have to deal with lots of sound clips with different types, algorithms and models. In this paper we explained a relational database model which will make the data handling easier. The database developed is different from other counterparts which are just providing the data. The relational database model described here provides easy maintenance as well as easy usage.

Although our goal of designing this database is for ESR, other areas that deal with environmental sounds can use it.

By improving the database by adding data from general databases mostly used and by adding more functionality we aim it to be a common database for research activities of ESR.

Conflict of Interest

The authors declare no conflict of interest.

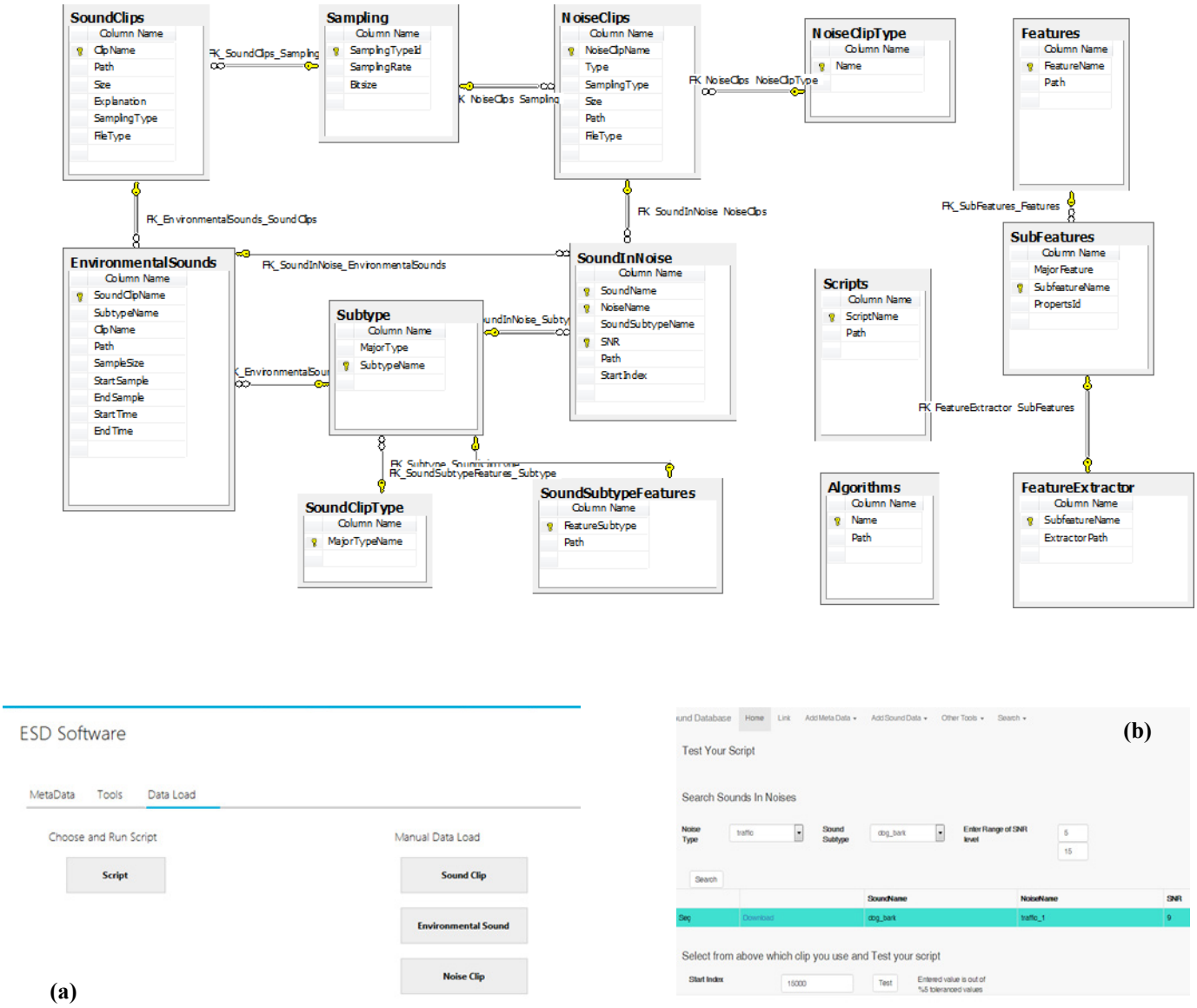


Figure 12 (a) Desktop application GUI for administrators to data load (b) WEB GUI for end user tools

References

[1] Y.Arslan and H. Canbolat, "A sound database development for environmental sound recognition", Signal Processing and Communications Applications Conference (SIU), 25th, 2017.

[2] T. Ahmed, M. Uppal and A. Muhammad, "Improving Efficiency and Realibility of Gunshot Detection Systems", IEEE, ICASSP 2013.

[3] P. Thumwarin, T. Matsuura and K. Yakoompai, "Audio forensics from gunshot for firearm identification", Proc. IEEE 4th Joint International Conference on Information and Communication Technology Electronic and Electrical Engineering Tailand, pp. 1-4, 2014.

[4] Chu, S.;Narayanan, S.; Kuo, C.-C.J.;Mataric, M.J.:Where am I? Scene recognition for mobile robots using audio features, in 2006 IEEE Int.Conf. on Multimedia and Expo. IEEE, 885–888, 2006.

[5] Yamakawa, N.; Takahashi, T.; Kitahara, T.; Ogata, T.; Okuno,H.G.: Environmental sound recognition for robot audition using Matching-Pursuit, in Modern Approaches in Applied Intelligence, in K.G. Mehrotra, C.K. Mohan, J.C. Oh, P.K. Varshney & M. Ali (Eds), Springer Berlin Heidelberg, 1–10, 2011.

[6] Chen, J.; Kam, A.H., Zhang, J.; Liu, N.; Shue, L.: Bathroom activity monitoring based on sound, in Pervasive Computing, in H.W. Gellersen, R.Want, & A. Schmidt (Eds), Springer Berlin Heidelberg, 47–61, 2005.

[7] Vacher, M.; Portet, F.; Fleury, A.;Noury, N.: Challenges in the processing of audio channels for ambient assisted living, in 2010 12th IEEE Int. Conf. on e-Health Networking Applications and Services (Healthcom), IEEE, 330–337, 2010.

[8] Wang, J.-C.; Lee, H.-P.; Wang, J.-F.; Lin, C.-B.: Robust environmental sound recognition for home automation. Automation Science and Engineering, IEEE Transactions on, 5 (1) (2008), 25–31.

[9] Bardeli, R.;Wolff,D.; Kurth, F.; Koch,M.; Tauchert, K.-H.; Frommolt, K.-H.: Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. Pattern Recognit. Lett., 31 (12) (2010), 1524–1534.

[10] Weninger, F.; Schuller, B.; Audio recognition in the wild: static and dynamic classification on a real-world database of animal vocalizations. in 2011 IEEE Int.Conf. on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2011, 337–340.

[11] P. Foggia, A. Saggese, N. Strisciuglio, M. Vento, and N. Petkov. Car crashes detection by audio analysis in crowded roads. In Advanced Video and Signal Based Surveillance e (AVSS), 2015 12th IEEE International Conference on, pages 1-6, Aug 2015.

[12] J. Rouas, J. Louradour, and S. Ambellouis, "Audio Events Detection in Public Transport Vehicle," Proc. of the 9th International IEEE Conference on Intelligent Transportation Systems, 2006.

[13] R. Radhakrishnan and A. Divakaran, "Systematic acquisition of audio classes for elevator surveillance," in Image and Video Communications and Processing 2005, vol. 5685 of Proceedings of SPIE, pp. 64–71, March 2005.

[14] P. K. Atrey, N. C. Maddage, and M. S. Kankanhalli, "Audio based event detection for multimedia surveillance," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '06), vol. 5, pp. 813–816, Toulouse, France, May 2006.

[15] Nakamura, S.; Hiyane, K.; Asano, F.; Nishiura, T.; Yamada, T.: Acoustical sound database in real environments for sound scene understanding and hands-free speech recognition, in LREC, 2000.

[16] DCASE 2016 WEB site available at: http://www.cs.tut.fi/sgn/arg/dcase2016/ (27.09.2017 last accessed)

[17] DCASE 2017 WEB site available at: http://www.cs.tut.fi/sgn/arg/dcase2017/ (27.09.2017 last accessed)

[18] K. J. Piczak, "ESC: Dataset for environmental sound classification", Proceedings of the ACM International Conference on Multimedia, 2015.

[19] Justin Salamon , Christopher Jacoby , Juan Pablo Bello, A Dataset and Taxonomy for Urban Sound Research, Proceedings of the 22nd ACM international conference on Multimedia, November 03-07, 2014.

[20] Brian Gygi and Valeriy Shafiro, "Development of the Database for Environmental Sound Research and Aplication (DESRA):Design, Fuctionality and Retrieval Considerations", EURASIP Journal on Audio, Speech, and Music Processing Volume 2010.

[21] http://www.auditorylab.org (28.09.2017 last accessed)

[22] W. W. Gaver, "What in the world do we hear? An ecological approach to auditory event perception," Ecological Psychology, vol. 5, no. 1, pp. 1–29, 1993.

[23] A. L. Brown, J. Kang, and T. Gjestland. Towards standardization in soundscape preference assessment. Applied Acoustics, 72(6):387-392, 2011

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 151-156 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Analysis of Economic Load Dispatch with a lot of Constraints Using Vortex Search Algorithm

Mustafa Saka[*,1], Ibrahim Eke[2], Suleyman Sungur Tezcan[1], Muslum Cengiz Taplamacioglu[1]

[1]*Gazi University, Electrical and Electronics Engineering Department, 06500, Turkey*

[2]*Kirikkale University, Electrical and Electronics Engineering Department, 71450, Turkey*

ARTICLE INFO

Article history:
Received: 31 October, 2017
Accepted: 17 November, 2017
Online: 11 December, 2017

Keywords:
Economic load dispatch
Valve point effect
Ramp rate limits
Prohibited zone
Transmission losses
Optimization
Vortex search algorithm

ABSTRACT

In modern powers systems, one of the most considerable topics is economic load dispatch (ELD). ELD is non-linear problem and it became non-convex and non-smooth problem with some constraints such as valve point loading effect. ELD is very crucial for energy generation and distribution in power systems. For solving ELD problem, a lot of methods were developed and used at different power systems. Vortex search algorithm (VSA) is proposed and applied for solving ELD problem in this paper. VSA method was developed in the form of stirring liquids. Transmission line losses, valve point loading effect, ramp rate limits and prohibited zones constraints were used to make the results of ELD problem the closest to the truth. The results which are obtained from VSA compared with PSO, CPSO, WIPSO, MFO, GA and MRPSO techniques. It can be clearly seen that VSA gave minimum cost values with optimum generator powers so it is very effective and useful method and it gave the best solutions for ELD.

1. Introduction

This paper is an extension of work originally presented in 4th International Conference on Electrical and Electronics Engineering [1]. The purpose of this work is solving economic load dispatch problem with a lot of constraints through with new optimization technique Vortex Search Algorithm.

Economic load dispatch of a power system is very important in terms of control and planning of that power system. Main goal of ELD is distributed total demand power among the committed thermal generation units with minimum production cost by satisfying set of equality and inequality constraints. If ELD problem is not solved for thermal power plants, demand power may be generated very costly. ELD problem can be basically modeled second order (quadratic) function [2]. However, this function may became more complex, non-smooth and non-convex with some constraints such as valve point loading effect, ramp rate limits, transmission line losses and prohibited zones.

Economic load dispatch plays very big role for operated power plants. For this reason, a lot of researcher studied this issue. A number of optimization techniques developed and applied to ELD problem. Quadratic Programming [3], Linear Programming [4], Non-Linear Programming [5], Lambda Iteration Method [6] etc. techniques were used for solving ELD problem. These traditionally techniques gives good results for basic ELD problem but these techniques may poor results when constraints and complexity are increased.

Together with the advances in the computer sciences, a lot of random search optimization techniques developed and used [7]. Different Evolution [8], Particle Swarm Optimization [9], Genetic Algorithm [10], Artificial Bee Colony [11], Harmony Search [12], Bacterial Foraging Optimization [13], Firefly Algorithm [14], Ant Colony Optimization [15] etc. are some of these techniques. User defined parameters are necessary for these optimization techniques. If the parameters are not chosen properly, the results obtained from these techniques may not be good results.

The organization of this paper as follows: Economic load dispatch, main objective of ELD, constraints of ELD and mathematical express of ELD are described in Section 2. Vortex Search Algorithm and its mathematical model are described in Section 3. Using test system, its parameters, obtained results and

[*]Corresponding Author: Mustafa SAKA, Gazi University, Graduate School of Naturel and Applied Sciences, Ankara, TURKEY, msaka@gazi.edu.tr

figures are described in Section 4. Finally evaluation of this paper is briefed in Section 5.

2. Economic Load Dispatch

There are a lot of operating cost for thermal power plants such as fuel cost, personal fees etc.. In these costs the biggest share is fuel cost of thermal generation units. For this reason solving economic load dispatch problem for thermal power plants is necessity. Main objective of economic load dispatch is keep the total fuel cost as minimum level while meet the total demand power. Basically defined cost function of ELD as a quadratic function as follows:

$$F_i = a_i + b_i \times P_i + c_i \times P_i^2 \quad (1)$$

2.1. Valve Point Loading Effect

Due to opening stream valves at the power systems losses are increased. This effect is called valve point loading effect. Due to system losses are increased wit valve point loading effect, total cost value is increased. Above equation has sinusoidal terms due to valve point loading effect. This situation can be seen in Figure 1. This transformed equation is expressed as follows:

$$F_i = a_i + b_i \times P_i + c_i \times P_i^2 + |e_i \times \sin(f_i \times (P_i^{min} - P_i)) \quad (2)$$

F_i represents resulting fuel cost value, P_i represent power of thermal generator, a_i, b_i, c_i are cost coefficients and e_i, f_i are valve point loading coefficients of thermal generator unit i.

Total cost value of system is obtained by summed cost values of every thermal generation units.

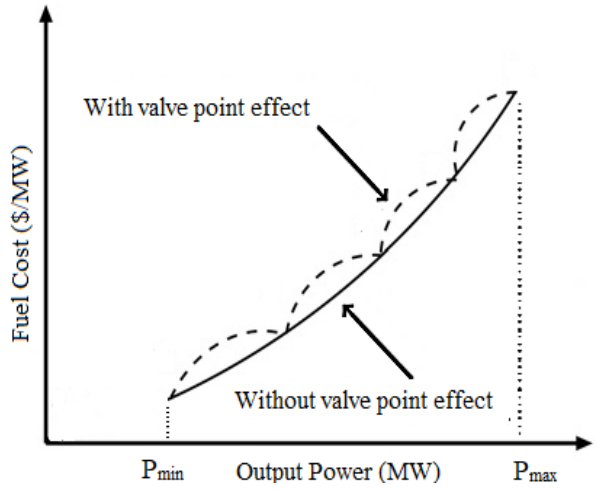


Figure 1 With and without valve point effect

2.2. Generators Limits

Thermal generators units must operate maximum and minimum power range. This power range can be different for different units:

$$P_{i,max} \geq P_i \geq P_{i,min} \quad (3)$$

Pi,min and Pi,max are represent minimum generator limit and maximum generator limit of unit i.

2.3. Power Balance

Total generated power at the thermal power plants meet the demand power by consumers. For this reason transmission line losses must be considered. The total generated power obtained sum of total demand power and total transmission line losses.

Transmission line losses is calculated as follows:

$$P_{loss} = \sum_{i=1}^{N} \sum_{j=1}^{N} P_i B_{ij} P_j + \sum_{i=1}^{N} B_{0i} P_i + B_{00} \quad (4)$$

Total generated power is calculated as follows:

$$\sum_{i=1}^{N} P_i = P_d + P_{loss} \quad (5)$$

N represent total thermal power plant. Pd and Ploss represent total demand power by consumer and total transmission line losses respectively. Bij, B0i and B00 are transmission loss coefficients.

2.4. Ramp Rate Limits

While a thermal generator unit is operating at a certain point, the operating point can only be increased to a certain level determined by the up ramp rate limit or decreased to a certain level determined by the down ramp rate limit.

This situation is shown as follows:

$$\max(P_i^{min}, P_i^0 - DR_i) \leq Pi \leq \min(P_i^{max}, P_i^0 + UR_i) \quad (6)$$

URi represents the up-ramp rate limit, DRi represent down ramp rate limit, is previous generated power and Pi present generated power of unit i.

2.5. Prohibited Zone

In some cases, thermal generation units do not worked and do not want to be worked some reasons such as mechanical corruption some particular power range (prohibited zone). These conditions can be expressed as follows:

$$P_i^{min} \leq Pi \leq P_{i,1}^{lowerbound} \quad (7)$$

$$P_{i,j-1}^{upperbound} \leq P_i \leq P_j^{lowerbound} \quad (8)$$

$$P_{i,ni}^{upperbound} \leq Pi \leq P_i^{max} \quad (9)$$

j represents number of prohibited zones of unit i and j=2, 3, 4 …ni. represents lower limit and represents upper limit of jth prohibited zone. Prohibited zone effect is shown in Figure 2.

3. Vortex Search Algorithm

Vortex search algorithm is a new optimization technique and inspired by stirring liquid materials.[17,18]. VSA is very influential and handy technique for solving economic load dispatch problem. problem. As a result of using an extensible step size modification arrangement a good balancing explorative and exploitative behavior of the search are obtained [18]. Vortex patterns can be represented in two dimensional space as a lot of nested circuits. The biggest and outer circuit is starting circuit of search space. The center of this circuit is calculated as follows:

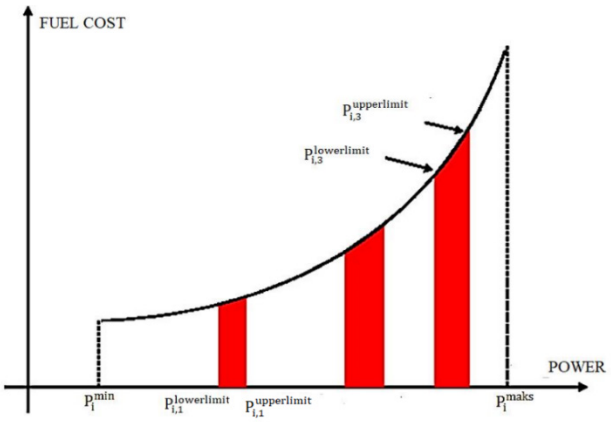


Figure 2 Shown of prohibited zones [16]

Upperlimit and lowerlimit are maximum and minimum constraints of the problem. A lot of candidate solutions are constituted into the outer circuit. The initial standard deviation is accepted as radius of this outer circuit and calculated as follows:

$$\mu_0 = (\text{upperlimit} + \text{lowerlimit}) / 2 \quad (10)$$

$$\sigma_0 = (\max(\text{upperlimit}) - \min(\text{lowerlimit})) / 2 \quad (11)$$

Generated candidate solutions must be controlled for ensure these solutions within the maximum and minimum limits of the problem. If these candidate solutions are not into the search space, they must be shifted into the search space. Some candidate solutions, which are not into the search space, shifted into the search space using equation as follows:

$$s_i^k = \text{lowerlimit}^i + (\text{upperlimit}^i - \text{lowerlimit}^i) \times \text{rand} \quad (12)$$

After this step a candidate solution, which is the best solution into the search space, is selected. This best solution is represented s'. This solution is memorized and current center μ0 is shifted to s'. New candidate solutions are constituted around this new center. All of candidate solutions are compared with previous best solution. If there is a better solution than s', it is selected as a new best solution [19]. This situation continues each iteration step. Radius of the circuits must be decreased every iteration. For this reason inverse gamma function (gammaincinv) is used and new radius is calculated every iteration as follows:

$$rt = \sigma 0 \times (1 / x) \times \text{gammaincinv}(x, a_t) \quad (13)$$

a_t is changeable parameter and it dependent a_0, t and MaxItr. This function shown as follows:

$$a_t = a_0 - (t / \text{MaxItr}) \quad (14)$$

Where t represents iteration number, MaxItr represents maximum iteration number. Due to cover the all search space a_0 is chosen 1. Outline flowchart of VSA method is shown in Figure 3:

4. Test System and Results

Vortex search algorithm was proposed and used for solving economic load dispatch problem. Six generation unit power system was selected. VSA algorithm was applied to this system considering various constraints. Each cases the number of iteration is limited to 1000 for obtained good solutions.

Case1 constraints are ramp rate limits and transmission losses;

Case2 constraints are ramp rate limits, prohibited zones and transmission losses;

Case3 constraints are transmission losses, valve point loading effect and ramp rate limits;

For every case transmission line losses coefficients are same and given as follows:

$B=10^{-4} \times$

0.17	0.12	0.07	-0.01	-0.05	-0.02
0.12	0.14	0.09	0.01	-0.06	-0.01
0.07	0.09	0.31	0.01	-0.10	-0.06
-0.01	0.01	0.00	0.24	-0.06	-0.08
-0.05	-0.06	-0.10	-0.06	1.29	-0.02
-0.02	-0.01	-0.06	-0.08	-0.02	1.5

$B_0=10^{-4} \times$

-3.9080	-1.2970	7.0470	0.5910	2.1610	-6.6350

$B_{00}=0.056$

4.1. Case 1: ELD with Ramp Rate

For this case two different constraints, which are ramp rate limits and transmission line losses, applied to system. Cost coefficients, ramp rate limits data and maximum – minimum limits of generators are shown in Table 1. Results of VSA was given Table 2 with different optimization techniques. Convergence behavior of VSA for this case was shown Figure 4. For this case selected power is 1263MW.

Table 1 Data of Case 1

Unit	a	b	c	P^0	UR	DR	Pmin	Pmax
1	240	7	0.0070	440	80	120	100	500
2	200	10	0.0095	170	50	90	50	200
3	220	8.5	0.0090	200	65	100	80	300
4	200	11	0.0090	150	50	90	50	150
5	220	10.5	0.0080	190	50	90	50	200
6	190	12	0.0075	150	50	90	50	120

Table 2 Results for Case 1

P(MW)	VSA	PSO	CPSO	WIPSO
P1	457.0630	493.24	471.66	454.39
P2	172.3751	114.63	140.03	164.279
P3	264.3900	263.41	240.06	264.223
P4	141.4373	139.71	149.97	123.21
P5	164.0545	179.65	173.78	167.22
P6	76.1690	84.83	99.97	120.00
P_{loss}	12.4889	12.22	12.38	12.24
Cost($)	15448	15489	15481.87	15453.13

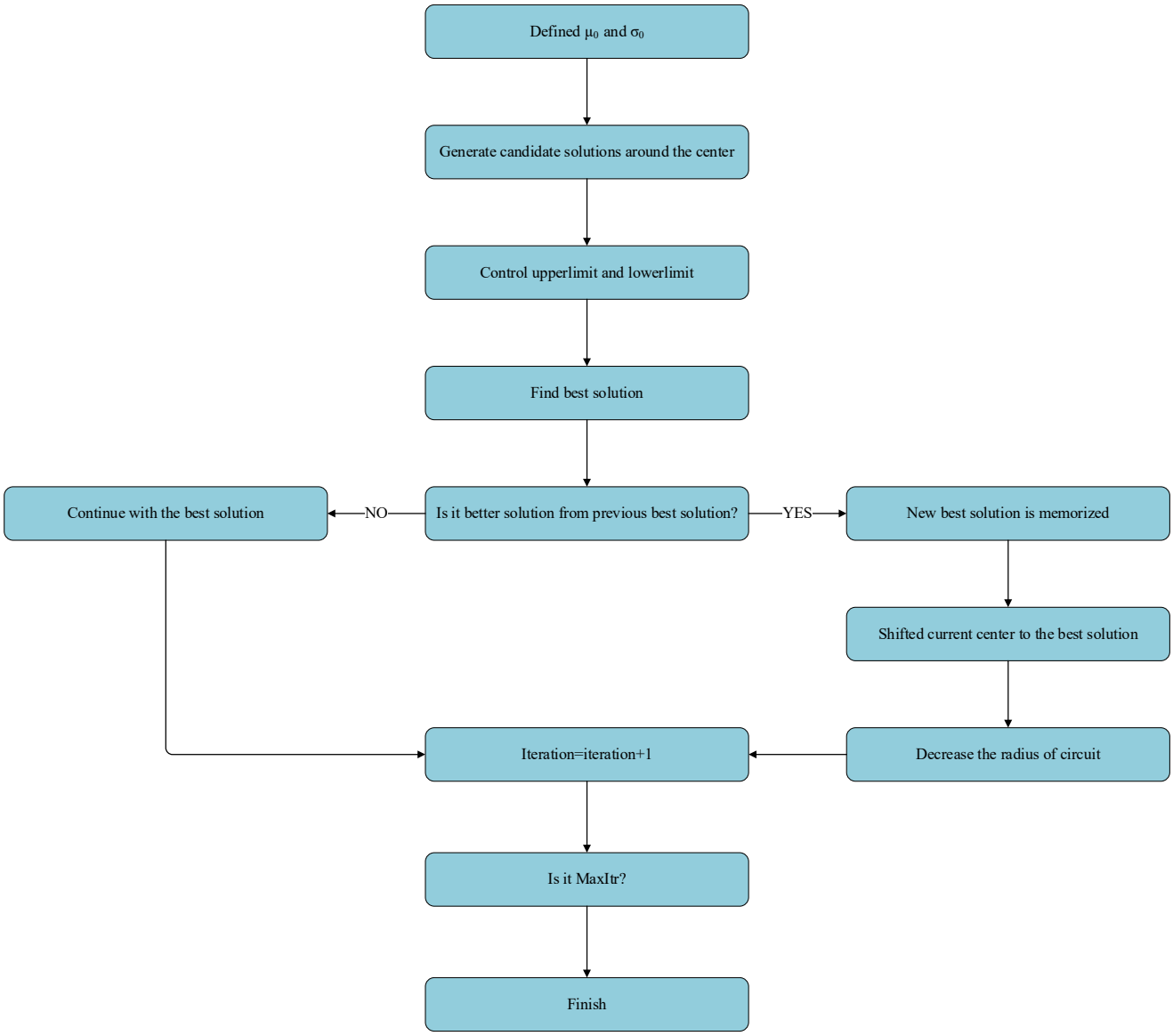


Figure 3 Outline flowchart of VSA

Results of vortex search algorithm compared with PSO, CPSO and WIPSO techniques from [20]. It is clear that from the Table 2 proposed VSA method has capable of the finding best solutions and minimum cost value. Figure 4 is shown convergence behavior of VSA for this case.

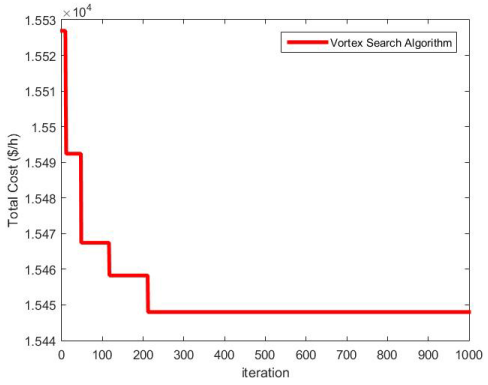


Figure 4 Convergence behavior of VSA for Case1

For this case three different constraints, which are ramp rate limits, prohibited zones and transmission line losses, applied to system. Cost coefficients was given in Table 3. Prohibited zone ranges and maximum – minimum limits of generators were given in Table 4. Ramp rate limits was given in Table 5. Convergence behavior of VSA for this case was shown Figure 5. 1263 MW power was selected for demand power. Results of VSA and other techniques results are shown in Table 6.

Table 3 Cost Coefficients

Unit	a	b	c
1	240	7	0.0070
2	200	10	0.0095
3	220	8.5	0.0090
4	200	11	0.0090
5	220	10.5	0.0080
6	190	12	0.0075

Table 4 Limits and Prohibited Zones of Generators

Unit	Pmax	Pmin	Prohibited Zone
1	500	100	210-240, 350-380
2	200	50	90-110, 140-160
3	300	80	150-170, 210-240
4	150	50	80-90, 110-120
5	200	50	90-110, 140-150
6	120	50	75-85, 100-105

Table 5 Ramp Rate Limits

Unit	P^0	UR	DR
1	440	80	120
2	170	50	90
3	200	65	100
4	150	50	90
5	190	50	90
6	150	50	90

Table 6 Results for Case 2

P(MW)	VSA	MFO	PSO	GA
P1	446.03	426.08	447.49	474.80
P2	181.09	199.80	173.32	178.63
P3	263.45	247.49	263.47	262.20
P4	133.96	136.94	139.05	134.28
P5	176.65	166.24	165.47	151.90
P6	74.53	98.93	87.12	74.18
P_{loss}	12.73	12.51	12.95	13.02
Cost($)	15447	15448.7	15450	15459

According to given results from Table 6, proposed VSA method has better cost value when compared with the MFO, PSO, GA techniques from [21].

Figure 5 is shown convergence behavior of VSA for this case.

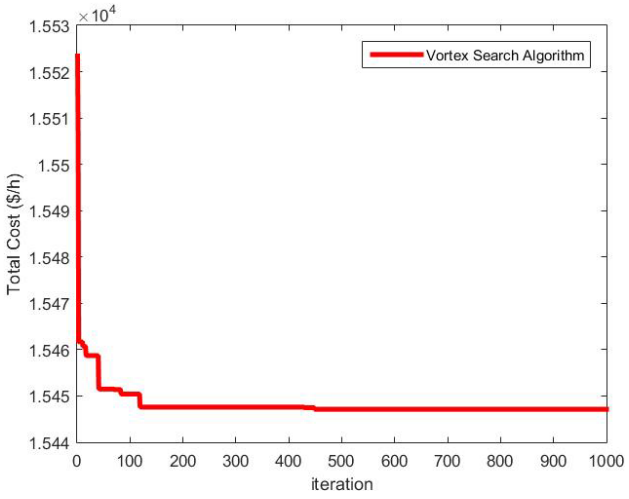


Figure 5 Convergence behavior of VSA for Case 2

4.3. Case 3: ELD with Valve Point Loading Effect

For this case three different constraints, which are transmission line losses, valve point loading effect and ramp rate limits, applied to system. Cost coefficients and valve point loading effect coefficients are given in Table 7. Obtained results were given in Table8 and compared with different methods. Convergence behavior of VSA for this case was shown Figure 6. 1263 MW power was selected for demand power.

Results of VSA and other techniques results was shown in Table 8. Maximum and minimum limits of generators are same with Table 4 and ramp rate limits are same with Table 5. Demand power is selected 1263 MW for this case.

Table 7 Valve point effect and cost coefficients

Unit	a	b	c	e	f
1	240	7	0.0070	300	0.031
2	200	10	0.0095	150	0.063
3	220	8.5	0.0090	200	0.042
4	200	11	0.0090	100	0.08
5	220	10.5	0.0080	150	0.063
6	190	12	0.0075	100	0.084

Table 8 Results for Case 3

P(MW)	VSA	PSO	CPSO	WIPSO	MRPSO
P1	495.29	443.03	467.55	437.82	442.07
P2	195.90	169.03	163.05	173.28	167.23
P3	235.79	262.02	253.41	271.97	267.09
P4	65.62	134.78	115.07	138.70	132.81
P5	197.82	147.47	169.45	146.98	155.02
P6	87.27	125.35	113.24	103.63	107.02
P_{loss}	12.92	18.68	18.70	18.08	18.03
Cost($)	15746	16372.9	16329.2	16327	16310.76

It is obviously seen that from the Table 8 VSA, method gave best cost value when compared with from PSO, CPSO, WIPSO and MRPSO techniques from [22].

Figure 6 is shown convergence behavior of VSA for this case

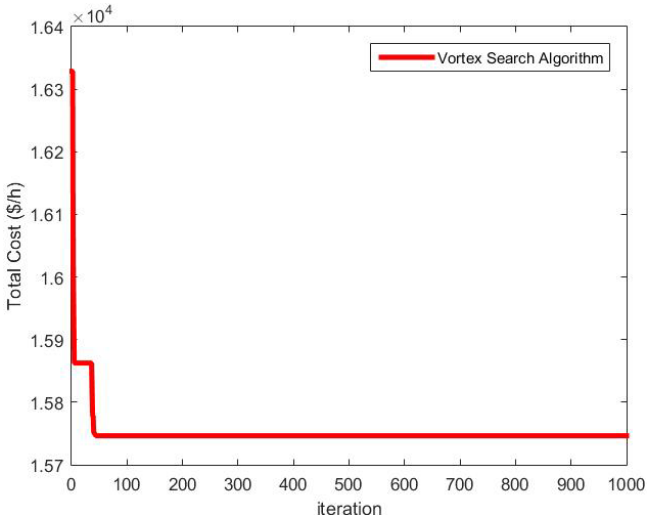


Figure 6 Convergence behavior of VSA for Case 3

5. Conclusion

In this paper vortex search algorithm was applied to six generator power system for solving economic load dispatch problem. ELD problem was became more complex and more difficult problem with valve point loading effect, ramp rate limits, transmission line losses and prohibited zones constraints. Three different situation were analyzed and for first case minimum cost value was found 15448 $, for second case minimum cost value was found 15447 $, for third case minimum cost value was found 15746 $. Obtained results from VSA compared with another techniques from the literature. These results clearly show that VSA is very capable,

feasible and effective method for solving non smooth and very complex economic load dispatch problem.

References

[1] M. Saka, I. EKE, S. S. Tezcan, M. C. Taplamacioglu., "Economic Load Dispatch Using Vortex Search Algorithm" in 4th International Conference on Electrical and Electronics Engineering, Ankara Turkey, 2017.

[2] T. Sen. H. D. Mathur, "A New Approach to Solve Economic Dispatch Problem Using A Hybrid ACO–ABC–HS Optimization Algorithm", Electrical Power and Energy Systems, 78, 735-744, 2016.

[3] L. S. Coelho, V. C. Mariani, "Combining of Chaotic Differential Evolution and Quadratic Programming for Economic Dispatch Optimization with Valve-Point Effect". IEEE Trans. Power Syst., 21, 989-996, 2006.

[4] A. Farag, S. Al-Baiyat, T. C. Cheng, "Economic Load Dispatch Multi Objective Optimization Procedures Using Linear Programming Techniques" IEEE Trans. Power Syst., 10, 731-738, 1995.

[5] L. Papageorgiou, E. Fraga, "A Mixed Integer Quadratic Programming Formulation for the Economic Dispatch of Generators with Prohibited Operating Zones" Electr. Power Syst. Res., 77, 1292-1296, 2007.

[6] C. L. Chen, S. C. Wang, "Branch-and-Bound Scheduling for Thermal Generating Units", IEEE Transactions on Energy Conversion, 8(2), 184–189, 1993.

[7] P. Bhui, N. Senroy, "A Unified Method for Economic Dispatch with Valve Point Effects" in IEEE Int. Con. Power Systems, 1-5, 2016.

[8] R. Balamurugan, S. Subramanian, "Differential Evolution-Based Dynamic Economic Dispatch of Generating Units with Valve-Point Effects," Electric Power Components and Systems, 36, 828-843, 2008.

[9] P. Sriyanyong, "An Enhanced Particle Swarm Optimization for Dynamic Economic Dispatch Problem Considering Valve-Point Loading," in Proc. the Forth IASTED International Conference on Power and Energy Systems (AsiaPES 2008), 167-172, 2008.

[10] C. L. Chiang, "Genetic-Based Algorithm for Power Economic Load Dispatch" IET Gener. Trans. Distrib., 1, 261-269, 2007.

[11] S. Hemamalini, S. P. Simon, "Artificial Bee Colony Algorithm for Economic Load Dispatch Problem with Non-Smooth Cost Functions" Electr. Power Compo. Syst., 38, 786-803, 2010.

[12] Y. C. Liang, J. R. C. Juarez, "A Normalization Method for Solving the Combined Economic and Emission Dispatch Problem with Meta-Heuristic Algorithms", Int. J. Electr. Power Energy Syst., 54, 163-186, 2014.

[13] R. Vijay, "Intelligent Bacterial Foraging Optimization Technique to Economic Load Dispatch Problem", Int. J. Soft Comput. Eng., 2, 55-59, 2012.

[14] X. S. Yang, S. S. S. Hosseini, A. H. Gandomi, "Firefly Algorithm for Solving Non-Convex Economic Dispatch Problems with Valve Loading Effect", Appl. Soft Comput., 12, 1180-1186, 2012.

[15] S. Pothiya, I. Ngamroo, W. Kongprawechnon, "Ant Colony Optimisation for Economic Dispatch Problem with Non-Smooth Cost Functions", Int. J. Electr. Power Energy Syst., 32, 478-487, 2010.

[16] M. Saka, "Economic Load and Emission Dispatch Analysis at Power Systems", M. Sc. Thesis, Gazi University, Ankara, 2017.

[17] B. Dogan, T. Olmez, "A New Metaheuristic for Numerical Function Optimization: Vortex Search Algorithm", Information Sciences, 293, 125-145, 2015.

[18] B. Dogan B., T. Olmez, "Vortex Search Algorithm for the Analog Active Filter Component Selection Problem", International Journal of Electronics and Comminications, 69, 1243-1253, 2015.

[19] M. Saka, I. Eke, S. S. Tezcan, M. C. Taplamacioglu, "Application of Vortex Search Algorithm to Economic Load Dispatch Including Transmission Losses with Valve-Point Loading Effect" in 1st International Mediterranean Science and Engineering Congress, Adana Turkey, 2016.

[20] N. Singh, Y. Kumar, "Economic Load Dispatch Using MRPSO with Generator Ramp Rate Limits Constraint" in Int. Con. Comp. Int. and Com. Networks, 632-637, 2012.

[21] N. Trivedi, A. H. Ranpariya, A. Kumar, P. Jangir, "Economic Load Dispatch Problem with Ramp Rate Limits and Prohibited Operating Zones Solve using Levy Flight Moth-Flame Optimizer" in Int. Conf. Energy Efficient Technologies for Sustainability, 442-447, 2016.

[22] N. Singh, Y. Kumar, "Economic Load Dispatch with Valve Point Loading Effect and Generator Ramp Rate Limits Constraint Using MRPSO", International Journal of Advanced Research in Computer Engineering & Technology, 2, 1472-1477, 2013.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 157-161 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Numerical evaluation of ABS parts fabricated by fused deposition modeling and vapor smoothing

Sung-Uk Zhang*

Electro Ceramic Center, Dept. of Automotive engineering, Dong-Eui University, Busan, 47340, South Korea

ARTICLE INFO

Article history:
Received: 26 October, 2017
Accepted: 01 December, 2017
Online: 10 December, 2017

Keywords:
Additive manufacturing
Fused deposition modeling
ABS
Vapor smoothing
Finite element analysis
Response surface analysis

ABSTRACT

The automotive industry has focused to use polymer materials in order to increase energy efficiency. So, the industry pays attention to use 3D printing technologies using several polymers. Among several 3D printer technologies, fused deposition modeling (FDM) is one of the popular 3D printing technologies due to an inexpensive extrusion machine and multi-material printing. FDM could use thermoplastics such as ABS, PLA, ULTEM so on. However, it has a problem related to the post-processing because FDM has relatively poor layer resolution. In this study, the mechanical properties of ABS parts fabricated by FDM were measured. The ABS parts were divided into one with vapor smoothing process and the other without the vapor smoothing process which is one of the post-processing methods. Using dynamic mechanical analysis (DMA) and dilatometer, temperature-dependent storage modulus and CTE for ABS specimens were measured. Based on the measured thermo-mechanical properties of ABS parts, finite element analysis was performed for an automotive bumper made of ABS. Moreover, response surface methodology was applied to study relationships among design parameters of thickness of the bumper, ambient temperature, and application of the vapor smoothing process. In result, a design guideline for a ABS product could be provided without time-consuming experiments

1. Introduction

The automotive industry has tried to increase energy efficiency for vehicles. In order to achieve the goal, one of ways is to reduce vehicle weight. If weight of the vehicle increases by 10%, fuel economy also increases by about 7% [1]. Therefore, researchers in the automotive industry have focused to use materials lighter than steel. One of the substitute materials is polymer. Currently, about 15% to 20% of total weight of a vehicle is in charge of polymer materials. The polymer materials are widely used for automobile parts including front ends, bumpers, fascia systems, grills, fenders, intake manifolds, engine mounts, door structures, panels, door impact beams, bonnet panels, under hood parts, pedal boxes, body structures, steering columns and so on [1]. Among several polymers, acrylonitrile butadiene styrene (ABS) is applied for bumpers, seats, dashboard, interior trim, exterior trim, and lighting and so on. This study focuses on ABS because fused deposition modeling (FDM), which is one of additive manufacturing techniques, could use ABS to make 3D printed structures.

Recently, additive manufacturing also known as 3D printing is considered as the key driver of the fourth industrial revolution. Compared to the conventional subtractive manufacturing such as turning, cutting, milling and grinding, 3D printing has unique advantages in design flexibility, personal fabrication, cost of geometry complexity, dimensional accuracy, time and cost efficiency in production run and so on [2]. Based on these merits, rapid development has been made in the several industrial areas such as aviation parts, manufacturing, and medical science [3]. However, the 3D printing has demerits of, scalability, trade-off between building and layer resolution, cost for mass manufacturing, material heterogeneity, need of anchor and support material, requirement for post-processing and so on [2]. So many researchers have solved the problems for the next generation of the 3D printing technique [2].

This study focuses on the post processing technique for the additive manufacturing. Additive manufacturing techniques have been classified into seven categories by ASTM international, which are shown in Figure.1.

Among them, FDM belongs to a category of material extrusion, which is an inexpensive and very popular 3D printing technique [2]. But FDM sometimes struggles with poor surface finish so that the post processing technique is necessary. The poor surface finish occurred by stair stepping effect could be treated by

*Corresponding Author: Sung-Uk Zhang, 82-51-890-1645, zsunguk@deu.ac.kr

polishing process such as sanding and vapor smoothing. The sanding may not be possible for tiny or complex geometries. On the other hand, the vapor smoothing could be effective for the complex shapes fabricated by 3D printing as shown in Figure. 2.

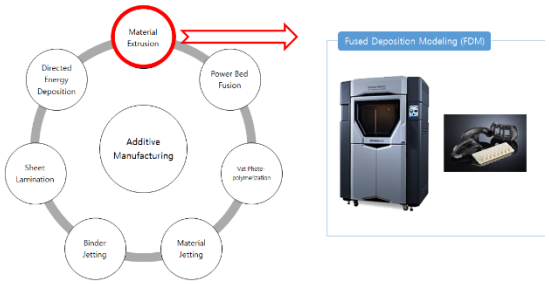


Figure.1. Seven categories for AM technologies

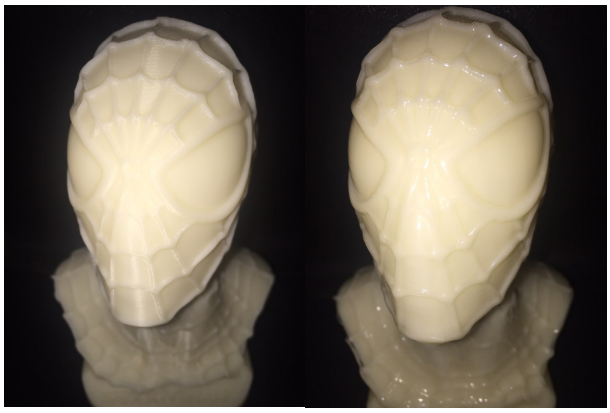


Figure.2. Vapor smoothing before and after

Many researchers have studied the vapor smoothing process and its effect in the mechanical properties. Kuo and Mao [4] developed an acetone-vapor smoothing system for ABS parts fabricated by FDM. They showed that surface roughness of the ABS parts could be significantly improved by using their system. Garg et al. [5] observed mechanical properties and surface roughness according to building parameters and vapor smoothing process for ABS specimens fabricated by FDM. In order to obtain the mechanical properties, they used ultimate testing machine (UTM). Singh et al. [6] studied vapor smoothing process for FDM parts. They had tried to improve the vapor smoothing process in term of surface roughness and dimensional accuracy by using DOE technique. Singh et al. [7] optimized building parameters of FDM and control parameters of vapor smoothing processes for hip joint as a biomedical application. They used Taguchi's method and response surface methodology to obtain the optimum level of process parameters. Cantrell et al. [8] measured mechanical properties of 3D printed ABS and PC parts with varying raster and building orientation of additive manufacturing. Gao et al. [9] investigates the impact of the vapor smoothing for 3D printed ABS parts in terms of the strength and elongation. They measured anisotropic mechanical properties of the ABS parts and showed that the material properties are affected by the length of process time for the vapor smoothing. Belter and Dollar [10] suggested a technique to strengthen 3D printed parts fabricated by FDM. They tried to remove voids in the printed parts by filling them with high-strength resins. They performed three-point bending test to obtain mechanical properties. Melenka et al. [11] evaluated mechanical properties and dimensional accuracy for 3D printed PLA parts fabricated by FDM. They used DOE technique to establish a relationship between building parameters and mechanical properties of the 3D printed parts.

In this study, temperature-dependent mechanical properties of 3D printed ABS parts fabricated by FDM and the acetone-vapor smoothing are measured by using dynamic mechanical analyzer (DMA) and dilatometer. Based on the measured data, finite element analysis and response surface methodology are performed in order to generate a design guideline for an automotive bumper made of ABS.

2. Methodology

The second section has four subsections. The first subsection includes geometry information of specimen for DMA and dilatometry. The second briefly explains about the vapor smoothing process. The third subsection describes about information measured by DMA and dilatometry. The last subsection introduces about finite element analysis and response surface methodology for an automotive bumper made of ABS.

2.1. Specimen fabrication for DMA and Dilatometry

Specimens for DMA and dilatometry were fabricated by using Fortus 3D printer which is based on Stratasys FDM technology. In modeler setup of Stratasys's software, building parameters were controlled, which are part interior style, visible surface style, and support style be set to sparse-low density, being enhanced, and being smart. The specimen for DMA has 27.8 mm length with 6.5 mm width and 2.0mm thickness and the size of specimen for dilatometry is 6 mm by 6 mm by 12 mm as shown in Figure. 3.

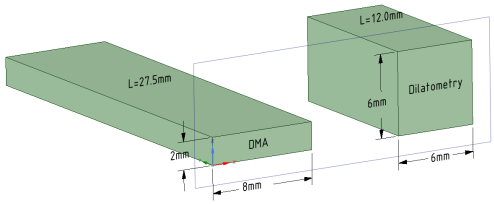


Figure.3. Geometry of the specimen for DMA and Dilatometer

2.2. Vapor smoothing process

Vapor smoothing process lets ABS plastic dissolve in acetone. So the secondary bonds between the ABS polymer chains is broken so that external layers of ABS 3D printed parts are softening. In result, this technique makes the surface finish of 3D printed parts be improved. A simple process for the vapor smoothing is shown in Figure. 4.

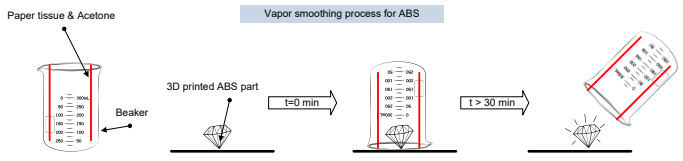


Figure.4. Vapor smoothing process

First of all, dust and debris from 3D printed ABS parts are removed. Small amount of acetone is poured on paper tissues. The tissues are attached inside of a beaker. Then, the 3D printed parts are covered with the beaker. Then, allow parts to be inside of the beaker for about 30 minutes. In this paper, amount of acetone was controlled from 5ml to 15ml in order to observe a relationship between the amount of acetone and mechanical properties of the 3D printed parts with the vapor smoothing.

2.3. Dynamic mechanical analyzer and Dilatometry

Dynamic Mechanical Analyzer (DMA), which known as DMA, is a powerful tool to characterize mechanical properties of a material as a function of temperature, time, frequency, stress, and so on. Q800 DMA was used, which is manufactured by TA instruments. Temperature-dependent storage modulus of ABS 3D printed parts could be observed by using DMA. As an applied mechanical load of DMA is a sinusoidal or oscillating force, complex modulus of E* could be obtained for polymer materials such as ABS, PLA, PC, and ULTEM, which consists of the storage modulus E' (real part) and the loss modulus E" (imaginary part) as follow

$$\mathbf{E}^* = E' + \mathbf{i}E'' \qquad (1)$$

$$\mathbf{tan}\delta = E''/E' \qquad (2)$$

where, tan δ is the ratio of loss modulus to storage modulus, which is an indicator of the energy lost. Figure.5 graphically shows temperature dependent storage modulus and tan δ measured by using DMA. In the graph of tan δ , the peak of tan δ indicates the glass transition temperature which means a midpoint between the glassy and rubbery states of a polymer.

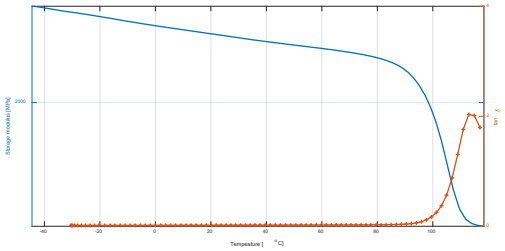


Figure.5.Temperature dependent storage modulus and tan δ obtained by using DMA

Moreover, a dilatometer is used to measure temperature dependent CTE of the polymer materials, which is an instrument to measure volume changes occurred by a physical or chemical event. In this study, the physical event is temperature change. The thermal expansion is expressed as

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right) \qquad (3)$$

where, α is a coefficient of thermal expansion (CTE), V is a volume, and T is a temperature. We used DIL-402C manufactured by NETZSCH. Using the dilatometry, we could obtain temperature dependent CTEs for ABS 3D printed parts.

2.4. Finite element analysis and response surface methodology for a 3D printed bumper made of ABS

A portion of a bumper is modeled and structurally evaluated by using finite element analysis as shown in Figure. 6 and Figure. 7. Fig 6 shows the geometry of the automotive bumper. Fig 7 graphically shows fixed boundary conditions, load boundary conditions, and body temperature condition on the geometry. In Figure. 7, the symbols of A, B and C mean boundary conditions of fixed support, pressures, thermal conditions.

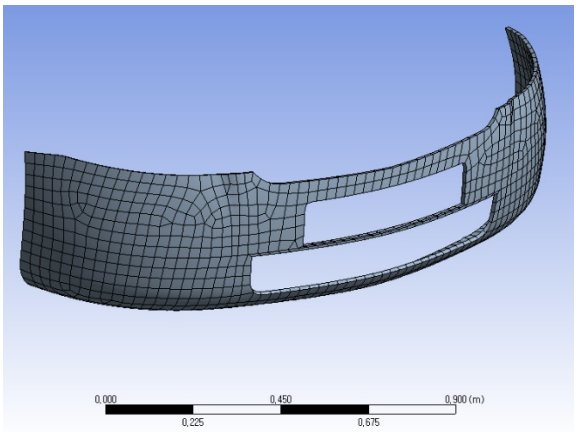


Figure.6. Finite element model for an automotive bumper

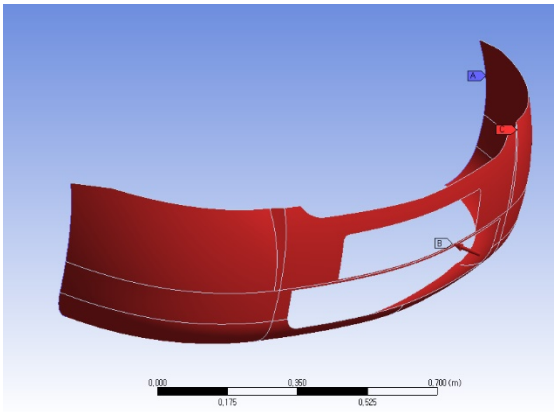


Figure.7. Boundary conditions for the finite element model

Input parameters for the response surface methodology were temperature, thickness of bumper, element size, and application of the vapor smoothing. The range of each parameter is shown in Table 1.

Table 1. The factor ranges for input parameters

Input parameter	[min, max]
Temperature [°C]	[-40, 70]
Thickness of Bumper [m]	[0.01,0.03]
Element Size [m]	[0.03,0.04]
Vapor smoothing	w/o, w/

Box-Behnken design was used for the response surface methodology. Design points were generated by using Minitab software as shown in Figure. 8. And the output parameter was a maximum of total deformation of the bumper.

3. Results and Discussion

The first describes thermo-mechanical properties of ABS specimens measured by DMA and dilatometry. The second subsection depicts numerical simulation for an automotive bumper made of ABS.

3.1. Thermo-mechanical properties of ABS specimens

Four different splits were measured, which were exposed to different amount of acetone vapor; 0ml, 5ml, 10ml and 15ml. Three specimens per each split were fabricated and measured in the same condition. Averaged thermo-mechanical properties of the

ABS specimens were obtained, which are the storage modulus and CTE as shown in Figure.9 and Figure.10.

	C1	C2	C3	C4	C5	C6	C7	C8-T
	StdOrder	RunOrder	PtType	Blocks	Temperature[C]	Thickness of Bumper [m]	Element size [m]	Vapor smoothing
1	1	1	2	1	-40	0.01	0.04	w/o
2	2	2	2	1	70	0.01	0.04	w/o
3	3	3	2	1	-40	0.03	0.04	w/o
4	4	4	2	1	70	0.03	0.04	w/o
5	5	5	2	1	-40	0.02	0.03	w/o
6	6	6	2	1	70	0.02	0.03	w/o
7	7	7	2	1	-40	0.02	0.05	w/o
8	8	8	2	1	70	0.02	0.05	w/o
9	9	9	2	1	15	0.01	0.03	w/o
10	10	10	2	1	15	0.03	0.03	w/o
11	11	11	2	1	15	0.01	0.05	w/o
12	12	12	2	1	15	0.03	0.05	w/o
13	13	13	0	1	15	0.02	0.04	w/o
14	14	14	0	1	15	0.02	0.04	w/o
15	15	15	0	1	15	0.02	0.04	w/o
16	16	16	2	1	-40	0.01	0.04	w/
17	17	17	2	1	70	0.01	0.04	w/
18	18	18	2	1	-40	0.03	0.04	w/
19	19	19	2	1	70	0.03	0.04	w/
20	20	20	2	1	-40	0.02	0.03	w/
21	21	21	2	1	70	0.02	0.03	w/
22	22	22	2	1	-40	0.02	0.05	w/
23	23	23	2	1	70	0.02	0.05	w/
24	24	24	2	1	15	0.01	0.03	w/
25	25	25	2	1	15	0.03	0.03	w/
26	26	26	2	1	15	0.01	0.05	w/
27	27	27	2	1	15	0.03	0.05	w/
28	28	28	0	1	15	0.02	0.04	w/
29	29	29	0	1	15	0.02	0.04	w/
30	30	30	0	1	15	0.02	0.04	w/

Figure.8. Design points generated by Minitab software

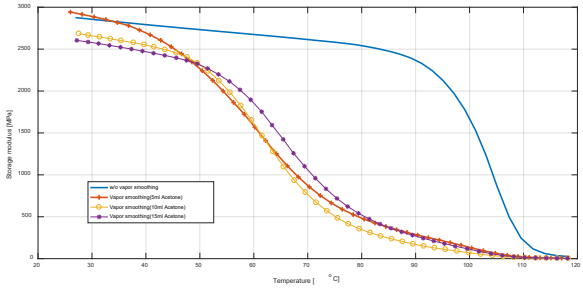


Figure.9. Average of temperature dependent storage moduli for four splits

Figure.9 graphically shows the difference between specimens without vapor smoothing process and ones with vapor smoothing process. The glass transition temperature of the specimens with the vapor smoothing process is much lower than one without the vapor smoothing process.

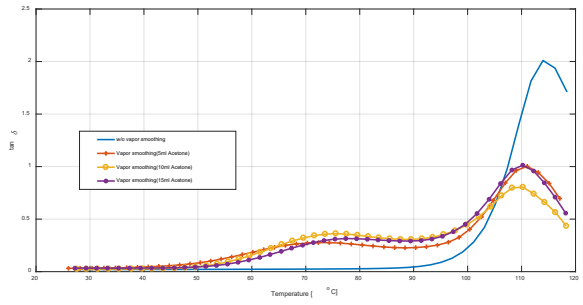


Figure.10. Average of tan δ for four splits

Figure.10 graphically shows the difference between specimens without vapor smoothing process and ones with vapor smoothing process in terms of tan δ. In the graph of first split, peak point of tan δ occurs at about 114ºC. However, the first peaks of the other specimens with the vapor smoothing are shown at 67ºC, 71ºC, and 74ºC. Figure.11 shows that CTEs vary as the temperature increases. CTE of the specimen without vapor smoothing is relatively constant.

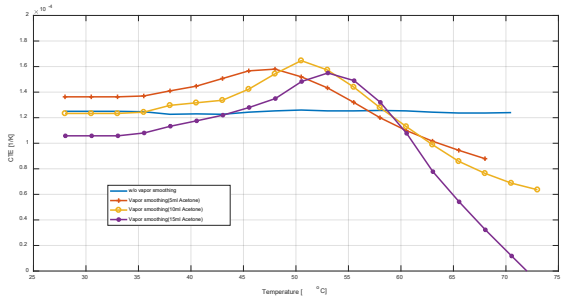


Figure.11. Average of the temperature dependent CTEs for four splits

However, CTEs of the others with vapor smoothing have large variations and have one peak between 45ºC and 60ºC. As the temperature increases more than 60ºC, CTEs of the samples with vapor smoothing decreases.

Figure.9, Figure.10 and Figure. 11 obviously shows which split does not have the vapor smoothing process. Figures represent that the vapor smoothing process weakens thermal stability of ABS 3D printed parts. However, above figures could not show any correlation between CTE and the amount of acetone vapor in the range of 5ml to 15ml.

3.2. Numerical evaluation for an automotive bumper made of ABS

The output parameter was the maximum value of total deformation of the automotive bumper as shown in Figure. 12. The maximum deformation occurs at the center of the bumper under the given boundary conditions.

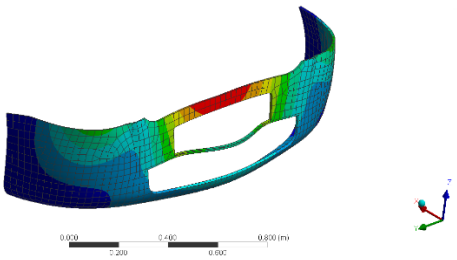


Figure.12. Total deformation of the automotive bumper

Using the response surface methodology and the finite element analysis, two regression equations with varying application of the vapor smoothing process were obtained as follow

$$D[m] = 0.1538 - 0.001126 \times T[C] - 5.97 \times Th[m] - 3.80 \times Elem[m] + 0.000006 \times T[C] \times T[C] + 108.4 \times Th[m] \times Th[m] + 51.7 \times Elem[m] \times Elem[m] - 0.01157 \times T[C] \times Th[m] + 0.02709 \times T[C] \times Elem[m] \quad (4)$$

$$D[m] = 0.182 - 0.000793 \times T[C] - 7.10 \times Th[m] - 3.81 \times Elem[m] + 0.000006 \times T[C] \times T[C] + 108.4 \times Th[m] \times Th[m] + 51.7 \times Elem[m] \times Elem[m] - 0.0116 \times T[C] \times Th[m] + 0.0271 \times T[C] \times Elem[m] + 0.2 \times Th[m] \times Elem[m] \quad (5)$$

where, D is the maximum deformation of the automotive bumper, T is the temperature, Th is the thickness of the bumper and Elem is the size of finite elements. When the vapor smoothing process is not applied, Equation (4) could be used. Otherwise, equation (5) could be used to estimate the maximum deformation of the bumper. R-square value of the response surface model was 81%. Two equations represent that the maximum deformation is

sensitive to the thickness of bumper and the temperature under given ranges. Based on two equations, contour plots were computed by using Minitab software as shown in Figure.12. According to the vapor smoothing process, two design guidelines could be obtained by using the response surface methodology, which might be useful information for bumper designers.

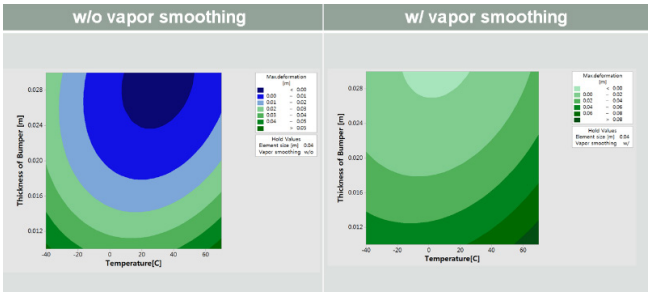


Figure.13. Contour plots according to the vapor smoothing process

4. Conclusion

In this study, the temperature-dependent mechanical properties of ABS parts were measured by used dynamic mechanical analyzer and dilatometry, which were fabricated by FDM and the vapor smoothing technique. The vapor smoothing process weakens thermal stability of ABS 3D printed parts. According to the temperature-dependent storage modulus of 3D printed ABS specimen, the glass transition temperature dramatically decreases by the process. The vapor smoothing process made relatively large variation in CTE of ABS 3D printed parts. Based on the measured thermos-mechanical properties of ABS specimen, finite element analysis and response surface methodology for an automotive bumper made of ABS were performed so that two design guidelines could be provided according to the vapor smoothing process.

Acknowledgment

This work was supported by Dong-Eui University Grant (201702720001) and the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (No. NRF-2017R1C1B5074219).

References

[1] Lyu MY, Choi TG. Research trends in polymer materials for use in lightweight vehicles 2015;16:213–20. doi:10.1007/s12541-015-0029-x.

[2] Gao W, Zhang Y, Ramanujan D, Ramani K, Chen Y, Williams CB, et al. The status, challenges, and future of additive manufacturing in engineering. Comput Des 2015;69:65–89. doi:10.1016/j.cad.2015.04.001.

[3] Chen L, He Y, Yang Y, Niu S, Ren H. The research status and development trend of additive manufacturing technology. Int J Adv Manuf Technol 2017;89:3651–60. doi:10.1007/s00170-016-9335-4.

[4] Kuo C-C, Mao R-C. Development of a Precision Surface Polishing System for Parts Fabricated by Fused Deposition Modeling. Mater Manuf Process 2016;31:1113–8. doi:10.1080/10426914.2015.1090594.

[5] Garg A, Bhattacharya A, Batish A. Chemical vapor treatment of ABS parts built by FDM: Analysis of surface finish and mechanical strength. Int J Adv Manuf Technol 2017;89:2175–91. doi:10.1007/s00170-016-9257-1.

[6] Singh R, Singh S, Singh IP, Fabbrocino F, Fraternali F. Investigation for surface finish improvement of FDM parts by vapor smoothing process. Compos Part B Eng 2017;111:228–34. doi:10.1016/j.compositesb.2016.11.062.

[7] Singh J, Singh R, Singh H. Repeatability of linear and radial dimension of ABS replicas fabricated by fused deposition modelling and chemical vapor smoothing process: A case study. Measurement 2016;94:5–11. doi:10.1016/j.measurement.2016.07.064.

[8] Cantrell JT, Rohde S, Damiani D, Gurnani R, DiSandro L, Anton J, et al. Experimental characterization of the mechanical properties of 3D-printed ABS and polycarbonate parts. Rapid Prototyp J 2017;23:811–24. doi:10.1108/RPJ-03-2016-0042.

[9] Gao H, Kaweesa D V., Moore J, Meisel NA. Investigating the Impact of Acetone Vapor Smoothing on the Strength and Elongation of Printed ABS Parts. Jom 2017;69:580–5. doi:10.1007/s11837-016-2214-5.

[10] Belter JT, Dollar AM. Strengthening of 3D printed fused deposition manufactured parts using the fill compositing technique. PLoS One 2015;10:1–19. doi:10.1371/journal.pone.0122915.

[11] Melenka GW, Schofield JS, Dawson MR, Carey JP. Evaluation of dimensional accuracy and material properties of the MakerBot 3D desktop printer. Rapid Prototyp J 2015;21:618–27. doi:10.1108/RPJ-09-2013-0093.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 162-174 (2017)
www.astesj.com

ASTES Journal
ISSN: 2415-6698

Mealy-to-Moore Transformation - A state stable design of automata

Mustafa Oezguel*, Florian Deeg, Sebastian M. Sattler

Chair of Reliable Circuits and Systems, Friedrich-Alexander-Universität Erlangenen-Nürnberg, Paul-Gordan-Str. 5, 91052 Erlangen, Germany

ARTICLEINFO

Article history:
Received: 27 October, 2017
Accepted: 02 December, 2017
Online: 10 December, 2017

Keywords :
Asynchronous Feedback
Functional Safety
Stabilization
Hazard-free
Race-free
Parallel de-composition
Dual-rail

ABSTRACT

The paper shows a method of transforming an asynchronously feedbacked Mealy machine into a Moore machine. The transformation is done in dual-rail logic under the use of the RS-buffer. The transformed machine stabilizes itself and is safe to use. The transformation is visualized via KV-diagrams and calculated with formulas. We will present three use-cases for a better understanding. To underpin the stated transformation a simulation is also presented.

1 Introduction

A Synchronous circuits entail several benefits. Firstly performance advantages, e.g. faster system perfomance, no clock synchronization, fewer power consumption and fewer electromagnetic emission. Secondly asynchronous circuits also have safety advantages, e.q. no clock skew. Asynchronous circuits can also lead to simplicity, by connecting different parts modularly [1, 2]. On the other hand asynchronous circuits provide risks, such as faults like hazards, glitches and so on [3]. Therefore it is very important to design asynchronous circuits correctly, especially in safety critical applications such as in automotive, where human lives are at stake. To prevent faults like hazards, a mealy-to-moore transformation will be presented, which creates a function stable asynchronous machine. The Mealy machine with the state transfer function δ and the output function λ, see figure 1 be given. Each Mealy machine can be transformed into an equivalent Moore machine [3] for example by increasing the arity of δ by $|x|$ and coding it with x, $z' = (z,x)$ with $\delta'(z',x) \mapsto (\delta(z,x),x)$, and the output function is only dependent on the new state variable z' with $\mu : (z') \mapsto \mu(z')$.
By comparing the two machines the pros and cons can be shown. Therefore the idea of transforming the machines into each other to profit from the benefits of both is obvious. Mealy machines have the advantage of requiring less states since one state can produce a number of different outputs in combination with the input. A Moore machine's state on the other hand only produces one output. A Mealy machine is also faster by reacting directly to the input. This feature however is not always wanted, since it can lead to undesired outputs (e.g. hazards, glitches) when the input is variable. A Moore machine is more stable in this regard, since it only indirectly reacts to input changes. The output only changes when transferring into the next state. Transforming a Mealy machine into a Moore machine is therefore useful in case a direct dependence on the input is to be avoided [2].
In the paper the transformation of a Mealy machine into a Moore machine is presented. The machine will be designed under the use of dual-rail logic and the RS-buffer, and will stabilize itself. Firstly the transformation is made by breaking the feedback up and nesting it in the RS-buffer, secondly by encoding the inputs as states, so that the output function is only dependent on pseudo states. With the method presented in this article, function stable sequential circuit parts can be seen as combinational logic and moved over other blocks at will. With this method, individual machines can be strung together to realize complex circuits for safety relevant applications.

2 Theory

This paper describes the underlying theory of the transformation and provides an illustrative example.

Figure 1 shows a fully asynchronous Mealy machine. The branches entering a node in the graph should end reflexively, so that only transient states are allowed which are conscious and triggered from the outside [2]. To consider all branches of the Mealy machine as locally reflexively

*Corresponding Author: Mustafa Oezguel, mustafa.oezguel@fau.de

https://dx.doi.org/10.25046/aj020621

concluded, the feedback should be moved over λ, making the output y the feedback. In order to set the correct state z for the transfer function δ, a function λ^{-1} is realized. This function generates the reduced state $\underline{z}$ from the feedback y. This equivalent transformation is outlined in figure 2.

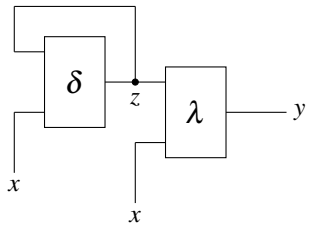


Figure 1: Fully asynchronous Mealy machine

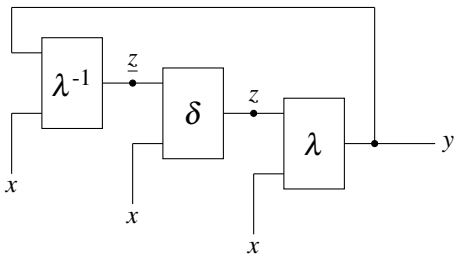


Figure 2: Equivalent transformation

The following applies:

$$\begin{aligned} \delta(\underline{z},x) &= \underline{z} \\ \delta\left(\lambda^{-1}(y,x),x\right) &= \lambda^{-1}(y,x) \\ \lambda\left(\delta\left(\lambda^{-1}(y,x),x\right),x\right) &= y \\ \text{mit } \lambda\left(\lambda^{-1}(y,x),x\right) &= y \end{aligned}$$

The next stage of transformation shows at first sight no feedback. Only the RS-buffer, which is in $\underline{\delta}$, has a nested feedback. This transformation is outlined in figure 3. Transient states can be triggered via the input, while state stable states are being held via the RS-buffer. In order to set the correct state $\underline{z}$ to μ, the function $\underline{\delta}$ must be well designed, which will be shown in the following sections.

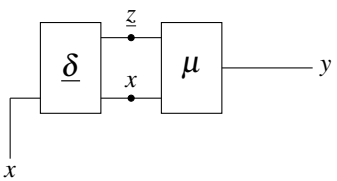


Figure 3: Moore transformation

The following applies:

$$\begin{aligned} \underline{\delta}(\underline{z},x) &= \underline{z} \\ \mu(\underline{z},x) &= y \end{aligned}$$

2.1 Dual-rail logic

For the implementation in dual-rail logic [4], the fully asynchronous circuit from figure 1 will be divided in the 1- and 0-share. This is done by partitioning the state transfer function and the output function, see figure 4.

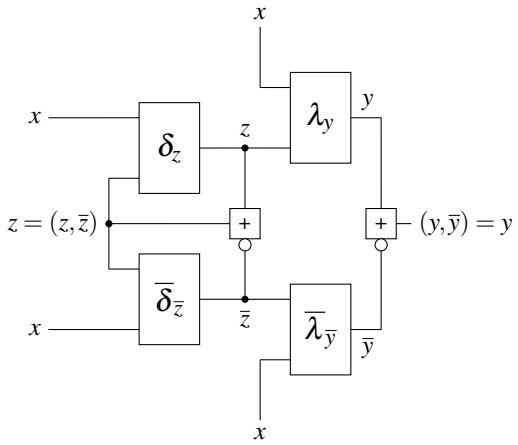


Figure 4: Mealy realized in dual-rail logic

The functions δ and λ are each w.l.o.g. realized in two blocks $\delta = (\delta_z, \overline{\delta}_{\overline{z}})$ and $\lambda = (\lambda_z, \overline{\lambda}_{\overline{z}})$. In order to guarantee this secure dual rail structure, RS-buffers are used. The mode of operation corresponds in a certain manner to the C-element [4, 5]. The function $\underline{\delta}$ of figure 3 will be designed by appropriate coding according to the RS-buffer, showing a general structure as outlined in figure 5.

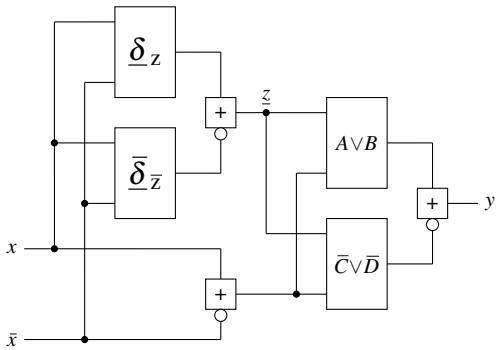


Figure 5: Transformed Moore

The stabilized 1-states, $\underline{\Lambda}_z \circ\!\!-\!\!\bullet \underline{z}$, are transformed into 1-outputs, $\Lambda_{\underline{y}} \circ\!\!-\!\!\bullet \underline{y}$. The corresponding Venn diagram of the stabilized 1-states in 1-outputs can be seen in figure 6. 1-states that appear as 0-outputs are declared as $\overline{\Lambda}_{\underline{z}}$, 0-states which appear as 1-outputs are declared as $\Lambda_{\overline{z}}$. The 1-partition is composed of:

$$\Lambda_{\underline{y}} \circ\!\!-\!\!\bullet \lambda_y(\underline{\delta}_z) + \lambda_y(\underline{\overline{\delta}}_{\overline{z}}) + \overline{\lambda}_{\overline{y}}(\underline{\delta}_z)$$

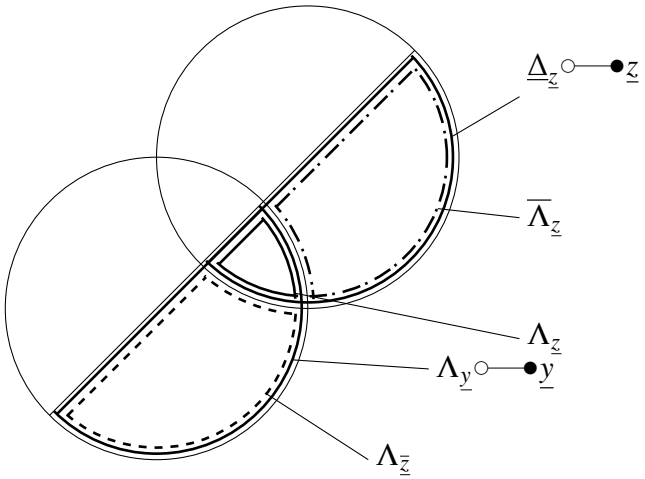


Figure 6: Venn diagram of 1-states in 1-outputs

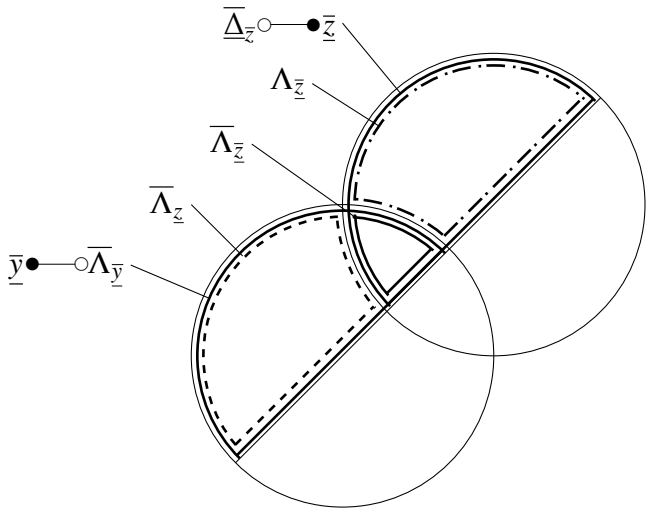


Figure 7: Venn diagram of 0-states in 0-outputs

The stabilized 0-states, $\overline{\underline{\Delta}}_{\overline{\underline{z}}} \circ\!\!-\!\!\bullet \overline{\underline{z}}$, are transformed into 0-outputs, $\overline{\Lambda}_{\overline{\underline{y}}} \circ\!\!-\!\!\bullet \overline{\underline{y}}$. The corresponding Venn diagram of the 0-states in 0-outputs is shown in figure 7. 0-states which appear as 1-outputs are declared as $\Lambda_{\overline{\underline{z}}}$, 1-states which appear as 0-outputs are declared as $\overline{\Lambda}_{\underline{z}}$. The 0-partition is composed of:

$$\overline{\Lambda}_{\overline{\underline{y}}} \circ\!\!-\!\!\bullet \overline{\lambda}_{\overline{y}}(\overline{\underline{\delta}}_{\overline{z}}) + \overline{\lambda}_{\overline{y}}(\underline{\delta}_{\underline{z}}) + \lambda_{y}(\overline{\underline{\delta}}_{\overline{z}})$$

To understand the coding, first the RS-buffer has to be understood.

2.2 RS-Buffer

The RS-Buffer used and its circuit symbol are represented in figure 8 and 9 and

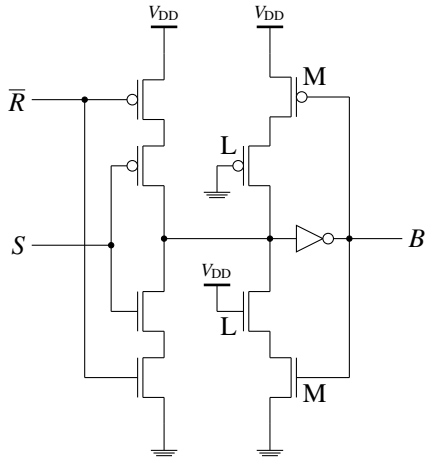


Figure 8: RS-Buffer (schematic)

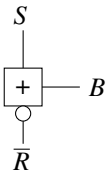


Figure 9: Circuit symbol of the used RS-Buffer

its truth table is specified in table 1.

R	0	0	1	1
S	0	1	0	1
B	B	1	0	B

Table 1: Truth table of the RS-Buffer

The RS-buffer is used to synchronize a dual-rail signal by only letting a state tranfer happen, if both incoming signals are disjoint (how it is desired) to each other. This means that there is no overlaying of signal, which could lead to fatal faults. The RS-buffer consists of the tri-state driver, which is the first state and the second stage the so called babysitter. The driving structure now provides 3 functions. Firstly the setting signal, which means the output is in the logical 1-state. Secondly the resetting signal, which means the output is in logical 0-state and least the hold state which holds the last state at the output by keeping the signal in the babysitter and having a high impedance state at the tri state. The babysitter consists of two complementary loops. The formula for the output B is therefore $B := \overline{R}(B \vee S) \vee S(B \vee \overline{R}) = \overline{R}S \vee \overline{R}B \vee SB$.

2.3 Design of the reduced state transfer function $\underline{\delta}$

As you can see the RS-buffer is coded via a 3-partitioning. There is the setting signal S=[1], the resetting signal R=[1] and the hold signal RS=[00]$\vee$[11]. The state transfer function of an automaton can in turn be split in 4 parts. Firstly

the transient 1-share, that means independent of the state the next state will be at logical 1, secondly the transient 0-share, that means output at logical 0 independent of the state, thirdly the state stable share of the state transfer function, that means the share, which is at constant input dependent on the value of the last state and the state instable state, which means a permanent change of the state at constant input, see figure 10. The resulting 3 parts, which are left, can now be easily compared to the 3-partitition of the RS-buffer. The transient 1-share $\underline{\delta}_z$ is realized by the setting signal of the RS-buffer, similar the transient 0-share $\overline{\underline{\delta}}_{\overline{z}}$ is realized by the resetting signal and the state stable share is realized by the hold signal of the RS-buffer.

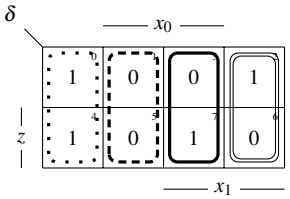


Figure 10: Parts of a state transfer function

To understand the transformation, first a few compositions must be defined:

$$
\begin{aligned}
{}_{z}\delta(z,x) &= \delta(z=1,x)\\
{}_{\overline{z}}\delta(z,x) &= \delta(z=0,x)\\
{}_{z}\overline{\delta}(z,x) &= \overline{\delta}(z=1,x)\\
{}_{\overline{z}}\overline{\delta}(z,x) &= \overline{\delta}(z=0,x)
\end{aligned}
$$

This work only concentrates on function stable machines, so that state instable parts are not allowed. State instable parts can be deduced via

$${}_{z}\overline{\delta} \wedge {}_{\overline{z}}\delta$$

Functions must not exhibit these parts. The other parts can be deduced by the following expressions

$$
\begin{aligned}
&\text{transient } 1: \underline{\delta}_z = {}_{z}\delta \wedge {}_{\overline{z}}\delta\\
&\text{transient } 0: \overline{\underline{\delta}}_{\overline{z}} = {}_{z}\overline{\delta} \wedge {}_{\overline{z}}\overline{\delta}\\
&\text{state stable}: {}_{z}\delta \wedge {}_{\overline{z}}\overline{\delta}
\end{aligned}
$$

This leads to the genereal formula for $\underline{\delta}_i$ for n states:

$$
\begin{aligned}
\underline{\delta}_{z_i} &= \underline{\delta}_{z_i} + \neg\overline{\underline{\delta}}_{\overline{z}_i}\\
&= ({}_{z_i}\delta_{z_i} \wedge {}_{\overline{z}_i}\delta_{z_i}) + \neg({}_{z_i}\overline{\delta}_{\overline{z}_i} \wedge {}_{\overline{z}_i}\overline{\delta}_{\overline{z}_i})\\
&= (\delta_{z_i}(z_i=1,x) \wedge \delta_{z_i}(z_i=0,x))\\
&+ \neg(\overline{\delta}_{\overline{z}_i}(z_i=1,x) \wedge \overline{\delta}_{\overline{z}_i}(z_i=0,x))
\end{aligned}
$$

with $i = 0...n-1$

2.4 Design of the output function

The detailed description of the Venn-diagrams given in figure 11 and 12 is reported in [6].

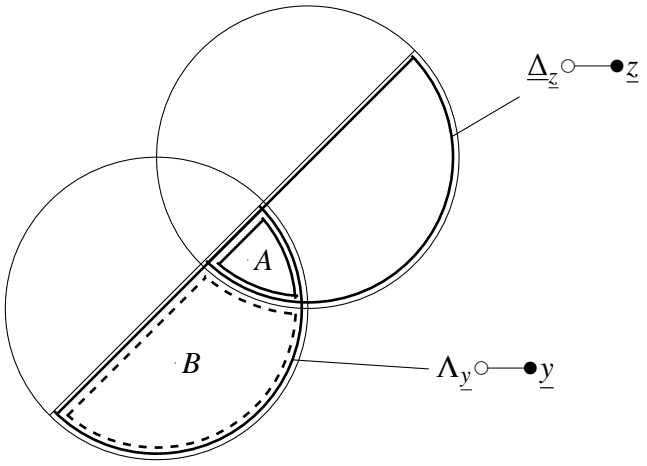


Figure 11: Venn-diagram of 1-states in 1-outputs

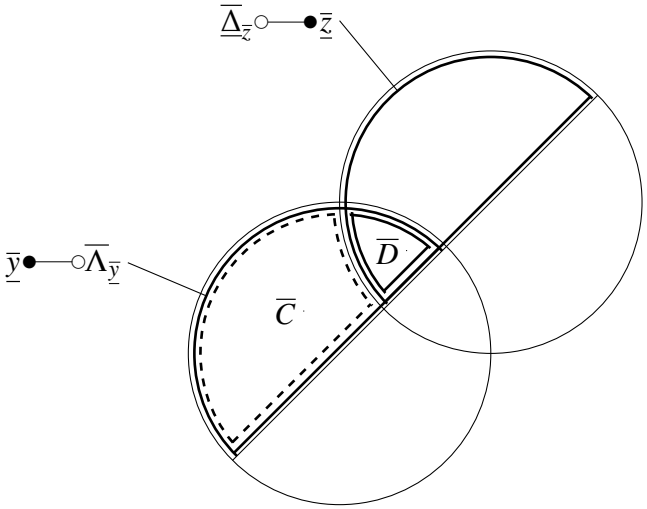


Figure 12: Venn-diagram of 0-states in 0-outputs

The coding of the parts A, B, $\overline{C}$ and $\overline{D}$ of the Venn-diagrams, see figure 11 and figure 12, is as follows:

$$
\begin{aligned}
&(z,y)\\
&(11) = A = \delta_z \wedge \lambda_y\\
&(01) = B = \overline{\delta}_{\overline{z}} \wedge \lambda_y\\
&(10) = \overline{C} = \delta_z \wedge \overline{\lambda}_{\overline{y}}\\
&(00) = \overline{D} = \overline{\delta}_{\overline{z}} \wedge \overline{\lambda}_{\overline{y}}
\end{aligned}
$$

The dual-rails y and $\overline{y}$ for y are now:

$$
\begin{aligned}
y &= A \vee B\\
\overline{y} &= \overline{C} \vee \overline{D}\\
y &= y + \neg\overline{y}
\end{aligned}
$$

The seperate parts of the machine have now been deduced and can simply be strung together to get the form of figure 5.

3 Use-Case

To get a better insight, three examples will be presented.

3.1 1-dimensional example

For $x = (x_1, x_0)$, $y = (y)$ and $z = (z)$ the transformation of the Mealy machine $(X, Y, Z, \delta, \lambda)$ with the state transfer function and output function

$$\delta(z,x) = x_1x_0 \vee z\overline{x}_1x_0 \qquad \lambda(z,x) = \overline{x}_1x_0 \vee z\overline{x}_1$$

will be executed. The automaton before the transformation can be seen in figure 13.

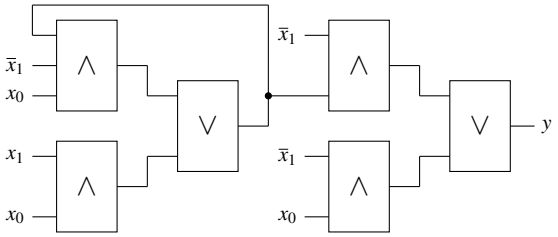


Figure 13: Mealy machine before the transformation

3.1.1 Design of $\underline{\delta}$

The KV-diagram of the state transfer function can be seen in figure 14.

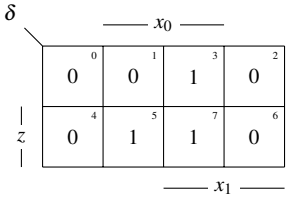


Figure 14: The state transfer function

To code the RS-buffer correctly, firstly the KV-diagram will be compacted, see figure 15.

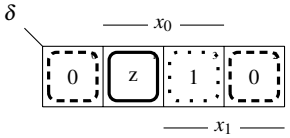


Figure 15: Compacted KV-diagram of δ

The parts of the state transfer function can easily be seen and the KV-diagrams for $\underline{\delta}_z$ and $\overline{\underline{\delta}}_{\overline{z}}$ can be deployed, see figure 16. For $\overline{\underline{\delta}}_{\overline{z}}$ 0 is coded as 1 and 1 is coded as 0 respectively, because of the dual-rail structure.

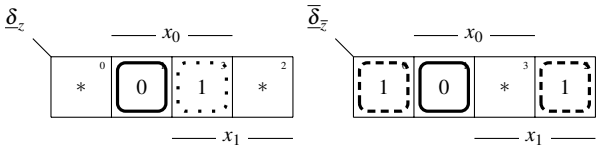


Figure 16: $\underline{\delta}_z$ and $\overline{\underline{\delta}}_{\overline{z}}$

For the next steps of the transformation, first the complement of δ has to be calculated:

$$\overline{\delta}(z,x) = \overline{z}\overline{x}_1 \vee \overline{x}_0$$

The parts of the state transfer function are:

$$\begin{aligned}
{}_z\delta &= \delta(z=1,x) = x_0 \\
{}_{\overline{z}}\delta &= \delta(z=0,x) = x_1x_0 \\
{}_z\overline{\delta} &= \overline{\delta}(z=1,x) = \overline{x}_0 \\
{}_{\overline{z}}\overline{\delta} &= \overline{\delta}(z=0,x) = \overline{x}_1 \vee \overline{x}_0
\end{aligned}$$

It has to be checked, if there are any instable parts:

$${}_{\overline{z}}\delta_z \wedge {}_z\overline{\underline{\delta}}_{\overline{z}} = x_1x_0 \wedge \overline{x_0} = 0$$

${}_{\overline{z}}\delta \wedge {}_z\overline{\delta}$ is 0 means the function has no instable parts. The functions $\underline{\delta}_z$ and $\overline{\underline{\delta}}_{\overline{z}}$ can now be calculated via

$$\begin{aligned}
\underline{\delta}_z &= {}_z\delta_z \wedge {}_{\overline{z}}\delta_z = x_1x_0 \\
\overline{\underline{\delta}}_{\overline{z}} &= {}_z\overline{\delta}_{\overline{z}} \wedge {}_{\overline{z}}\overline{\delta}_{\overline{z}} = \overline{x}_0
\end{aligned}$$

The coding for $\underline{\delta}$ of the RS-buffer is as follows:

$$\begin{aligned}
&(r\ s) \\
\text{Set:} &(*1) = x_1x_0\ (\underline{\text{transient 1-share}}) = \underline{\delta}_z \\
\text{Reset:} &(1*) = \overline{x}_0\ (\underline{\text{transient 0-share}}) = \overline{\underline{\delta}}_{\overline{z}} \\
\text{Hold:} &(**) = \overline{x}_1x_0\ (\underline{\text{state stable share}})
\end{aligned}$$

Now $\underline{\delta}$ can be derived:

$$\begin{aligned}
\underline{\delta} &= \underline{\delta}_z + \neg\overline{\underline{\delta}}_{\overline{z}} \\
&= x_1x_0 + \neg(\overline{x}_0)
\end{aligned}$$

3.1.2 Design of λ

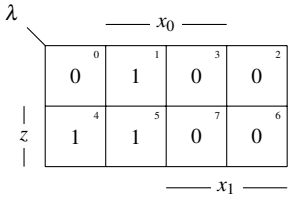


Figure 17: Output function λ

The stabilized state transfer function has now been designed, now the output function must be designed. Therefore the output function has to be layed over the state transfer function and the parts have to be specified:

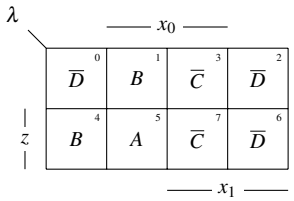


Figure 18: Parts of the output function

The parts are:

$$A(z,x) = z\overline{x}_1x_0$$
$$B(z,x) = \overline{z}\overline{x}_1x_0 \vee z\overline{x}_1\overline{x}_0$$
$$\overline{C}(z,x) = x_1x_0$$
$$\overline{D}(z,x) = x_1\overline{x}_0 \vee \overline{z}\overline{x}_0$$

This leads to the output function:

$$y = A \vee B = \overline{x}_1x_0 \vee z\overline{x}_1$$
$$\overline{y} = \overline{C} \vee \overline{D} = x_1 \vee \overline{z}\overline{x}_0$$
$$y = y + \neg\overline{y} = \overline{x}_1x_0 \vee z\overline{x}_1 + \neg(x_1 \vee \overline{z}\overline{x}_0)$$

The resulting automaton has the form of figure 19 with the determined functions $y = A \vee B$ and $\overline{y} = \overline{C} \vee \overline{D}$.

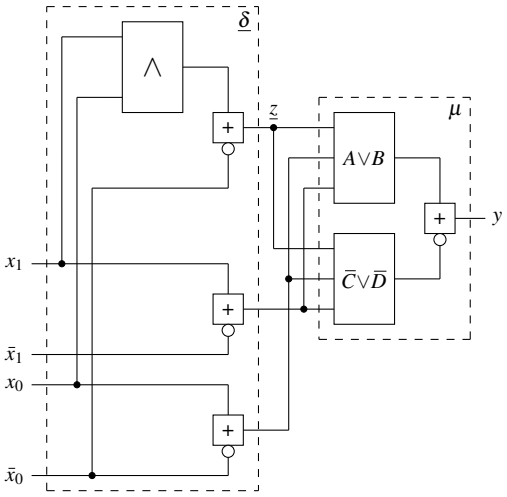


Figure 19: Transformed Moore

The truth table of the automaton:

#	z	x_1	x_0	δ	$\underline{\delta}_z$	$\overline{\underline{\delta}}_{\overline{z}}$	$\underline{\delta}$	λ	A	B	$\overline{C}$	$\overline{D}$	$A \vee B$	$\overline{C} \vee \overline{D}$	y
0	0	0	0	0	*	1	0	0	*	*	*	1	*	1	0
1	0	0	1	0	*	*	0	1	*	1	*	*	1	*	1
2	0	1	0	0	*	1	0	0	*	*	*	1	*	1	0
3	0	1	1	1	1	*	1	0	*	*	1	*	*	1	0
4	1	0	0	0	*	1	0	1	*	1	*	*	1	*	1
5	1	0	1	1	*	*	1	1	1	*	*	*	1	*	1
6	1	1	0	0	*	1	0	0	*	*	*	1	*	1	0
7	1	1	1	1	1	*	1	0	*	*	1	*	*	1	0

Table 2: Partial truth table of the resulting automaton

3.2 2-dimensional example

For $x = (x_1, x_0)$, $y = (y)$ and $z = (z_1, z_0)$ the transformation of the Mealy machine $(X, Y, Z, \delta, \lambda)$ with the state transfer functions and output function

$$\delta_{z_0}(z,x) = \overline{x}_0 \vee z_1x_1 \qquad \delta_{z_1}(z,x) = z_0x_1x_0 \vee z_1x_1\overline{x}_0$$
$$\lambda(z,x) = z_1x_1 \vee z_0\overline{x}_1\overline{x}_0$$

will be executed. The machine before the transformation is outlined in figure 20.

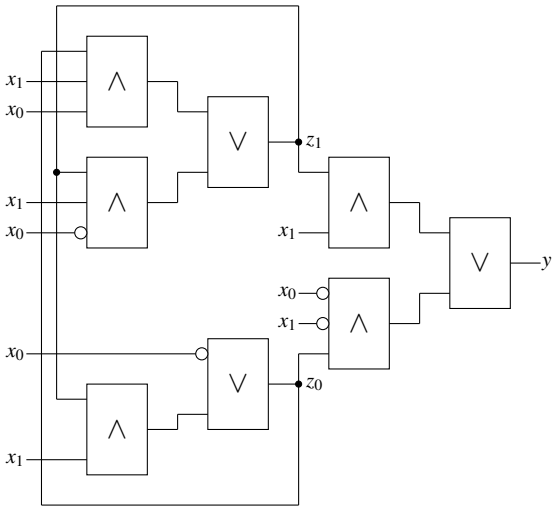


Figure 20: Automaton before the transformation

3.2.1 Design of $\underline{\delta}_{z_0}$

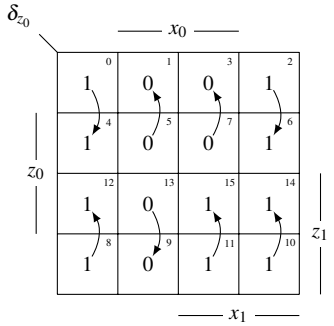


Figure 21: KV-diagram of δ_{z_0}

The KV-diagram of figure 21 will be lossless compacted given by the pointers in the diagram to the diagram of figure 22.

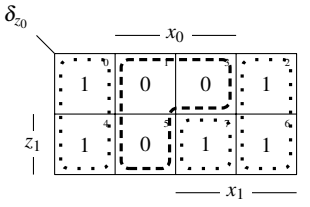


Figure 22: compacted KV-diagram of δ_{z_0}

With the compacted KV-diagram the RS-buffer for δ_{z_0} can be coded by specifying the functions $\underline{\delta}_{z_0}$ and $\overline{\underline{\delta}}_{\overline{z}_0}$, see

figure 23. The partial function $\underline{\delta}_{z_0}$ delivers the 1's of δ_{z_0} as 1's, while $\overline{\underline{\delta}}_{\bar{z}_{z_0}}$ delivers the 0's of δ_{z_0} as 1's. In this state transfer function, there is no state stable part only transient parts. Making the stable $\underline{\delta}_{z_0}$ a normal dual-rail structure.

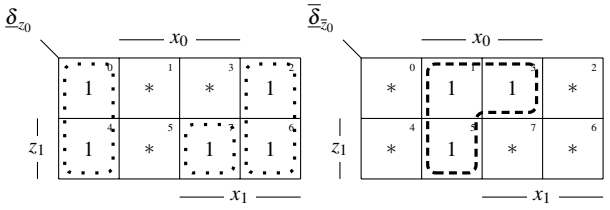


Figure 23: $\underline{\delta}_{z_0}$ and $\overline{\underline{\delta}}_{\bar{z}_0}$

State instable parts are not allowed in the state transfer function:

$$\bar{z}_0\delta_{z_0} \wedge_{z_0} \overline{\delta}_{\bar{z}_0} = (\bar{x}_0 \vee z_1x_1) \wedge (\bar{z}_1x_0 \vee \bar{x}_1x_0) = 0$$

Now the resulting $\underline{z}_0$ can be calculated:

$$\begin{aligned}
\underline{\delta}_{z_0}(z,x) &= {}_{z_0}\delta_{z_0} \wedge_{\bar{z}_0} \delta_{z_0} = \bar{x}_0 \vee z_1x_1 \\
\overline{\underline{\delta}}_{\bar{z}_0}(z,x) &= {}_{z_0}\overline{\delta}_{\bar{z}_0} \wedge_{\bar{z}_0} \overline{\delta}_{\bar{z}_0} = \bar{z}_1x_0 \vee \bar{x}_1x_0 \\
\underline{z}_0 &:= \underline{\delta}_{z_0} + \neg(\overline{\underline{\delta}}_{\bar{z}_0}) \\
\underline{z}_0 &:= \bar{x}_0 \vee z_1x_1 + \neg(\bar{x}_1x_0 \vee \bar{z}_1x_1)
\end{aligned}$$

3.2.2 Design of $\underline{\delta}_{z_1}$

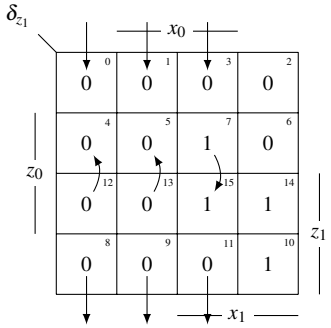


Figure 24: KV-diagram of δ_{z_1}

For $\underline{\delta}_{z_1}$ the same approach applies. First the KV-diagram, see figure 24, will be compacted:

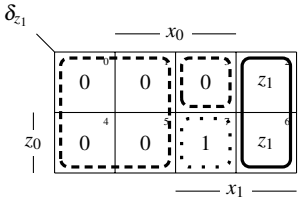


Figure 25: compacted KV-diagram of δ_{z_1}

Now the functions $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{\bar{z}_1}$ can be deduced:

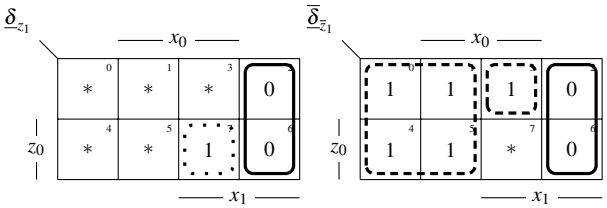


Figure 26: KV-diagrams for $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{\bar{z}_1}$

First δ_{z_1} is checked for instable parts:

$$\bar{z}_1\delta_{z_1} \wedge_{z_1} \overline{\delta}_{\bar{z}_1} = z_0x_1x_0 \wedge (\bar{z}_0x_0 \vee \bar{x}_1) = 0$$

No instable parts, so the RS-buffer can be coded for δ_{z_1} by determining the functions $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{\bar{z}_1}$:

$$\begin{aligned}
\underline{\delta}_{z_1}(z,x) &= {}_{z_1}\delta_{z_1} \wedge_{\bar{z}_1} \delta_{z_1} = z_0x_1x_0 \\
\overline{\underline{\delta}}_{\bar{z}_1}(z,x) &= {}_{z_1}\overline{\delta}_{\bar{z}_1} \wedge_{\bar{z}_1} \overline{\delta}_{\bar{z}_1} = \bar{x}_1 \vee \bar{z}_0x_0 \\
\underline{z}_1 &:= \underline{\delta}_{z_1} + \neg(\overline{\underline{\delta}}_{\bar{z}_1}) \\
\underline{z}_1 &:= z_0x_1x_0 + \neg(\bar{x}_1 \vee \bar{z}_0x_0)
\end{aligned}$$

3.2.3 Design of λ

The KV-diagram of λ can be seen in 27. Because of the two-dimensional state, λ has to be divided into eight parts. They exist as follows:

(z_1, z_0, y)	$(z_1, z_0, \bar{y})$
$(001) = A$	$(001) = \overline{E}$
$(011) = B$	$(011) = \overline{F}$
$(101) = C$	$(101) = \overline{G}$
$(111) = D$	$(111) = \overline{H}$

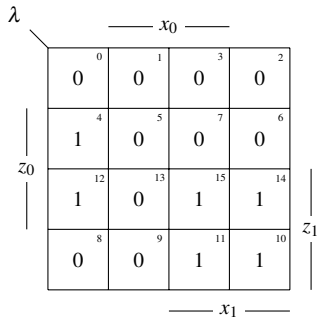


Figure 27: KV-diagram of λ

The parts of λ are set up in the KV-diagram, see figure 28.

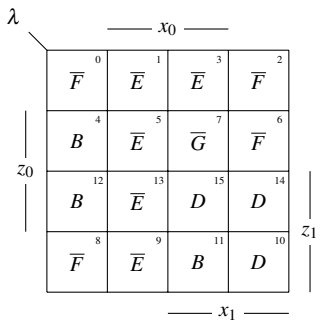


Figure 28: Parts of λ

The output functions y and $\bar{y}$ can be inferred from this:

$$
\begin{aligned}
y &= A \vee B \vee C \vee D \\
&= z_0\bar{x}_1\bar{x}_0 \vee z_1\bar{z}_0x_1x_0 \vee z_1x_1\bar{x}_0 \vee z_1z_0x_1 \\
&= z_1x_1 \vee z_0\bar{x}_1\bar{x}_0 \\
\bar{y} &= \overline{E} \vee \overline{F} \vee \overline{G} \vee \overline{H} \\
&= \bar{x}_1x_0 \vee \bar{z}_1\bar{z}_0x_0 \vee \bar{z}_0\bar{x}_0\bar{x}_1 \vee \bar{z}_1x_1\bar{x}_0 \vee \bar{z}_1\bar{z}_0\bar{x}_0 \vee \bar{z}_1z_0x_1x_0 \\
&= \bar{x}_1x_0 \vee \bar{z}_1x_1 \vee \bar{z}_0\bar{x}_1
\end{aligned}
$$

The resulting automaton can be seen in 30. As you can see in table 3, situation 7 leads to a cycle, by stepping to configuration 11, which steps back to configuration 7, and so on. So to have a state stable design in multistate machines, it is also necessary to regard the transitions between the states. A rule has to be found, to filter these race conditions. In the following transformation the auomaton of subsection b is stabilized, by changing δ_{z_1}.

3.3 Stable 2-dimensional example

The state transfer function δ_{z_1} has now been changed, to make it a stable automaton. For $x = (x_1, x_0)$, $y = (y)$ and $z = (z_1, z_0)$ the transformation of the Mealy machine $(X, Y, Z, \delta, \lambda)$ with the state transfer functions and output function

$$\delta_{z_0}(z,x) = \bar{x}_0 \vee z_1x_1 \qquad \delta_{z_1}(z,x) = z_0x_1x_0 \vee z_1x_1$$

$$\lambda(z,x) = z_1x_1 \vee z_0\bar{x}_1\bar{x}_0$$

will be executed. The machine before the transformation is outlined in figure 29.

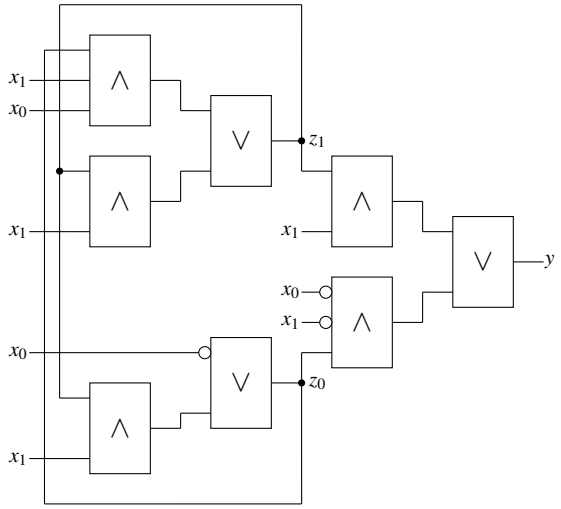


Figure 29: Automaton before the transformation

3.3.1 Design of $\underline{\delta}_{z_0}$

The KV-diagram of figure 31 will be lossless compacted given by the pointers in the diagram to the diagram of figure 32.

#	z_1	z_0	x_1	x_0	δ_{z_1}	δ_{z_0}	$\underline{\delta}_{z_1}$	$\overline{\underline{\delta}}_{\bar{z}_1}$	$\underline{\delta}_{z_1}$	$\underline{\delta}_{z_0}$	$\overline{\underline{\delta}}_{\bar{z}_0}$	$\underline{\delta}_{z_0}$	λ	A	B	C	D	$\overline{E}$	$\overline{F}$	$\overline{G}$	$\overline{H}$	y	$\bar{y}$	y
0	0	0	0	0	0	1	*	1	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
1	0	0	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
2	0	0	1	0	0	1	*	*	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
3	0	0	1	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
4	0	1	0	0	0	1	*	1	0	1	*	1	1	*	1	*	*	*	*	*	*	1	*	1
5	0	1	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
6	0	1	1	0	0	1	*	*	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
7	0	1	1	1	1	0	1	*	1	*	1	0	0	*	*	*	*	*	*	1	*	*	1	0
8	1	0	0	0	0	1	*	1	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
9	1	0	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
10	1	0	1	0	1	1	*	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1
11	1	0	1	1	0	1	*	1	0	1	*	1	1	*	1	*	*	*	*	*	*	1	*	1
12	1	1	0	0	0	1	*	1	0	1	*	1	1	*	1	*	*	*	*	*	*	1	*	1
13	1	1	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
14	1	1	1	0	1	1	*	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1
15	1	1	1	1	1	1	1	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1

Table 3: Truth table of the transformed automaton

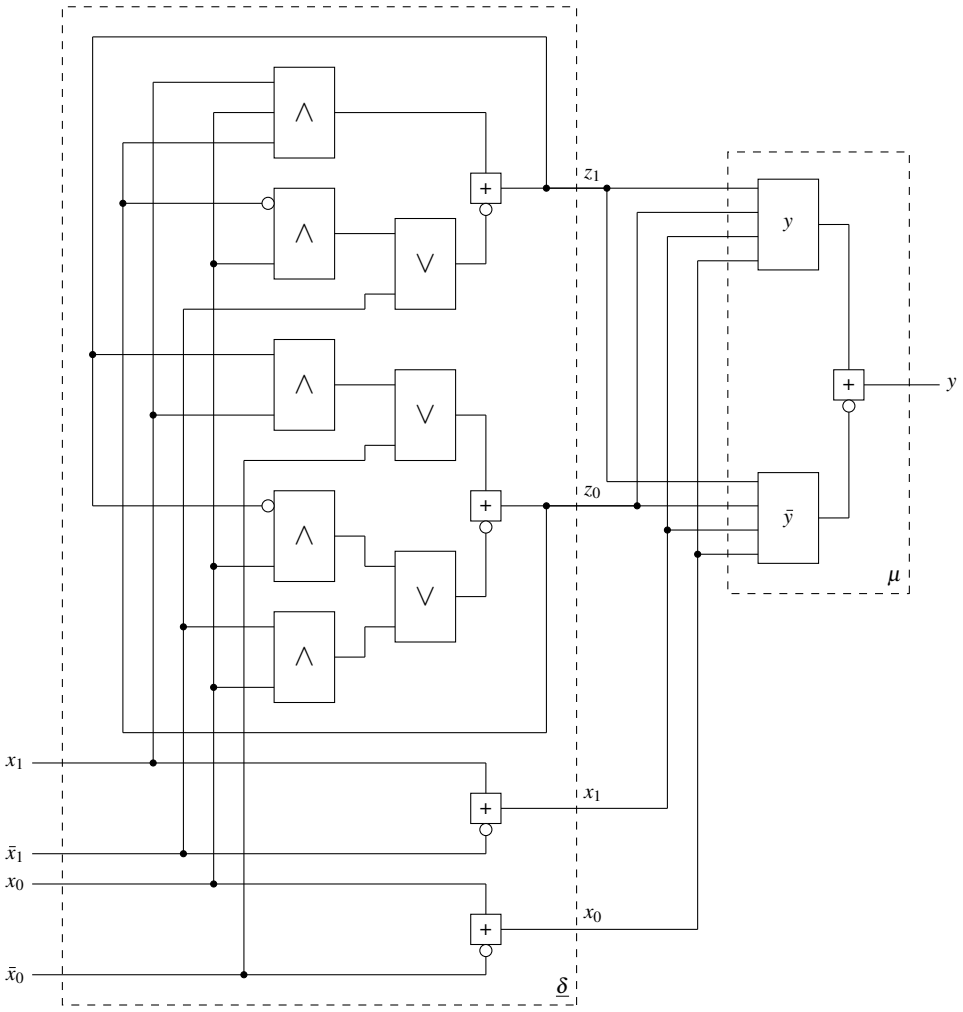


Figure 30: Resulting automaton

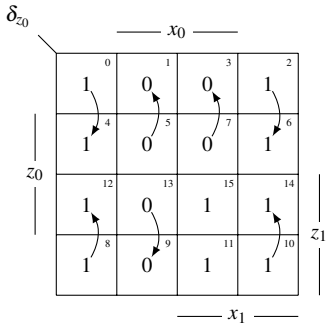


Figure 31: KV-diagram of δ_{z_0}

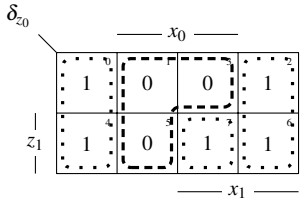


Figure 32: compacted KV-diagram of δ_{z_0}

With the compacted KV-diagram the RS-buffer for δ_{z_0} can be coded by specifying the functions $\underline{\delta}_{z_0}$ and $\overline{\delta}_{\bar{z}_0}$, see figure 33.

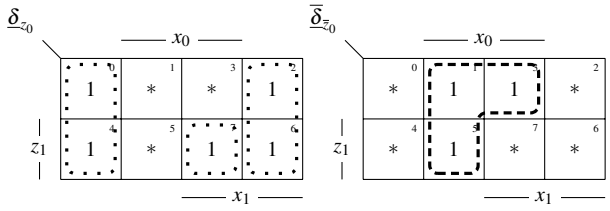


Figure 33: $\underline{\delta}_{z_0}$ and $\overline{\delta}_{\bar{z}_0}$

State instable parts are not allowed in the state transfer function:

$$_{\bar{z}_0}\delta_{z_0} \wedge {}_{z_0}\overline{\delta}_{\bar{z}_0} = (\bar{x}_0 \vee z_1 x_1) \wedge (\bar{z}_1 x_0 \vee \bar{x}_1 x_0) = 0$$

Now the resulting $\underline{z}_0$ can be calculated:

$$\underline{\delta}_{z_0}(z,x) = {}_{z_0}\delta_{z_0} \wedge {}_{\bar{z}_0}\delta_{z_0} = \bar{x}_0 \vee z_1 x_1$$
$$\overline{\delta}_{\bar{z}_0}(z,x) = {}_{z_0}\overline{\delta}_{\bar{z}_0} \wedge {}_{\bar{z}_0}\overline{\delta}_{\bar{z}_0} = \bar{z}_1 x_0 \vee \bar{x}_1 x_0$$
$$\underline{z}_0 := \underline{\delta}_{z_0} + \neg(\overline{\underline{\delta}}_{\bar{z}_0})$$
$$\underline{z}_0 := \bar{x}_0 \vee z_1 x_1 + \neg(\bar{x}_1 x_0 \vee \bar{z}_1 x_1)$$

3.3.2 Design of $\underline{\delta}_{z_1}$

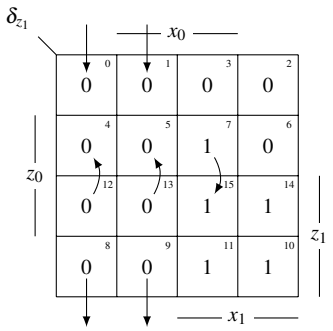


Figure 34: KV-diagram of δ_{z_1}

For $\underline{\delta}_{z_1}$ the same approach applies. First the KV-diagram, see figure 34, will be compacted:

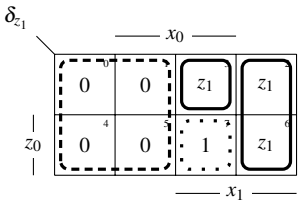


Figure 35: compacted KV-diagram of δ_{z_1}

Now the functions $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{\bar{z}_1}$ can be deduced:

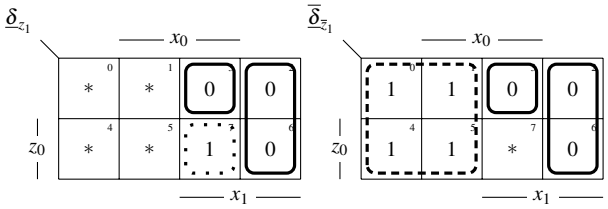


Figure 36: KV-diagrams for $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{\bar{z}_1}$

First δ_{z_1} is checked for instable parts:

$$_{\bar{z}_1}\delta_{z_1} \wedge _{z_1}\overline{\delta}_{\bar{z}_1} = z_0x_1x_0 \wedge \bar{x}_1 = 0$$

No instable parts, so the RS-buffer can be coded for δ_{z_1} by determining the functions $\underline{\delta}_{z_1}$ and $\overline{\underline{\delta}}_{z_1}$:

$$\underline{\delta}_{z_1}(z,x) = _{z_1}\delta_{z_1} \wedge _{\bar{z}_1}\delta_{z_1} = z_0x_1x_0$$
$$\overline{\underline{\delta}}_{\bar{z}_1}(z,x) = _{z_1}\overline{\delta}_{\bar{z}_1} \wedge _{\bar{z}_1}\overline{\delta}_{\bar{z}_1} = \bar{x}_1$$
$$\underline{z}_1 := \underline{\delta}_{z_1} + \neg(\overline{\underline{\delta}}_{\bar{z}_1})$$
$$\underline{z}_1 := z_0x_1x_0 + \neg(\bar{x}_1)$$

3.3.3 Design of λ

The KV-diagram of λ can be seen in 37. Because of the two-dimensional state, λ has to be divided into eight parts. They exist as follows:

(z_1,z_0,y)	$(z_1,z_0,\bar{y})$
$(001) = A$	$(001) = \overline{E}$
$(011) = B$	$(011) = \overline{F}$
$(101) = C$	$(101) = \overline{G}$
$(111) = D$	$(111) = \overline{H}$

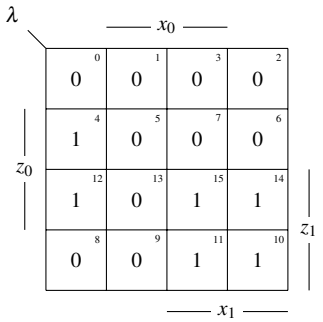


Figure 37: KV-diagram of λ

The parts of λ are set up in the KV-diagram, see figure 38.

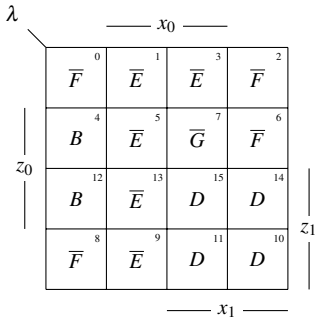


Figure 38: Parts of λ

The output functions y and $\bar{y}$ can be inferred from this:

$$\begin{aligned} y &= A \vee B \vee C \vee D \\ &= z_0\bar{x}_1\bar{x}_0 \vee z_1x_1 \\ \bar{y} &= \overline{E} \vee \overline{F} \vee \overline{G} \vee \overline{H} \\ &= \bar{x}_1x_0 \vee \bar{z}_1\bar{z}_0x_0 \vee \bar{z}_0\bar{x}_1\bar{x}_0 \vee \bar{z}_1x_1\bar{x}_0 \vee \bar{z}_1z_0x_1x_0 \\ &= \bar{x}_1x_0 \vee \bar{z}_1x_1 \vee \bar{z}_0\bar{x}_1 \end{aligned}$$

4 Test of Theory

To confirm the theory presented in the article, the state transfer function of the one dimensional example was analyzed making it a Medvedev. To simulate the circuit it was modelled with NAND gates and simulated in LT-Spice. The modeled NAND structure then was realized on the circuit board and the output was captured via an oszilloscope. The Medvedev machine $(X,Y,Z,\delta,1)$ with $x = (x_1,x_0)$, $y = (y)$ and $z = (z)$ with the state transformation function and output function be given:

The table and circuit:

#	z_1	z_0	x_1	x_0	δ_{z_1}	δ_{z_0}	$\underline{\delta}_{z_1}$	$\overline{\underline{\delta}}_{\bar{z}_1}$	$\underline{\delta}_{z_1}$	$\underline{\delta}_{z_0}$	$\overline{\underline{\delta}}_{\bar{z}_0}$	$\underline{\delta}_{z_0}$	λ	A	B	C	D	$\overline{E}$	$\overline{F}$	$\overline{G}$	$\overline{H}$	y	$\bar{y}$	y
0	0	0	0	0	0	1	*	1	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
1	0	0	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
2	0	0	1	0	0	1	*	*	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
3	0	0	1	1	0	0	*	*	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
4	0	1	0	0	0	1	*	1	0	1	*	1	1	*	1	*	*	*	*	*	*	1	*	1
5	0	1	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
6	0	1	1	0	0	1	*	*	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
7	0	1	1	1	1	0	1	*	1	*	1	0	0	*	*	*	*	*	*	1	*	*	1	0
8	1	0	0	0	0	1	*	1	0	1	*	1	0	*	*	*	*	*	1	*	*	*	1	0
9	1	0	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
10	1	0	1	0	1	1	*	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1
11	1	0	1	1	1	1	*	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1
12	1	1	0	0	0	1	*	1	0	1	*	1	1	*	1	*	*	*	*	*	*	1	*	1
13	1	1	0	1	0	0	*	1	0	*	1	0	0	*	*	*	*	1	*	*	*	*	1	0
14	1	1	1	0	1	1	*	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1
15	1	1	1	1	1	1	1	*	1	1	*	1	1	*	*	*	1	*	*	*	*	1	*	1

Table 4: Truth table of the transformed automaton

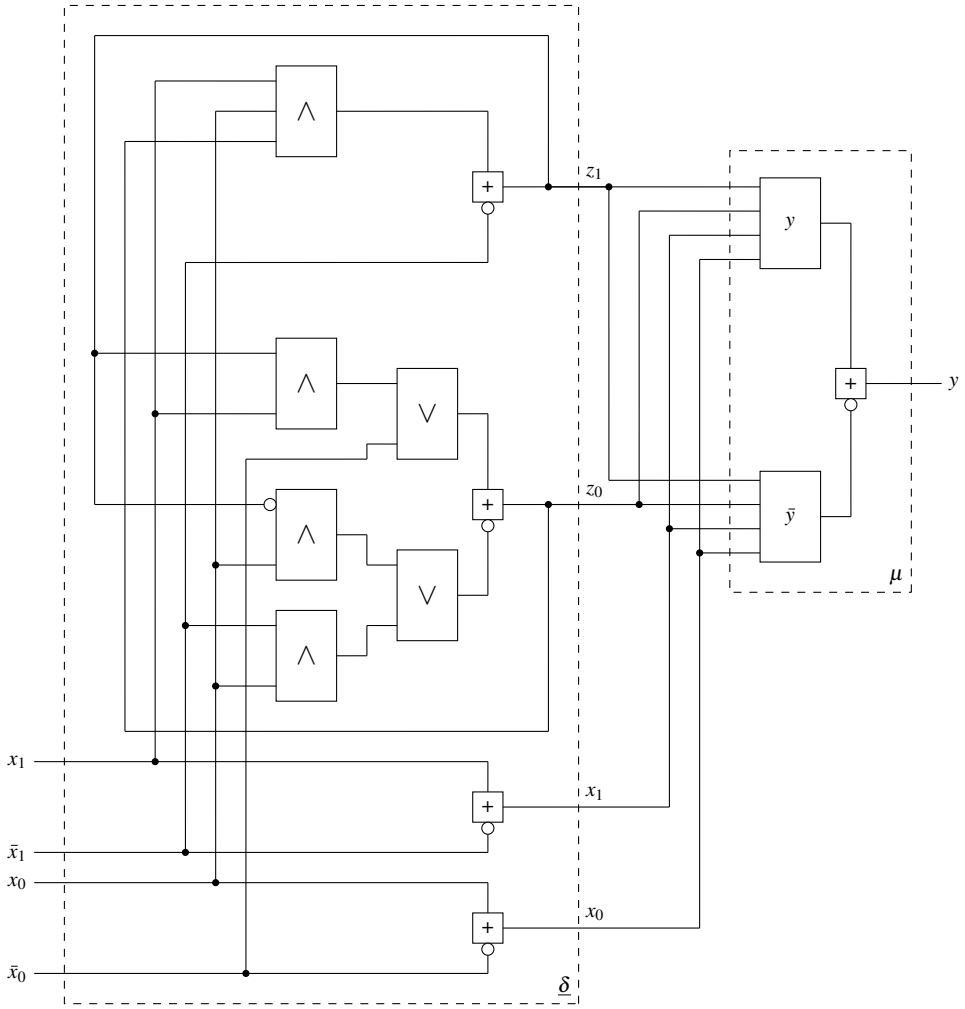


Figure 39: Resulting automaton

$$\delta(z,x) = zx_0\bar{x}_1 \vee x_1$$

The KV-Diagramm:

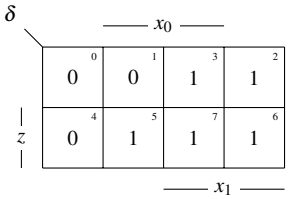


Figure 40: KV-diagram of δ

The table:

z	x_1	x_0	δ
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

Table 5: Truth table of the automaton

The graph of the automaton:

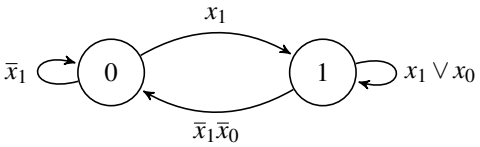


Figure 41: Graph of the automaton

NAND gates have been used to realize the structure. Therefore the DNF was transformed into a NAND strucutre by twice negating the DNF.

$$\delta(z,x) = \overline{\overline{zx_0\bar{x}_1 \vee x_1}} = \overline{\overline{zx_0\bar{x}_1} \wedge \bar{x}_1}$$

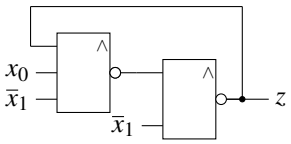


Figure 42: machine before the transformation

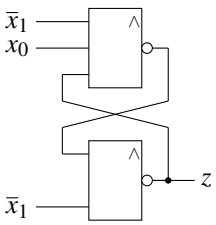


Figure 43: Different view of the automaton

Figure 43 shows two asynchronous feedbacks, which can lead to a race. On transition of x_1 from logical 1 to 0, while x_0 stays on logical 1, the feedback, which is first on logical 0, will make the other to a stable logical 1. Therefore on this transition, it can not be forecasted, whether the output will be on logical 1 or 0.

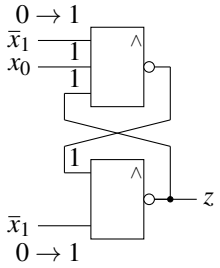


Figure 44: Race condition

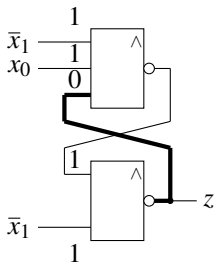


Figure 45: Race z = 0

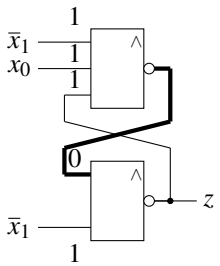


Figure 46: Race z =1

To simulate the race condition, the circuit was modelled and simulated in LT Spice. The simulation shows a race.

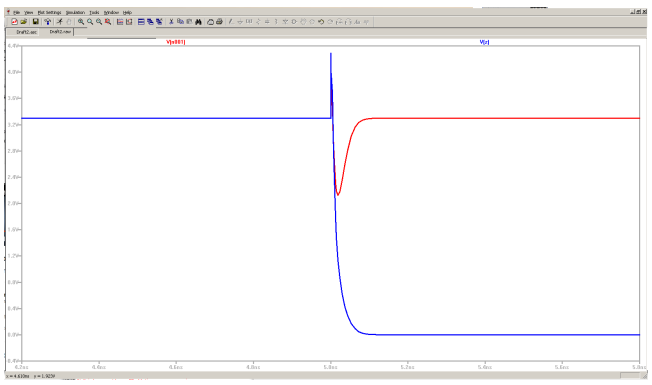


Figure 47: Simulated race situation

The circuit was realized in reality and the race was captured via an oscilloscope.

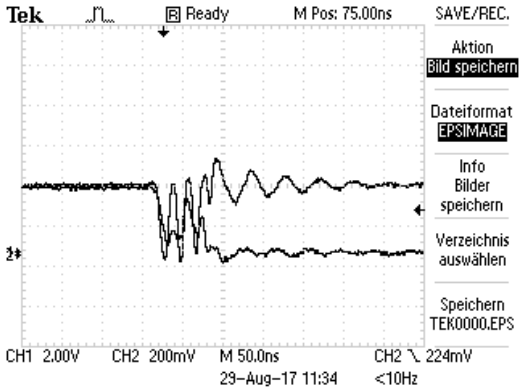


Figure 48: Race situation in reality

Then the transformation, which was presented in this journal, was made leading to the following structure, which does not have any contending feedbacks anymore. The only feedback this structure needs is nested in the RS-Buffer.

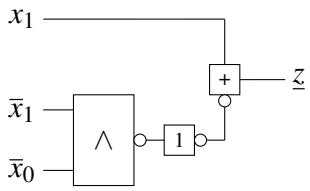


Figure 49: Transformed Moore

This structure does not show a race problem. To prove this statement the race condition of the old automaton has been set to the inputs. In figure 50 the state z and the rising edge of $\bar{x}_1$ is shown. It's clearly to see that the state doesn't change, not even a fluctuation of the signal can be recognized. Therefore the transformed machine is determined and stable.

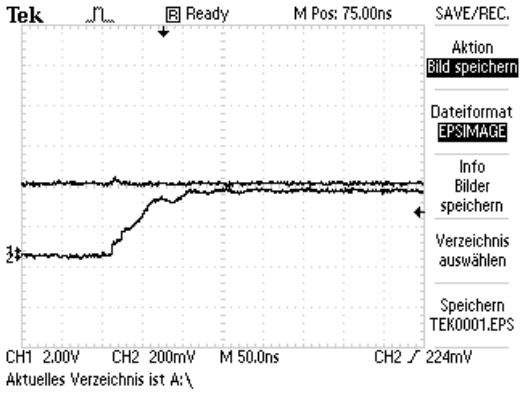


Figure 50: Race condition applied to the transformed automaton

5 Conclusion and Future Work

An asynchronous function stable Mealy was transformed into a state stable Moore under the use of the RS-buffer with suitable coding. The transformation provides general asynchronous design rules, e.q. the isolation of critical situations via the hold function. Due to the stabilisation by RS-buffer with the hold function, no races will be propagated through the circuit. In this journal three examples have been presented, a 1-dimensional and two 2-dimensional. Although the stable 2-dimensional transformation has produced an asynchronous feedback, this feedback leads either to a confirmation or a hold of the RS-buffer. The feedback has to be seen as a starting condition which gets triggered by the input and a change of the feedback will only affect the RS-buffer if the input changes. This design guideline can be used for all circuit designs, to create a stable, reliable system. The goal of the design is to remove the outer feedback and nest it in the RS-buffer. The inputs have been coded as pseudo states and stabilized via the RS-buffer. The output function has also been designed in dual-rail. This leads to a circuit, that looks at first sight as a combinational circuit. With this method you can simplify complex automata and combine them. The results have to be simulated to get a detailed look at the behavior of the transformed Moore. As it was shown, one dimensional machines get stabilized by this design, but for more dimensional machines it is necessary to design stable functions and regard the transitions between the states, not only concentrating on one state. So a rule has to be found, which can check for instabilities between states and stabilize machines in these constellations. Additionally the formula for the Transformation can be programmed, to create an automated process of transformation. Additionally the output function λ has to be checked for functional Hazards and addapted to guarantee that they are no longer existing.

References

[1] Birtwhistle, G.: Davis, A.: Asynchronous Digital Circuit Design, Springer London, 1995

[2] Van Berkel, K.: Burgess, R.: Kessels, J.L.W.: Peeters, A.: Roncken, M.: Schalij, F.: A fully asynchronous low-power error corrector for the DCC player, IEEE J. Solid-State Circuits, vol. 29, Dec. 1994. - P. 1429-1439

[3] Uygur, G.: Sattler, S.M.: A Real-World Model of Partially Defined Logic. 12th International Workshop on Boolean Problems 2016, Freiberg

[4] Mokhov, A.: Khomenko, V.: Sokolov, D.: Yakovlev, A.: On Dual-Rail Control Logic for Enhanced Circuit Robustness, ACSD '12 Proceedings of the 2012, 12th International Conference on Application of Concurrency to System Design, Pages 112-121, June 27 - 29, 2012, IEEE Computer Society Washington, DC, USA

[5] Shirvani, P.: Mitra, S.: Ebergen, J.: Roncken,M.: DUDES: a fault abstraction and collapsing framework for asynchronous circuits, in Advanced Research in Asynchronous Circuits and Systems (ASYNC), Proceedings, Sixth International Symposium on, 2000, pp. 73–82

[6] Özgül, M.: Deeg, F.: Sattler, S.M.: Mealy-to-Moore Transformation, 20th IEEE International Symposium on DDECS 2017, Dresden

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 175-185 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

Data Exchange Strategies for Aggregating Geographical Distribution of Demands for Location-Dependent Information Using Soft-State Sketches in VANETs

Akihiro Yamada[*,1], Susumu Ishihara[2]

[1] *Graduate School of Integrated Science and Technology, Shizuoka University, 4328011, Japan*

[2] *College of Engineering, Academic Institute, Shizuoka University, Hamamatsu, 4328011, Japan, ishihara.susumu@shizuoka.ac.jp*

ARTICLEINFO

Article history:
Received: 30 October, 2017
Accepted: 02 December, 2017
Online: 14 December, 2017

Keywords:
VANET
Data Dissemination
Aggregation

ABSTRACT

When we drive a car, we often want to know the circumstances of roads between the current location and the destination. Ishihara et al. proposed a Vehicular ad-hoc NETworks (VANET) -based information sharing system, called the Real-Time Visual Car Navigation System, in which a driver can obtain Location-Dependent Information (LDI), such as photographs or videos of his/her Point of Interest (PoI), by telling the onboard device, e.g., a car navigation device, the PoI. The simplest way to provide LDI to vehicles is to flood the LDI to all vehicles, but unneeded LDI may be disseminated, and network resources may be wasted. To disseminate LDI to many vehicles that demand it with low communication traffic, we have proposed a data dissemination scheme based on Demand maps (Dmaps). In this scheme, each vehicle has a Dmap, a data set representing the geographical distribution of the strength of demands for LDI. To keep the Dmap up to date, each vehicle exchanges a subset of data constituting a Dmap (Dmap Information: DMI) with other vehicles. Vehicles preferentially send new LDI or forwarded LDI strongly demanded to the area where the LDI is frequently demanded based on Dmaps. If vehicles frequently send DMI, the accuracy of the Dmap becomes high, however, communication traffic becomes large. In this paper, we propose strategies for controlling the frequency of sending DMI and strategies for selecting DMI to be sent to improve the Dmap accuracy with low communication traffic. Simulation results show that one of the proposed strategies, the LDI Transfer Zone (LTZ) strategy, achieves highly accurate Dmaps with a small amount of communication traffic.

1 Introduction

If a driver can visually know the current circumstances of roads on a route to his/her destination, such as traffic jams, roadwork, accidents, and beautiful sights, he/she may select a better (faster, easier, or more fun) route to the destination. Currently, there are some applications that inform a driver of traffic information to make driving more comfortable, such as Google maps and ETC2.0 [1] in Japan, etc. In addition, Pioneer has provided a cellular network-based system that allows a driver to visually know the current circumstances of roads predefined by the system [2]. However, to the best of our knowledge, there is no application that visually informs a driver of the current circumstances of roads that are designated by the driver.

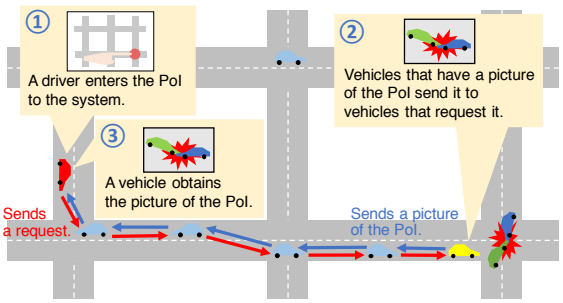


Figure 1. Real-Time Visual Car Navigation System

[*]Corresponding Author: Akihiro Yamada, Shizuoka University, 3-5-1 Johoku, Naka-ku Hamamatsu, Japan, akihiro@ishilab.net

Ishihara et al. proposed a Vehicular Ad-hoc NETworks (VANET)-based road traffic information sharing system, called the *Real-Time Visual Car Navigation System* [3–5]. This system allows driver to obtain photographs/videos of his/her Point of Interest (PoI) that is inputted to an onboard device, e.g., a car navigation device. In this system, when a driver inputs his/her PoI to an onboard device, the vehicle sends a request message to vehicles in the vicinity of the PoI. Then, vehicles that have a photograph/video of the PoI send it to the requesting vehicle (Figure 1).

In this system, we assume that vehicles use only VANET to send request messages to vehicles in the vicinity of a PoI and to send photographs/videos of PoI to requesting vehicles. If we use a simple pull-based strategy for sending the photographs/videos in the *Real-Time Visual Car Navigation System*, a requesting vehicle sends request messages to vehicles in the vicinity of a PoI, and vehicles that received a request message respond by sending photographs/videos of the PoI to the requesting vehicle. With this strategy, it is difficult for vehicles to certainly send the photographs/videos to the requesting vehicles with low latency because the connectivity between vehicles is intermittent due to the high mobility of vehicles. Furthermore, if multiple vehicles that are close together demand photographs/videos of the same PoI at the same time, the same data and the same request messages are repeatedly sent to the same location at the same time. Consequently, network resources may be wasted. Cooperative caching techniques such as ones proposed in [6] can be used to mitigate this problem. Another problem of pull-based schemes is that it is difficult to constantly deliver newly generated Location-Dependent Information (LDI) of PoI items to vehicles having interest in the PoI.

On the other hand, if all vehicles that have photographs/videos of the PoI send them based on a simple push-based strategy for certainly sending the photographs/videos to requesting vehicles, all vehicles send the photographs/videos whenever vehicles receive or generate them. Thus, unneeded data may be disseminated and limited network resources are wasted.

To ensure that newer LDI is sent to requesting vehicles with low communication traffic, Ishihara et al. have proposed a data dissemination scheme based on *Demand maps* (Dmaps) [4]. A Dmap is a data set representing the geographical distribution of the strength of demands for LDI. This scheme assumes that every vehicle has a Dmap. Vehicles grasp the geographical distribution of the strength of demands for LDI using a Dmap, and they send LDI to areas where it is strongly demanded via an efficient route. To keep Dmaps on vehicles up to date, each vehicle broadcasts several data constituting its Dmap (Dmap Information: DMI) with a beacon that it sends periodically. When vehicles receive other DMI, they update their Dmap by merging the received DMI with their own.

Since vehicles send LDI in accordance with Dmaps, the accuracy of Dmaps, i.e. the reproducibility of real geographical distribution of the strength of demands in Dmaps, should be high. Each vehicle should frequently update its Dmap by frequently exchanging DMI with other vehicles to make the accuracy of Dmaps high. However, if each vehicle frequently sends DMI, communication traffic becomes high. Therefore, strategies to increase the accuracy of Dmaps and control communication traffic are needed. In [4], Ishihara, et al. design the basic specification of Dmap-based data dissemination, however, they did not design the details of how to send DMI.

In this paper, to solve the tradeoff between Dmap accuracy and communication traffic, we propose strategies that control the frequency of sending DMI and strategies that each vehicle selects DMI to be sent. In addition, we evaluate the accuracy of Dmaps and communication traffic in the case that vehicles use the proposed strategies through simulations.

The contributions of this paper are summarized as follows.

1. We propose a strategy for controlling the frequency of broadcasting DMI. In this strategy, each vehicle determines whether it sends DMI based on the frequency of updating its DMI.

2. We propose three strategies for selecting DMI to be sent: i) the Recently Update First (RUF) strategy, which considers how new each DMI is, ii) the LDI Transfer Zone-based (LTZ) strategy, which considers the relation between each PoI and each location where requests occur, and iii) a hybrid strategy.

3. We evaluated the effectiveness of the proposed strategies through simulations and confirmed that the LTZ strategy most effectively reduces communication traffic and makes the accuracy of Dmaps high.

This paper is an extension of work originally presented in 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA) [7]. In this paper, we conduct simulations with more appropriate parameters, and provide a more detailed explanation of the proposed strategies and Dmap-based data dissemination. In addition, we improve the implementation of our proposed strategies in an existing simulator, and obtain more reliable simulation results. The description of related work is updated.

The remainder of this paper is composed as follows. In the next section, we review related work. In Section 3, we recapitulate the data dissemination scheme based on Dmaps and introduce the data structure of Dmap based on soft-state sketches. In Section 4, we propose strategies for sending DMI to improve the accuracy of Dmaps with low communication traffic. Then, in Section 5, we evaluate the effectiveness of the proposed strategies. Finally, we conclude this paper in Section 6.

2 Related Work

In Dmap-based data dissemination, we assume that a vehicle that has generated LDI appends the location where it generated the LDI. In addition, vehicles can know locations where the LDI is demanded based on their Dmap. Vehicles carry out the routing of LDI by referring to a location where the LDI is generated and locations where the LDI

is demanded. This concept is similar to Named Data Networking (NDN). In NDN, vehicles give a name, e.g. a location where data item is generated, to each data item and deliver the data item to vehicles requesting the data item associated with that name.

In [8], Grassi et al. proposed Navigo, an interest forwarding scheme in NDN. In Navigo, each vehicle that has received or generated an interest forwards it based on the shortest path between the location where it exists and the location where many vehicles having the demanded data exist. This scheme aims to deliver interests to locations where many vehicles having the required data are. In Navigo, vehicles send requests (interests) to many vehicles having the required data by sending the requests to the locations of these vehicles. In the Dmap-based scheme, vehicles register their requests to their Dmap instead of sending requests to the PoIs. In addition, vehicles having LDI can know not only a request for the LDI, but also the geographical distribution of the strength of demands for the LDI, and utilize this information for efficiently routing the LDI.

In [9], Ahmed et al. proposed CODIE as a data dissemination scheme using NDN. In CODIE, vehicles that forward an interest append the hop count of the interest from the requesting vehicle, and vehicles having the required data forward the data in accordance with the hop count. In this scheme, the same LDI may be disseminated to the same region at the same time, and network resources may be wasted. On the other hand, in the Dmap-based scheme, each vehicle grasps regions where each LDI is highly desired as well as its distribution. Therefore vehicles may be able to disseminate demanded LDI to many vehicles with a small number of the LDI transfers. In addition, Bian et al. introduced the NDN architecture to VANETs and proposed forwarding and caching schemes [10].

In Dmap-based data dissemination, requesting vehicles register their requests to their Dmaps. In the publish/subscribe (pub/sub)-based approach [14], e.g. [15] and [16] , requesting vehicles register their interests and their IDs to broker nodes. We regard that Dmap-based data dissemination is a kind of pub/sub based approach. A Dmap can be regarded as information on subscription and subscriber, and locations where requesting vehicles exist can be regarded as subscribers. In the Dmap-based scheme, each requesting vehicle does not register their IDs to a Dmap. However, locations where it requests LDI and locations where the LDI is generated are registered to a Dmap when it demands LDI. By disseminating LDI based on locations where requesting vehicles exist, the Dmap-based scheme can efficiently disseminate LDI to requesting vehicles compared with disseminating LDI based on IDs of requesting vehicles.

In our scheme, each vehicle sends DMI with a beacon that is sent periodically. In this paper, we discuss controlling the frequency of sending DMI. Many schemes for controlling the frequency of sending beacons have been proposed for sharing traffic information with low communication traffic. Dynamic beaconing (DynB) [11], UV-CAST [12], and DV-Cast [13] control the frequency of sending beacons adaptively to the density of vehicles and communication condition. In this paper, we propose a strategy that vehicles control the frequency of sending DMI adaptively to the relationship between the frequency of updating their DMI and the frequency that neighboring vehicles update their DMI and sharing a Dmap with low communication traffic. Strategies that use density of vehicles and communication condition, such as DynB, UV-CAST, and DV-Cast can be applied to our strategy.

Dmap-based data dissemination proposed in [4] and [17] can be categorized into a kind of data dissemination system for sharing Location-Dependent Information (LDI) in VANETs. Traffic information is typical LDI treated in VANETs. Various applications for sharing traffic information in VANETs, such as SOTIS [18], Trafficview [19], CASCADE [20], and Zone Flooding [21] have been proposed so far. They have a function for aggregating LDI obtained in a distributed manner. For example, in SOTIS, vehicles periodically broadcast a packet including traffic information, e.g. their location, velocity, and direction. When a vehicle receives information from other vehicles, it aggregates all the known information on each road segment to one average value. In those applications, unneeded information may be sent because vehicles broadcast traffic information without considering how strongly the traffic information is required. On the other hand, Dmap-based data dissemination aims to reduce redundant information dissemination by disseminating strongly required information in accordance with a Dmap.

In this work, a Dmap is represented with soft-state sketches proposed in [22]. A soft-state sketch is a data structure to stochastically estimate the number of unique data for aggregating data and reducing communication traffic. A soft-state sketch is represented by an array of non-negative values. In [22], soft-state sketches are used for counting the number of available parking spaces every load segment and sharing this information among vehicles. Vehicles use a set of soft-state sketches for estimating the number of available parking spaces in a load segment. In [22], each vehicle has a set of soft-state sketches every load segment. When each vehicle finds available parking spaces, it inputs the initial TTL of available parking spaces information to a soft-state sketch corresponding to the load segment where the vehicle exists. In addition, vehicles append sets of soft-state sketches to a beacon that is sent periodically, and update their information by merging their own soft-state sketches with received ones.

On the other hand, in this paper, we use soft-state sketches for counting the number of demands for LDI and sharing this information among vehicles. In our scheme, each vehicle has a set of soft-state sketches for every pair of each location where LDI is generated and each location where LDI is required. Then, each vehicle inputs the initial TTL to a soft-state sketch corresponding to its request when it requests LDI, and sends sets of soft state sketches as DMI with a beacon.

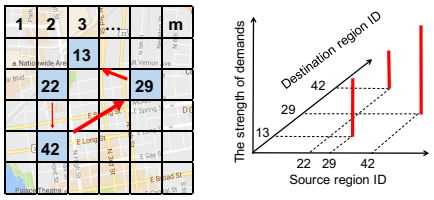
In [22], to reduce communication traffic for sending sets of soft-state sketches, a method for compressing soft-state sketches is proposed. However, we aim to reduce communication traffic by controlling the frequency of sending sets of soft-state sketches and selecting sets of soft-state sketches that vehicles should send.

3 A data dissemination scheme based on Demand map

In this section, we describe the details of the data dissemination scheme based on a Demand map (Dmap). In this scheme, each vehicle has a Dmap, which is a data set representing the geographical distribution of the strength of demands for LDI. Each vehicle forwards strongly demanded LDI items preferentially in accordance with its Dmap.

3.1 Demand map (Dmap)

A Dmap represents the geographical distribution of the strength of demands for LDI. We assume that the strength of demands for LDI is measured by the number of demands generated by vehicles. If multiple vehicles in geographically distant locations demand LDI of the same PoI, e.g. vehicles in the east end of a city and vehicles at the west end of the city require LDI of the city center, the routes for delivering the LDI to the vehicles in distant locations will be different. Thus, we divide an area into multiple regions, and a Dmap manages the strength of demands for LDI of each region every each region where demands are generated.



(a) Regions in a Dmap (b) Concept of a Dmap

Figure 2. Demand map

A Dmap is created as follows. Vehicles divide an area supported by their Dmap into multiple small regions and give a unique ID to each small region as shown in Figure 2 (a). We refer to a region where a requesting vehicle exists as a source region, and refer to a region where the requested LDI is generated as a destination region. For instance, in Figure 2, if a vehicle in region 29 requires LDI of region 13, we call region 29 a source region and call region 13 a destination region. A Dmap records the number of demands, i.e. the strength of demands, every pair of each source region and each destination region. For instance, in Figure 2 (a), if the start point of an arrow is a source region, and the end point of an arrow is destination region, and width of an arrow indicates the strength of demands corresponding to the arrow, a Dmap is represented as shown in Figure 2 (b).

3.2 How vehicles use a Demand map for disseminating Location-Dependent Information (LDI)

A Dmap on each vehicle is used for decision of routing of LDI. When a Dmap is updated or each vehicle generates new LDI or receives LDI, each vehicle refers to its Dmap and knows the number of demands per pair of a source region and a destination region. In addition, the vehicle refers to LDI time stamps, and it preferentially selects LDI that was required many times from sets of LDI that are current. If vehicles that are far from a source region send LDI, it may become outdated during delivery, even though LDI should be new. Therefore, vehicles that are far from the source region do not send LDI.

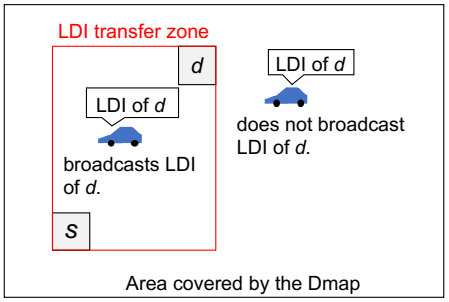
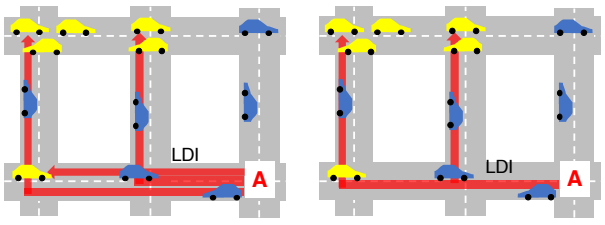


Figure 3. Disseminating LDI of s to d in accordance with a Dmap

In addition, LDI should be delivered from the destination region to source regions via a route as short as possible. To realize this, we defined a zone where LDI is delivered as an LDI Transfer Zone (LTZ). An LTZ is a rectangular zone that consists of a diagonal whose ends are a source region and a destination region (Figure 3). When a vehicle sends LDI, it determines whether it sends the LDI based on whether it is in the LTZ corresponding to the LDI. A vehicle sends an LDI item only when it is in the LTZ corresponding to the pair of the destination region and the source region of the LDI (Figure 3). If the vehicle is in one of the LTZs corresponding to LDI that is strongly demanded from many source regions, it aggressively broadcasts the LDI item.

LDI is delivered via many source regions. Thus, in the case that LDI is demanded from multiple source regions at the same time, if duplicate routes exist, LDI is not passed through the duplicate route multiple times, and redundant dissemination may be reduced compared with simple pull-based strategy. For instance, in Figure 4 (a), if yellow cars request LDI of A, the LDI is sent via routes represented by red arrows multiple times and duplicate dissemination occurs. However, in Figure 4 (b), if all cars have the same content of a Dmap, the LDI is sent via a route represented by red arrows once, then, duplicate disseminations do not occur.



(a) Simple pull-based strategy (b) Dmap-based strategy

Figure 4. Comparison of Dmap-based strategy and simple pull-based strategy

We assume that vehicles send new LDI demanded by many vehicles to the area where the demanding vehicles exist so that the LDI can be delivered via a route satisfying the aforementioned requirements. A routing strategy based on a propagation function proposed in [23] can be used for

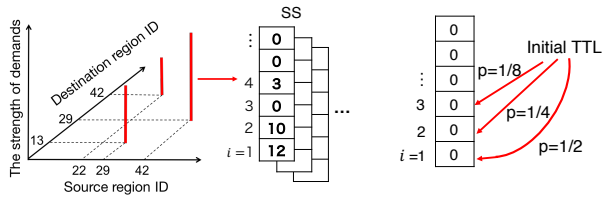
this purpose.

By this operation, vehicles that want the LDI will be able to obtain the LDI with a low data delivery delay and a small amount of traffic. For achieving high reliability and low data traffic for delivering LDI, the behavior of neighboring vehicles that have the same LDI and potentially broadcast the LDI should be considered. The detailed design of data forwarding algorithms based on a Dmap and behavior of the neighboring vehicles is beyond the scope of this paper.

3.3 Data structure of Demand map

One main scheme to create a Dmap is to collect all log data concerning demands that are generated by all vehicles, e.g. source region ID, destination region ID, car ID, and Time to Live (TTL) of demands, etc. However, if vehicles create their Dmap in accordance with this main scheme, all vehicles broadcast all log data concerning demands, and data size that vehicles have to broadcast and manage in their local storage increase in proportion to the number of demands and the number of vehicles.This leads to low scalability. To reduce the required communication traffic for maintaining Dmaps and the required storage on vehicles, a Dmap should be represented by a suitable data structure and needs algorithms to handle the data that can aggregate many demands generated by different vehicles in different locations.

Therefore, in [4] and [17], Ishihara et al. proposed to represent a Dmap by using soft-state sketches proposed in [22], a data structure to aggregate data and stochastically count the number of unique data. In [22], soft-state sketches are used for managing observation of free parking spaces in the same area by multiple vehicles. Meanwhile, we use soft-state sketches for counting the number of demands for the same pair of source region and destination region.



(a) Dmap with soft-state sketch (b) Putting TTL

Figure 5. Soft-state sketches in a Dmap

To manage a Dmap, each vehicle has a set of h ($h \geq 1$) sketches for a pair of source region and destination region (Figure 5 (a)). A sketch is an array of non-negative integers as shown in Figure 5 (a). Hereafter, we refer to a set of sketches as an SS for simplicity. Since an SS is used for representing DMI for a pair of source region and destination region, we will use these words interchangeably.

When a vehicle requires LDI of a PoI, it randomly selects a sketch from the SS corresponding to the pair of the vehicle's current region, i.e. the source region of the demand, and PoI, i.e. the destination region of the demand (Figure. 6). Then, it puts the initial TTL for the demand into i-th element from the smallest index of elements in the selected sketch with a probability of $1/2^i$ (Figure 5 (b)). The TTL value in each element is decremented as time progresses until the TTL value becomes 0. Since the probability that the TTL value is put into an element with a small index i is high and the probability that the TTL value is put into an element with a large index i is low, elements with a small index easily have non-zero values, but ones with a large index are hardly filled with a non-zero value. Therefore, the more demands occur, the more easily elements with large indexes are filled with non-zero values and the longer the run length of non-zero values from an element with the smallest index of a sketch is. Thus, the number of generating demands for a pair of source region and destination region, i.e. the strength of demands for the pair, is estimated based on the run length of non-zero values from an element with the smallest index of each sketch in an SS for the pair.

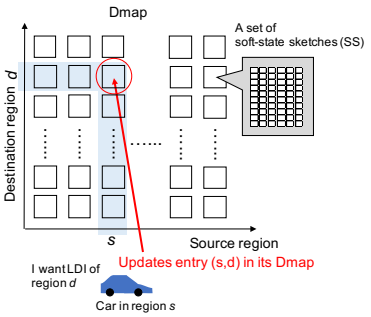


Figure 6. Behavior of a vehicle when it demands LDI

To keep Dmaps on vehicles up to date and share Dmaps among vehicles, each vehicle selects several SSs and sends them with a beacon that is sent periodically. When a vehicle receives SSs from another vehicle, it merges each of the received SSs with a corresponding SS in its Dmap. To merge the received SSs with the corresponding SSs in a local Dmap, the vehicle selects a larger value in the same element of the same index of sketches in each of received SSs and the corresponding SS in the local Dmap. Then the selected value is set to the element of the sketch of the SS in the receiver vehicle's Dmap as shown in Figure 7.

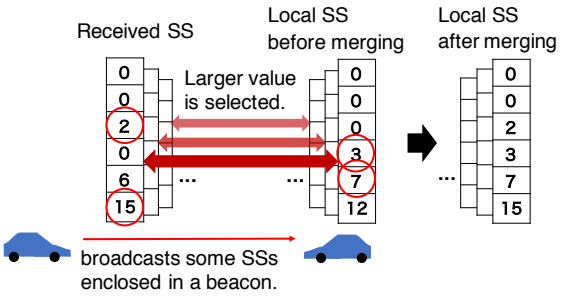


Figure 7. Merging with local SSs and other's

The number of demands for a pair of a source region and a destination region recorded in an SS consisting of m sketches $(S_1, \ldots, S_m)$ is estimated by the following formula presented in [22],

$$C(S_1, \ldots, S_m) := m \cdot \frac{2^{\sum_{j=1}^{m} Z(S_j)/m} - 2^{-\kappa \cdot \sum_{j=1}^{m} Z(S_j)/m}}{\varphi}, \quad (1)$$

with $\kappa \sim 1.75$, $\varphi \sim 0.77351$. S_j is a sketch in an SS, and m is the number of sketches composing an SS. $Z(S_j)$ is the run length of non-zero values from an element with the smallest index in S_j. In addition, the standard error of the estimate is $0.78/\sqrt{m}$. Therefore, the larger the number of sketches that compose an SS, the higher the accuracy of the estimate. However, data size that each vehicle has and sends increases in proportion to the number of sketches that compose an SS.

4 Strategies for efficiently sending DMI

In this section, we propose a strategy that controls the frequency of sending DMI and strategies that select DMI that vehicles should send for sending DMI with making the accuracy of Dmap high and making communication traffic low. To reflect current demands in a Dmap certainly, i.e. to improve the accuracy of a Dmap, it is desirable for each vehicle to frequently receive SSs for all pairs of all source regions and all destination regions from many vehicles and make its local Dmap up to date. However, if all vehicles frequently send all SSs that they have, the communication resource will be wasted because unneeded SSs may be sent and vehicles may send duplicate SSs. Therefore, we propose strategies for making the accuracy of Dmap high with low communication traffic.

4.1 Strategy controlling the frequency of sending DMI — Frequent strategy

We propose *selecting frequently updating vehicles* strategy, a strategy that controls the frequency of sending DMI (= SSs), to reduce the communication traffic for sending SSs without deteriorating the accuracy of Dmaps. Hereafter, we refer to this strategy as *Frequent* strategy for simplicity. Based on this strategy, when a vehicle sends a beacon, it determines whether it sends SSs with a beacon. The key idea of this strategy is that vehicles that frequently receive and update SSs may have SSs that certainly reflect the actual number of demands. On the other hand, vehicles that rarely receive and update SSs may have SSs that incorrectly reflect the actual number of demands. In addition, vehicles that frequently generate demands and update their SSs may have many SSs that are still not reflected in the SSs of other vehicles. Therefore the vehicles should send their SSs frequently. Therefore, in the Frequent strategy, vehicles that frequently update their SSs have a high priority to send SSs. Here, *updating* an SS means that one of the elements in a sketch of the SS is overwritten by putting a larger value due to a new demand generation or receiving a corresponding SS that has at least one larger value for any of elements in sketches in the receiver vehicle's SS from a neighboring vehicle.

In this strategy, each vehicle counts the total number of updates to its SSs during a certain past period, τ_u. We refer to the number as the Total Count of Updating Sketches (TCUS), for simplicity. In this strategy, vehicles whose TCUS is large are regarded as vehicles that frequently update their SSs. Vehicles whose TCUS is larger than one of their neighbor vehicles have a higher priority to send their SSs.

The operation of vehicles in this strategy is summarized follows. Each vehicle sends its TCUS with a beacon periodically. When a vehicle receives a beacon from another vehicle, it stores the received TCUS with the ID of the vehicle that has sent the TCUS. When a vehicle sends a beacon, it sorts TCUSs received from vehicles in the vicinity and its TCUS in descending order. Then, if its TCUS is in the top-k group of TCUSs, it decides to send SSs with a beacon (Figure 8). Otherwise, it does not send any SSs.For instance, as shown in Figure 8, if TCUS of car A is 4th from the largest TCUS and $k > 4$, car A sends SSs.

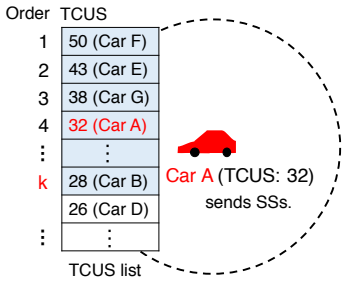


Figure 8. Frequent strategy

4.2 Strategies for selecting DMI that vehicles send

Each vehicle should select SSs to send so that Dmaps on vehicles, including itself, can achieve high accuracy with low communication traffic by eliminating redundant information that will not update the SSs of other vehicles. To achieve this, we propose three strategies, the Recently Updated First (RUF) strategy, the LDI Transfer Zone (LTZ) strategy, and the RUF + LTZ Hybrid strategy.

4.2.1 Recently Updated First (RUF) strategy

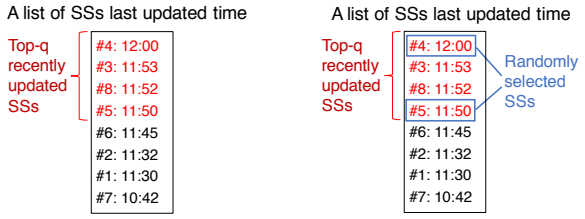
Recently updated SSs may have a large TTL value and may update SSs on vehicles that receive them, while SSs that are updated long ago may have a small TTL value and may not update SSs of the receiver vehicles. Sending SSs that will not update SSs of other vehicles wastes limited communication resources.

Therefore, in the RUF strategy, we give a high priority to recently updated SSs. When a vehicle updates an SS, it records the current time with a pair of a source region ID and a destination region ID. If a vehicle decides to send SSs with a beacon, it selects top-q recently updated SSs from its local SSs that have been updated so far (Figure 9 (a)).

However, if each vehicle selects SSs only based on the last update time of them, SSs that have the same pair of source region and destination region may frequently be sent and SSs that have other pairs will lose a chance to be sent. For instance, if an SS for a pair of source region s and destination region d is the most recently updated SS in a vehicle, the vehicle sends the SS. Then, when neighbor vehicles receive and update the SS, the SS for a pair of s and d becomes the most recently updated SS in the vehicles and the

vehicles will send it. The SS for a pair of s and d is repeatedly sent and SSs for other pair are rarely sent. Consequently, for some pairs of a source region and a destination region, the transition of the number of demands on a vehicle may not be reflected in SSs, and the Dmap accuracy may become low.

To provide the opportunity for not only the most recently updated SSs to be sent, in this strategy, each vehicle randomly selects up to n SSs from the top-q recently updated SSs and sends them (Figure 9 (b)).



(a) Selecting recently updated SSs (b) n SSs from the top-q SSs

Figure 9. RUF strategy

4.2.2 LDI Transfer Zone (LTZ) strategy

In Dmap-based data dissemination, we assume that vehicles send LDI for a pair of source region s and destination region d only when they are in the LTZ corresponding to s and d, as mentioned in Section 3.2. Therefore, vehicles in the LTZ corresponding to s and d should certainly know the number of demands for LDI for the pair of s and d. On the other hand, vehicles that are not in the LTZ do not have to know the number of demands for LDI. If the SSs for the pair of s and d are sent by only vehicles in the LTZ corresponding to s and d, the SSs are exchanged among only vehicles in the LTZ. As a result, the accuracy of the number of demands only for a pair of s and d composing an LTZ containing a vehicle holding the Dmap becomes high because noisy unneeded SSs are not sent and SSs that affect other vehicles SSs for a pair of s and d composing the LTZ are sent.On the other hand, exchanging SSs among vehicles outside the LTZ is not effective because vehicles do not need SSs that they will not use when they disseminate LDI.

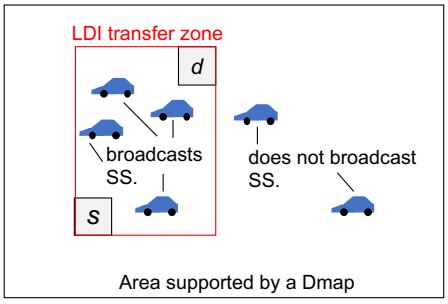


Figure 10. LTZ strategy

In the LTZ strategy, vehicles send SSs for the pair of s and d only when they are in the LTZ corresponding to s and d, and vehicles do not send the SSs when they are not in the LTZ. To this end, when each vehicle decides to send SSs, it first selects only SSs whose pair of source region s and destination region d satisfies $P \in A_{s,d}$. Here, P denotes the vehicle's current region, and $A_{s,d}$ denotes the LTZ for the pair of s and d (Figure 10). If the number of selected SSs is larger than the maximum number of SSs that can be included in a beacon packet, the vehicle randomly selects up to n SSs from the selected SSs.

4.2.3 RUF + LTZ Hybrid strategy

To leverage the advantages of both the RUF and LTZ strategies, we propose the RUF + LTZ Hybrid strategy. In this strategy, each vehicle sends more recently updated SSs only when it is in the LTZ corresponding to the SSs. Each vehicle selects and sends SSs as described below. If a vehicle v decides to send SSs with a beacon, it selects the top-q recently updated SSs from its local SSs that have been updated so far, like in the RUF strategy. Note that a vehicle selects SSs only based on whether SSs are more recently updated in this phase. We refer to the set of these SSs as $\boldsymbol{L}_v$. Then, it selects SSs whose pair of source region s and destination region d satisfies $P \in A_{s,d}$ from $\boldsymbol{L}_v$. We refer this set of selected SSs as $\boldsymbol{L}'_v$. Finally, v randomly selects up to n SSs from $\boldsymbol{L}'_v$ and broadcasts them. This algorithm is shown in Algorithm 1.

Algorithm 1 RUF + LTZ strategy

When vehicle v attempts to send some sets of sketches
$\boldsymbol{L}_v, \boldsymbol{L}'_v \leftarrow \phi$
$\boldsymbol{L}_v \leftarrow$ top q recently updated sets of sketches on v
for all (s,d) of a set of sketches in $\boldsymbol{L}_v$ **do**
  if $P_v \in A_{s,d}$ **then**
    add $S_{v,s,d}$ to $\boldsymbol{L}'_v$
  end if
end for
select n items randomly from $\boldsymbol{L}'_v$

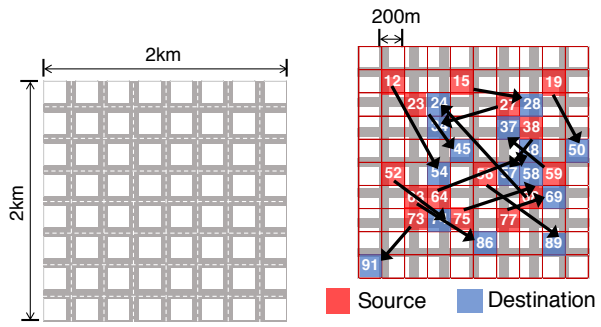
5 Evaluation of strategies for efficiently sending DMI

In this section, we evaluate the effectiveness of the proposed strategies through simulations. We focus on the accuracy of Dmaps and communication traffic required for sending DMI in this evaluation.

5.1 Simulation settings

5.1.1 Simulation scenario

In each scenario of our simulations, each vehicle broadcasts SSs to update and to share their Dmap using only VANET. We used Scenargie [24], a network simulator developed by Space-Time Engineering, and we implemented our proposed strategies and comparative strategies described later. In our scenarios, there are 14 double lane roads in a square area that has 2 km on each side as shown in Figure 11 (a). All intersections have traffic lights.



(a) Simulation area (b) Regions in Dmap

Figure 11. Simulation area and regions for a Dmap

All vehicles move on the roads in the simulation area by a mobility log generated using the SUMO road traffic simulator [25]. In addition, we conducted simulations of vehicle mobility by changing the inter-vehicle arrival rate. These vehicle mobility simulation logs are used for evaluating how the density of vehicles affects the accuracy of Dmaps and the effectiveness of the proposed strategies.

Table 1. Pairs of source region ID and destination region ID of demands generated in the simulations

Source	Destination	Source	Destination
12	54	23	45
19	50	15	28
52	74	63	86
56	89	27	34
38	57	73	91
75	58	77	69
59	37	64	48
68	24		

Table 2. Simulation parameters

Parameters	Values
Inter-vehicle arrival rate [/h·lane]	100, 300
Interval of sending beacons [sec]	1
Tx power [dBm]	10
Interval of generating demands [sec]	1
The effective period of a demand	63
Simulation time [s]	900
MAC Protocol	IEEE802.11p
Propagation model	ITU-R P.1411
App. max. data size [bytes]	500
Size per SS [bytes]	192
τ_u	30
k	10, 17
q	7
n	2
p	0.8

We divided the simulation area into small square regions 200×200m^2 and gave an ID to each small region for creating a Dmap of each vehicle Figure 11 (b). In Figure 11 (b), vehicles in blue regions demanded LDI of red regions corresponding to the blue regions. Table 1 also summarizes pairs of a source region and a destination region of demands that vehicles generated in the simulations. Vehicles in each source region shown in Table 1 demanded LDI of the destination region corresponding to the source region every second when they were moving towards the direction of the destination region. For instance, a vehicle in region 12 demanded LDI of region 54 every second only when it was moving eastbound or southbound.

We repeated simulation runs ten times with different random number seeds. Table 2 summarizes parameters used in the simulations. We determined the values of n, k and the interval of sending beacons based on preliminary simulations. We selected values of these parameters so that communication traffic could be as small as possible and the reproducibility of real demands in Dmaps (the ratio of coincidence described later) could be high based on the results of the preliminary simulations using the same scenarios.

5.1.2 Comparative strategies

To evaluate the effectiveness of proposed strategies, we used three comparative strategies, the *All* strategy, the *Random Timing* (RandT) strategy, and the *Random Selecting* (RandS) strategy. The All strategy and RandT strategy are for controlling the frequency of sending SSs and used for evaluating the effectiveness of the Frequent strategy. The RandS strategy is for selecting SSs to be sent, and used for evaluating the effectiveness of the LTZ strategy, RUF strategy, and RUF + LTZ Hybrid strategy.

In the All strategy, all vehicles do not control the frequency of sending SSs and send SSs with every beacon. In the RandT strategy, each vehicle decides to send SSs with a probability of p when it sends a beacon. In the RandS strategy, vehicles that send SSs randomly select n SSs to send from SSs that have been updated so far.

5.2 Simulation metrics

We evaluated the performance of the proposed strategies focusing on the accuracy of a Dmap in each vehicle and communication traffic required for sending DMI (= SSs). Here, the accuracy of a Dmap means the reproducibility of the actual number of demands in the Dmap. Since vehicles send LDI based on the estimated number of demands given by the Dmap, the reproducibility of the actual number of demands in the Dmap should be high. The reproducibility of the actual number of demands should be evaluated by using a relative metric because the influence of the error is different according to how the actual number of demands is. Therefore, for evaluating the accuracy of Dmaps, we use the ratio of coincidence between the actual number of demands and the estimated number of demands in Dmaps, M.

M for a pair of source region and destination region is defined as follows.

$$M = 1 - \min\{1, |R - D|/R\} \quad (0 \le M \le 1). \qquad (2)$$

where D denotes the estimated number of valid demands (TTL > 0) that are generated in a source region for a destination region, and R denotes the actual number of valid demands for the same combination of a source region and a destination region as D. In addition, we assume that R is not 0 in this formula. The closer M is to 1, the higher the accuracy of the Dmap. As mentioned above, only vehicles in an LTZ for the pair of source region s and destination region d should surely know the number of demands for LDI for the pair of s and d because we assume that vehicles send

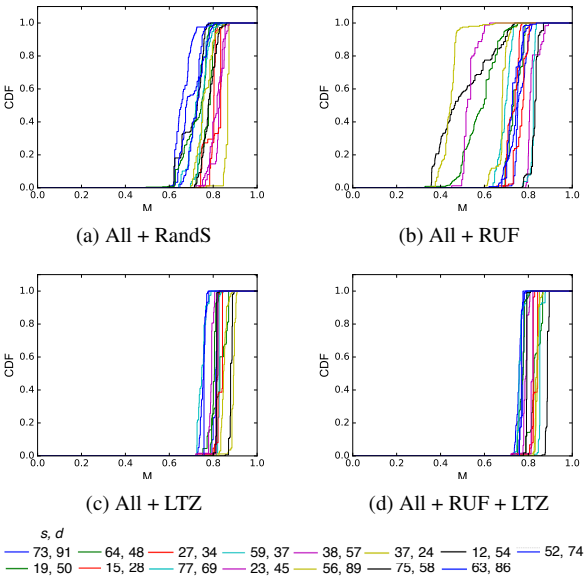
the LDI only when they are in LTZ corresponding to s and d. Therefore, we evaluated M for the pair of source region s and destination region d in a Dmap in only vehicles that are in the LTZ for the pair of s and d at the time point of the data analysis.

We also evaluated the communication traffic required for sending SSs. It is the average value of the product of the data size of an SS and the total number of broadcasted SSs in each simulation.

5.3 Simulation results

5.3.1 The ratio of coincidence M

Figure 12 shows the CDF of M calculated 800 seconds after the start of the simulation with vehicle traffic of 300 vehicles/lane·hour when the All strategy is used for controlling the frequency of sending SSs, and we evaluated the performance of the proposed strategies for selecting SSs. The vertical axis is the ratio of vehicles that have an SS representing M below each M on the horizontal axis. Each curve in the graph corresponds to a pair of a source region and a destination region of demands shown in Table 1.

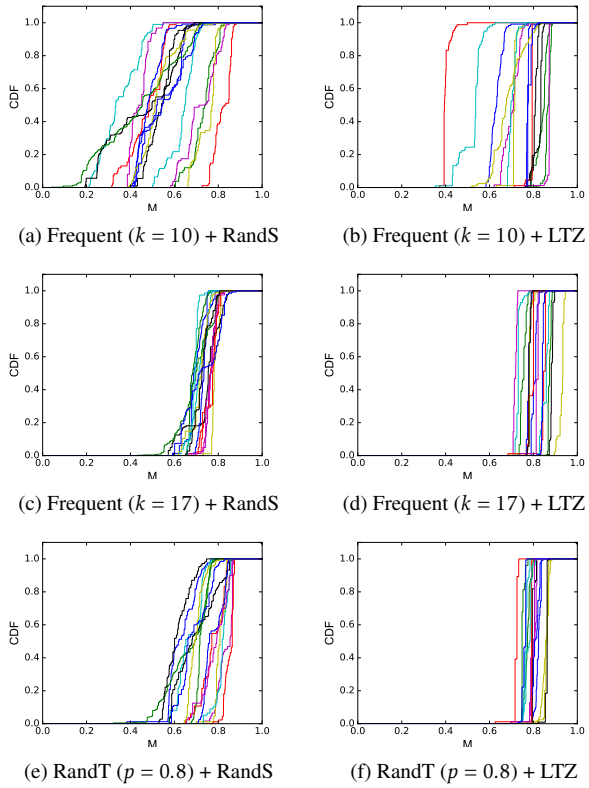


(a) All + RandS (b) All + RUF

(c) All + LTZ (d) All + RUF + LTZ

Figure 12. CDF of the Dmap accuracy (M) at 800 sec (inter-vehicle arrival rate: 300 vehicles/lane·hour)

As shown in Figure 12, the LTZ strategy and the RUF + LTZ strategy outperform the other strategies. When the RandS strategy is used, M of pairs of s and d (Hereafter, we call them sd-pairs) ranges from 0.6 to 0.86. With the LTZ strategy and the RUF + LTZ strategy, M of sd-pairs is improved compared with M when the RandS strategy is used. M of sd-pairs when the LTZ strategy or the RUF + LTZ strategy is used ranges from 0.7 to 0.86. This is because SSs for the pair of s and d are preferentially sent in LTZ corresponding to the SSs. The reason why M in the LTZ strategy is similar to M in the RUF + LTZ strategy is summarized as follows. SSs are frequently sent in their corresponding LTZ if the LTZ strategy is used, then the time stamp of the SSs on vehicles in the LTZ is frequently updated. Therefore, SSs are frequently sent in the LTZs when vehicles select recently updated SSs in RUF + LTZ strategy. Therefore, similar SSs are selected both in the LTZ strategy and the RUF + LTZ strategy. Consequently, M in the LTZ strategy is similar to M in the RUF + LTZ strategy. On the other hand, when only the RUF strategy is used for selecting SSs, M is deteriorated. The reason is that SSs of the same sd-pairs tend to be sent repeatedly and transmission of SSs of other sd-pairs is suppressed.



(a) Frequent (k = 10) + RandS (b) Frequent (k = 10) + LTZ

(c) Frequent (k = 17) + RandS (d) Frequent (k = 17) + LTZ

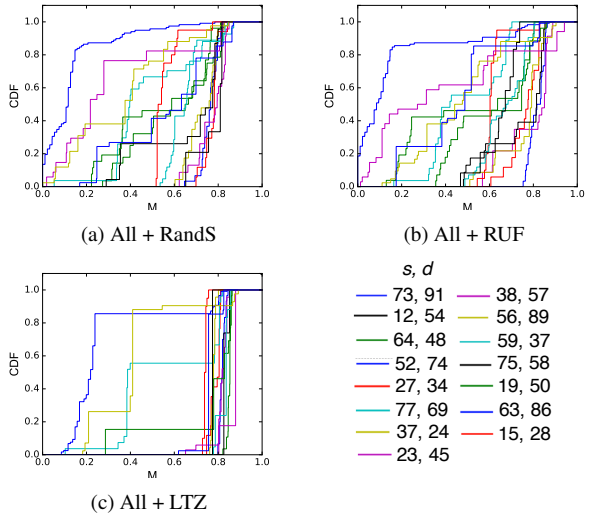
(e) RandT (p = 0.8) + RandS (f) RandT (p = 0.8) + LTZ

Figure 13. CDF of the Dmap accuracy (M) at 800 sec (inter-vehicle arrival rate: 300 vehicles/lane·hour)

Figure 13 shows the CDF of the Dmap accuracy (M) when the Frequent strategy and the RandT strategy were used for controlling the frequency of sending DMI (= SSs) with the RandS strategy or the LTZ strategy for selecting SSs. When vehicles use the RandT strategy and the RandS strategy, M is deteriorated compared to that in the case when the All strategy and the RandS strategy are used. This is simply because the number of exchanged SSs is decreased. However, when vehicles use the RandT strategy and the LTZ strategy, M is similar to one in the case when the All strategy and the LTZ strategy are used. The reason is that SSs are frequently exchanged in LTZ corresponding to them when the LTZ strategy is used even if the small number of vehicles send SSs. When the Frequent strategy with $k = 10$ is used, M is more deteriorated than that in the case when the RandT strategy is used. The reason is that the number of vehicles that send SSs is excessively restrained in the Frequent strategy with $k = 10$. On the other hand, when the Frequent strategy with $k = 17$ is used, M approximated M in the cases when the All strategy is used because the number of vehicles that send SSs was hardly not restrained.

Figure 14 shows CDF of M calculated at 800 seconds from the start of the simulation with 100 vehicles/lane·hour

vehicle traffic. Compared with Figure 12 (a)–(c), we can see that M is deteriorated when the vehicle density is low. This is because a sufficient number of SSs is not exchanged due to lack of connectivity between vehicles. Interestingly, even with poor connectivity, the LTZ strategy outperforms other strategies and most of the sd-pairs achieve high M around 0.8. The reason is that SSs are easily updated in LTZ corresponding to the SSs because SSs are frequently sent in the LTZ.



(a) All + RandS (b) All + RUF

(c) All + LTZ

Figure 14. CDF of M at 800 sec (inter-vehicle arrival rate: 100 vehicles/lane·hour)

5.3.2 Communication traffic

Figure 15 shows communication traffic in each combination of strategies when vehicle traffic is 300 vehicles/lane·hour. We can see that the LTZ strategy most effectively reduces communication traffic among strategies for selecting SSs. The reason is that each SS is sent only in the LTZ corresponding to the SS and is not sent outside the LTZ.

Concerning strategies for controlling the frequency of sending SSs, the Frequent strategy ($k = 10$) most effectively reduces the communication traffic. However, when the Frequent strategy (k=10) is used, M becomes low as shown in Figure 13 (a), (b) because SSs are rarely updated due to the limited chance to send SSs.

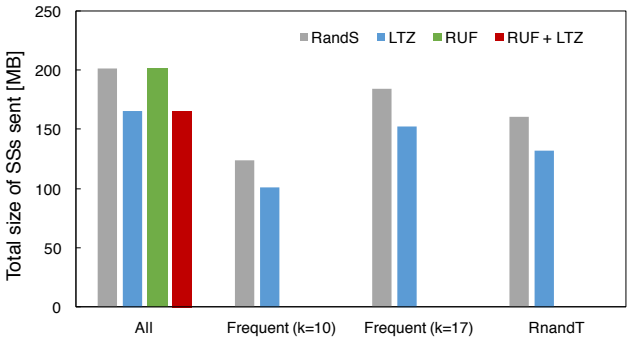


Figure 15. Communication traffic (Inter-vehicle arrival rate: 300 vehicles/lane·hour)

Summarizing the simulation results of the ratio of coincidence M and communication traffic, we can conclude that the LTZ strategy is the most useful to improve both the Dmap accuracy and communication traffic needed for exchanging SSs. The reason is that this strategy increases the frequency of SS exchange in their corresponding LTZ and reduces the frequency of exchanging them outside the LTZ. Concerning strategies for controlling the frequency of sending SSs, the Frequent strategy reduces the communication traffic. However, it deteriorates the Dmap accuracy. Therefore, the Frequent strategy does not solve the tradeoff between the Dmap accuracy and communication traffic. In addition, when both the LTZ strategy and RandT strategy are used, communication traffic and the Dmap accuracy is most effectively improved.

6 Conclusions

In this paper, we proposed a strategy for controlling the frequency of sending DMI (SSs) and strategies for selecting SSs that vehicles should send for improving the accuracy of a soft-state sketch-based Dmap on vehicles used for Dmap-based location dependent data dissemination services on VANETs, such as a real-time visual car navigation system. The simulation results revealed that one of the proposed strategies for selecting SSs, the LTZ strategy, effectively reduces communication traffic and achieves high Dmap accuracy, i.e. the reproducibility of the actual number of demands in Dmaps. The key feature of the LTZ strategy is that it is designed in consideration of where the exchanged information, i.e. SSs, may be used. In this strategy, vehicles send SSs in only locations where SSs may be used and do not send SSs in locations where SSs may not be used. Thus, it can reduce communication traffic and improve the accuracy of SSs in locations where SSs is used. On the other hand, the strategy for controlling the frequency of sending SSs, the Frequent strategy, which gives a high priority to vehicles frequently updating SSs, deteriorates the Dmap accuracy, while it drastically reduces communication traffic required for exchanging SSs.

The demand generation patterns used in the evaluation in the paper is limited. Thus, the performance evaluation of various demand generation patterns and realistic demand generation patterns are needed because there are many demand generation patterns in real life. For example, a demand generation pattern that frequency of demands change as time progress or change depending on a distance between a source region to a destination region should be considered. The mobility pattern of vehicles will also affect the performance of sharing Dmap information. For example, if an isolated vehicle has important Dmap information of a distant area, giving a high priority to exchanging the important Dmap information will contribute to that many vehicles knowing the distributions of demands for LDI. If many vehicles in the distant area want to obtain LDI of an area that has low connectivity from them in terms of store-carry-forward-based data delivery, such demand information will be carried by isolated vehicles.

For delivering LDI based on Dmaps, additional information is needed for each vehicle to determine whether it forwards an LDI item or not. For example, if its neighboring vehicle has sent the same LDI item recently, it should

not forward the LDI item to reduce the redundant data transmission without loss of the reliability of data delivery. Besides, if there are multiple isolated routes between the PoI and the destination of the LDI item, whether selecting one of the routes or using all the routes will affect the reliability of the LDI data delivery. In addition, we should design a criterion for selecting LDI that vehicles send, a method for calculating efficient forwarding routes based on Dmaps, etc. Our proposed strategies are important factors in the routing of LDI because vehicles disseminate LDI based on Dmaps in Dmap-based data dissemination. The design of a data forwarding algorithm based on Dmaps and additional information remains as our future work.

Acknowledgment This work was supported by JSPS KAKENHI Grant Number 15H02689.

References

[1] MLIT Ministry of Land, Infrastructure, Transport and Tourism: ITS, http://www.mlit.go.jp/road/road_e/p1_its.html#a2 (Accessed 14th September 2017)

[2] Pioneer news release: Pioneer introduces new CYBER NAVI car navigation systems for Japan Market, http://pioneer.jp/en/news/press/index/1617/May2013.

[3] J. Okamoto and S. Ishihara, "Distributing location-dependent data in VANETs by guiding data traffic to high vehicle density areas," in Vehicular Networking Conference, 2010. https://doi.org/10.1109/VNC.2010.5698248

[4] S. Ishihara, N. Nakamura, and Y. Niimi, "Demand-based Location Dependent Data Dissemination in VANETs" in the 19th annual international conference on Mobile computing & networking, 2013. https://doi.org/10.1145/2500423.2504580

[5] N. Nakamura, Y. Niimi, and S. Ishihara, "Live VANET CDN: Adaptive data dissemination scheme for location-dependent data in VANETs" in Vehicular Networking Conference, 2013. https://doi.org/10.1109/VNC.2013.6737595

[6] L. Yin and G. Cao, "Supporting cooperative caching in ad hoc networks" IEEE Trans. Mobile Computing, **5**(1), 77–89, 2006. https://doi.org/10.1109/TMC.2006.15

[7] A. Yamada and S. Ishihara. "Data Exchange Strategies for Aggregating Geographical Distribution of Demands for Location-Dependent Information Using Soft-State Sketches in VANETs" in Advanced Information Networking and Applications, 2017 IEEE 31st International Conference on, 2017. https://doi.org/10.1109/AINA.2017.42

[8] G. Grassi, D. Pesavento, and G. Pau, "Navigo: Interest forwarding by geolocations in vehicular Named Data Networking" in World of Wireless, Mobile and Multimedia Networks, 2015 IEEE 16th International Symposium on a, 2015. https://doi.org/10.1109/WoWMoM.2015.7158165

[9] S. H. Ahmed, S. H. Bouk, and M. A. Yaqub, "CODIE: Controlled Data and Interest Evaluation in Vehicular Named Data Networks" IEEE Trans. Vehicular Technology, **65**(6), 3954–3963, 2016. https://doi.org/10.1109/TVT.2016.2558650

[10] C. Bian , T. Zhao, X. Li , W. Yan, "Boosting named data networking for data dissemination in urban VANET scenarios" Vehicular Communications, **2**(4),195–207, 2015. https://doi.org/10.1016/j.vehcom.2015.08.001

[11] C. Sommer, S. Joerer, M. Segata, "How shadowing hurts vehicular communications and how dynamic beaconing can help" IEEE Trans. Mobile Computing, **14**(7), 1411–1421, 2015. https://doi.org/10.1109/TMC.2014.2362752

[12] W. Viriyasitavat, O.K. Tonguz, and F. Bai, "UV-CAST: an urban vehicular broadcast protocol" in Vehicular Networking Conference, 2010. https://doi.org/10.1109/VNC.2010.5698266

[13] O.K. Tonguz, N. Wisitpongphan, and F. Bai, "DV-CAST: A distributed vehicular broadcast protocol for vehicular ad hoc networks" IEEE Wireless Communications, **17**(2), 47–57, 2010. https://doi.org/10.1109/MWC.2010.5450660

[14] P. Th. Eugster, P. A. Felber, R. Guerraoui, and A. M. Kermarrec. "The many faces of publish/subscribe" ACM computing surveys, **35**(2), 114–131, 2003. https://doi.org/10.1145/857076.857078

[15] I. Leontiadis and C. Mascolo, "Opportunistic Spatio-Temporal Dissemination System for Vehicular Networks" in the 1st international MobiSys workshop on Mobile opportunistic networking, 2007. https://doi.org/10.1145/1247694.1247702

[16] T. Pandey, D. Garg, and M.M. Gore, "Publish/subscribe based information dissemination over VANET utilizing DHT" Frontiers of Computer Science, **6**(6), 713–724, 2012. https://doi.org/10.1007/s11704-012-1154-7

[17] Y. Niimi and S. Ishihara, "Demand map-based data dissemination scheme for location dependent data in VANETs" in Mobile Computing and Ubiquitous Networking, 2015 Eighth International Conference on, 2015. https://doi.org/10.1109/ICMU.2015.7061047

[18] L. Wischhof, A. Ebner, H. Rohling, "Information dissemination in self-organizing inter-vehicle networks" IEEE Trans. Intelligent Transportation Systems, **6**(1), 90–101, 2005. https://doi.org/10.1109/TITS.2004.842407

[19] T. Nadeem, S. Dashtinezhad, C. Liao, L. Iftode, "TrafficView: traffic data dissemination using car-to-car communication" ACM SIGMOBILE Mobile Computing and Communications Review, **8**(3), 6–9, 2004. https://doi.org/10.1145/1031483.1031487

[20] K. Ibrahim, M.C. Weigle, "Accurate data aggregation for VANETs" in the fourth ACM international workshop on Vehicular ad hoc networks, 2007. https://doi.org/10.1145/1287748.1287761

[21] J. Bronsted, L.M. Kristensen, "Specification and performance evaluation of two zone dissemination protocols for vehicular ad- hoc networks" in Specification and performance evaluation of two zone dissemination protocols for vehicular ad-hoc networks, 2006. https://doi.org/10.1109/ANSS.2006.43

[22] C. Lochert, B. Scheuermann, and M. Mauve, "A probabilistic method for cooperative hierarchical aggregation of data in VANETs" Ad Hoc Networks, **8**(5), 518–530, 2010. https://doi.org/10.1016/j.adhoc.2009.12.008

[23] D. Sormani, G. Turconi, P. Costa, D. Frey, M. Migliavacca, and L. Mottola, "Towards Lightweight Information Dissemination in Inter-Vehicular Networks" in the 3rd international workshop on Vehicular ad hoc networks, 2006. https://doi.org/10.1145/1161064.1161069

[24] Space-Time Engineering, https://www.spacetime-eng.com/ (Accessed 28th September 2017)

[25] SUMO — Simulation of Urban Mobility, http://sumo.dlr.de/wiki/Main_Page (Accessed 28th September 2017)

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 186-193 (2017)
www.astesj.com
Special issue on Advancement in Engineering Technology

ASTES Journal
ISSN: 2415-6698

Fault Tolerant Control Based on PID-type Fuzzy Logic Controller for Switched Discrete-time Systems: An Electronic Throttle Valve Application

Wafa Gritli*, Hajer Gharsallaoui, Mohamed Benrejeb

Université de Tunis El Manar, LA.R.A Automatique, Ecole Nationale d'Ingénieurs de Tunis, BP 37, Le Belvédère, 1002 Tunis, Tunisie.

A R T I C L E I N F O

Article history:
Received: 09 November, 2017
Accepted: 01 December, 2017
Online: 14 December, 2017

Keywords:
Electronic throttle valve
Switched systems
Fault tolerant control
PID-type FLC
Unknown input observer
Lyapunov theory
LMIs

A B S T R A C T

This paper deals with the problem of Fault Tolerant Control (FTC) using PID-type Fuzzy Logic Controller (FLC) for an Electronic Throttle Valve (ETV) described by a switched discrete-time systems with input disturbances and actuator faults. In order to detect the faults, Unknown Input Observers (UIOs) are designed and formulated in terms of Lyapunov theory and Linear Matrix Inequalities (LMIs). This approach is designed in order to minimize the error between the desired flat trajectory generated using the flatness property and the estimated state provided from differents UIOs and to maintain asymptotic stability under an arbitrary switching signal, even in the presence of actuator faults. The simulation results have shown the effectiveness of the proposed approach.

1 Introduction

Actuator faults may cause undesired system behaviour and sometimes lead to instability, hence it is necessary to develop Fault Tolerant Control (FTC) methods against actuator faults of uncertain nonlinear systems. In the past decades some FTC design methods have been proposed for several classes of nonlinear systems with actuator faults [1-3]. It consists in computing control laws by taking into account the faults affecting the system in order to maintain acceptable performances and to preserve stability of the system in the faulty situations [4]. From the point of view of FTC strategies, the literature considers two main groups of techniques: passive and active ones. In passive FTC, the faults are treated as uncertainties. Therefore, the control is designed to be robust only to the specified faults [4]. Active FTC techniques consist in adapting the control law using the information given by the Fault Detection and Isolation (FDI) block [5]. The informations issued from the FDI block are used by the FTC module to reconfigure the control law in order to compensate the fault and to ensure an acceptable system performances.

The study of this problem was extended to switched systems in [6-8]. In [7], a switched discrete-time system with state delay has been considered. The design method is based on the construction of a filter and a fault estimation approach. In [8], an adaptive fuzzy tracking control method for a class of switched nonlinear systems with arbitrary switchings and with actuator faults has been proposed. The proposed control scheme guarantee the stability of the whole switched control system based on the common Lyapunov function stability theory and attenuate the effect of the actuator faults on the control performance by designing a new fuzzy controller to accommodate uncertain actuator faults. In [9], an observer has been built to detect the fault when it occurs. The problem of FTC for switchied linear systems is addressed by using a nominal control law designed in the absence of any fault, associated with fault detection, localization and reconfiguration techniques to maintain the stability of the system under an arbitrary switching signal in the presence of sensor faults. A state trajectory tracking has been proposed in [4] for actuator faults and observer bank based on controllers with switching mechanism for sensor faults has been also presented in [10]. A nonlinear observer based controller, adopting the so-

*Corresponding Author: Wafa Gritli, Email: wafa.gritli@enit.rnu.tn

https://dx.doi.org/10.25046/aj020623

called parallel distributed compensation structure, has been designed to choose an adequate state estimate to compensate the effects of the faults on the system in [11]. Switched systems are dynamical systems consisting of a collection of continuous-time subsystems. Switched systems have attracted more attention due to their significance in the modelling of many engineering applications, such as chemical processes, robot manipulators and power systems. Stability analysis and synthesis of switched systems have been made using Lyapunov function to ensure stability of the switched systems such in [12].

In this paper, in order to acheive the FTC for a switched discrete-time systems, a set of PID-type Fuzzy Logic Controller (FLC) is implemented to minimize the error between the desired flat trajectory and the estimated state and to maintain the stability of the system in the presence of actuator faults. The estimated state is provided from an Unknown Input Observer (UIO). Based on residual analysis, a switching strategy using stateflow is then designed. The global stability for the switched systems is studied by Lyapunov theory and expressed as a Linear Matrix Inequalities (LMIs).

The paper is organized as follows. In Section 2, the FTC problem statement is formulated. In Section 3, the fault detection method using UIO is introduced. In Section 4, the PID-type FLC is described. Section 5 deals with the flatness property. The proposed approach is applied to an ETV in Section 6. Finally, the conclusion is drawn in Section 7.

2 FTC problem statement

Consider the discrete time switched system, which can be formulated such that

$$x(k+1) = A_j x(k) + B_j u(k) + E_j d(k) + B_j f_a(k) \quad (1)$$

$$y(k) = C_j x(k)$$

$x(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}^p$ the control input, $y(k) \in \mathbb{R}^o$ the measured output, $d(k) \in \mathbb{R}^p$ the unknown disturbance input and $f_a(k) \in \mathbb{R}^p$ the actuator faults. $A_j \in \mathbb{R}^{n\times n}$, $B_j \in \mathbb{R}^{n\times p}$, $C_j \in \mathbb{R}^{o\times n}$ and $E_j \in \mathbb{R}^{n\times p}$ are the known constant matrices for $j \in \psi = \{1,2,...,m\}$ and m the number of models, $m > 1$.

In this paper, a FTC structure based on PID-type FLC, given by Figure 1, is proposed to maintain the trajectory tracking and to preserve stability of ETV in the presence of both input disturbance $d(k)$ and actuator faults $f_a(k)$.

According to this structure, the FTC approach needs to detect actuator faults firstly and then to design the j^{th} control law, denoted $u_j(k)$ given by (2), in order to minimize the error between the desired flat trajectory, generated using the flatness property and the estimated state, provided from the j^{th} UIO.

$$u_j(k) = -K_{c,j}\hat{x}_j(k) + y_j^d(k) \quad (2)$$

$K_{c,j}$ is the j^{th} gain control law.

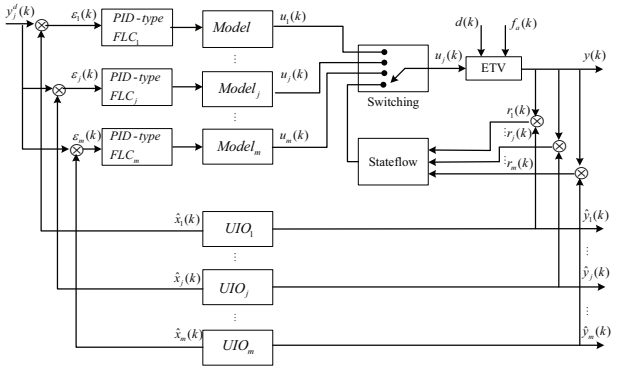


Figure 1: Proposed structure of FTC approach

In this case, LMI-based UIOs for the switched system (1) are designed using Lyapunov function for arbitary switching signal. Then, residuals $r_j(k)$ are determined by the UIOs and employed to achieve actuator faults detection.

3 Unknown input observers design and stability analysis

In this section, let us consider $f_a(k) = 0$, then the model (1) becomes

$$x(k+1) = A_j x(k) + B_j u(k) + E_j d(k) \quad (3)$$

$$y(k) = C_j x(k)$$

$j \in \psi$ denotes the j^{th} model. The structure of full rank UIOs can be formulated by

$$z(k+1) = F_j(k) + T_j B_j u(k) + K_j y(k) \quad (4)$$

$$\hat{x}(k) = z(k) + H_j y(k)$$

where $\hat{x}(k)$ is the estimated state vector $x(k)$, $z(k)$ is the state vector of full rank UIOs. F_j, T_j, K_j and H_j are unknown matrices which need to be designed.

Lemma: [13] Equation (4) is UIO of the switched system (3), if and only if the following conditions are satisfied

- $rank(C_j E_j) = rank(E_j)$
- $(C_j A_{j1})$ is observable

with

$$A_{j1} = A_j - E_j\left[(C_j E_j)^T C_j E_j\right]^{-1}(C_j E_j)^T C_j A_j \quad (5)$$

According to the above, the formulation of UIOs is constructed for each subsystem. In the next part, the multiple Lyapunov function will be used to realize the design of parameters for UIOs of switched system.

Theorem: [14] In the condition of arbitrary switching signal, for the system (3), if

$$(H_j C_j - I)E_j = 0 \quad (6)$$

$$T_j = (I - H_j C_j)$$

$$F_j = (A_j - H_j C_j A_j - K_{j1} C_j)$$
$$K_{j2} = F_j H_j$$

hold, and there exist symmetric matrices $P_j > 0$, $\forall j \in \psi$ such that

$$\begin{bmatrix} -P_j & P_j F_j \\ F_j^T P_j & -P_j \end{bmatrix} \leqslant 0, \ \forall j \in \psi \tag{7}$$

then the parameters of UIOs can be designed.

Proof: Define the state estimation error as $e(k) = x(k) - \hat{x}(k)$, the dynamic equation can be derived as

$$\begin{aligned} e(k+1) &= (A_j - H_j C_j A_j - K_{j1} C_j) e(k) \\ &- \left[F_j - (A_j - H_j C_j A_j - K_{j1} C_j)\right] z(k) \\ &- \left[K_{j2} - (A_j - H_j C_j A_j - K_{j1} C_j) H_j\right] y(k) \\ &- \left[T_j - (I - H_j C_j)\right] B_j u(k) - (H_j C_j - I) E_j d(k) \end{aligned} \tag{8}$$

with

$$K_j = K_{j1} + K_{j2} \tag{9}$$

In order to make the error decoupled from known control input $u(k)$, measured output $y(k)$ and unknown input $d(k)$, we should let

$$(H_j C_j - I) E_j = 0 \tag{10}$$

$$T_j - (I - H_j C_j) = 0$$
$$F_j - (A_j - H_j C_j A_j - K_{j1} C_j) = 0$$
$$K_{j2} - (A_j - H_j C_j A_j - K_{j1} C_j) H_j = 0$$

It can be concluded that

$$H_j = E_j \left[(C_j E_j)^T C_j E_j \right]^{-1} (C_j E_j)^T \tag{11}$$

$$A_{j1} = A_j - E_j \left[(C_j E_j)^T C_j E_j \right]^{-1} (C_j E_j)^T C_j A_j \tag{12}$$

$$F_j = A_j - H_j C_j A_j - K_{j1} C_j = A_{j1} - K_{j1} C_j \tag{13}$$

and the error dynamics is given by

$$e(k+1) = F_j e(k) \tag{14}$$

Equation (4) is UIOs of the switched system (3) if the estimation error tends asymptotically to zero despite the presence of an unknown input $d(k) \neq 0$.

Consider the following Lyapunov function

$$V_j(e(k)) = e(k)^T P_j e(k) \tag{15}$$

For $P_j = P_j^T > 0$, then, $V_j(e(k)) > 0$ holds, the ΔV given by (16) is negative

$$\begin{aligned} &V_i(e(k+1)) - V_j(e(k)) \\ &= e(k)^T F_j^T P_i F_j e(k) - e(k)^T P_j e(k) \\ &= e(k)^T (F_j^T P_i F_j - P_j) e(k); (i \in \psi, j \in \psi, i \neq j) \end{aligned} \tag{16}$$

if the following inequalities are satisfied

$$F_j^T P_i F_j - P_j \leqslant 0 \tag{17}$$

Thus, the error system (14) is stable asymptotically.

According to Schur complement lemma, the inequalities (17) can be rewritten as

$$\begin{bmatrix} -P_i & P_i F_j \\ F_j^T P_i & -P_j \end{bmatrix} \leqslant 0 \tag{18}$$

By substituting $F_j = A_{j1} - K_{j1} C_j$, the above matrix inequalities become

$$\begin{bmatrix} -P_i & P_i (A_{j1} - K_{j1} C_j) \\ (A_{j1} - K_{j1} C_j)^T P_i & -P_j \end{bmatrix} < 0 \tag{19}$$

and

$$\begin{bmatrix} -P_i & (P_i A_{j1} - W_{ij} C_j) \\ (P_i A_{j1} - W_{ij} C_j)^T & -P_j \end{bmatrix} < 0 \tag{20}$$

for $W_{ij} = P_i K_{j1}$.

Since $K_{j1} = P_j^{-1} W_{ij}$, we can obtain the value of K_{j1} then F_j, K_{j2} and $K_j = K_{j1} + K_{j2}$ from P_j and W_{ij} solutions of LMIs.

The aim of fault detection is to generate a residual signal $r_j(k)$, given by (21), which is sensitive to $f_a(k)$ in the presence of actuator faults which is the purpose of [15].

$$r_j(k) = y(k) - C_j \hat{x}_j(k) \tag{21}$$

4 PID-type fuzzy logic controller design

In this study, the FTC approach is based on PID-type FLC. To adjust the input and the output scaling factors of this controller, Genetic Algorithm (GA) optimization technique has been proposed in order to improve the performance of the controller.

4.1 PID-type fuzzy logic controller description

We consider a PID-type FLC structure as shown in Figure 2, [16], where $K_e \in \mathbb{R}^+$ and $K_d \in \mathbb{R}^+$ are the input scaling factors, $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$ the output scaling factors.

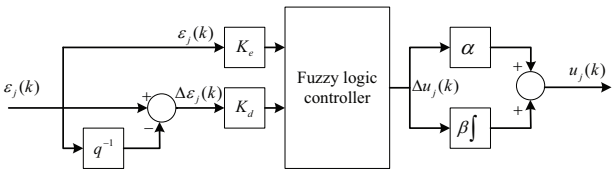


Figure 2: The PID-type FLC

The FLC inputs variables, the error $\varepsilon_j(k)$ between the desired flat trajectory $y_j^d(k)$ and the estimated state $\hat{x}(k)$, and error variation $\Delta\varepsilon_j(k)$ are given by the equations (22) and (23) where T_e is the sampling period.

$$\varepsilon_j(k) = -K_{c,j} \hat{x}(k) + y_j^d(k) \tag{22}$$

$$\Delta\varepsilon_j(k) = \frac{\varepsilon_j(k) - \varepsilon_j(k-1)}{T_e} \tag{23}$$

The output variable $\Delta u_j(k)$ of a such controller is the variation of the control law signal $u_j(k)$ which can be defined as (24).

$$\Delta u_j(k) = \frac{u_j(k) - u_j(k-1)}{T_e} \tag{24}$$

The output of the PID-type FLC is given by (25), [17]

$$\begin{aligned} u_j(k) &= \alpha\Delta u_j(k) + \beta \int \Delta u_j(k)\ \mathrm{d}t \\ &= \alpha(A + QK_e u_j(k) + DK_d\Delta u_j(k)) \\ &+ \beta \int (A + QK_e u_j(k) + DK_d\Delta u_j(k))\ \mathrm{d}k \\ &= \alpha A + \beta Ak + (\alpha K_e Q + \beta K_d D)u_j(k) \\ &+ \beta K_e Q \int u_j(k)\ \mathrm{d}k + \alpha K_d D\Delta u_j(k) \end{aligned} \tag{25}$$

where the terms Q and D are given by (26) and (27), [10].

$$Q = \frac{\Delta u_{(i+1)j} - \Delta u_{ij}}{\varepsilon_{i+1} - \varepsilon_i} \tag{26}$$

$$D = \frac{\Delta u_{i(j+1)} - \Delta u_{ij}}{\Delta\varepsilon_{j+1} - \Delta\varepsilon_j} \tag{27}$$

The fuzzy controllers with product−sum inference method, centroid defuzzification method and triangular uniformly distributed membership functions for the inputs and a crisp output proposed in [18], are used, in our study.

The linguistic levels, assigned to the variables $\varepsilon_j(k)$, $\Delta\varepsilon_j(k)$ and $\Delta u_j(k)$, are given in Table 1 as follows: NL: Negative Large; N: Negative; ZR: Zero; P: Positive; PL: Positive Large.

$\varepsilon_j(k)$ / $\Delta\varepsilon_j(k)$	N	ZR	P
N	NL	N	ZR
ZR	N	ZR	P
P	ZR	P	PL

Table 1: Fuzzy rules-base

4.2 Optimization of scaling factors using genetic algorithm

To adjust the input and the output scaling factors (K_e, K_d) and (α, β), GA is used in order to obtain their optimal values.

At first, an initial chromosome population is randomly generated. The chromosomes are candidate solutions to the problem. Then, the fitness values of all chromosomes are evaluated by calculating an objective function. So, based on the fitness of each individual, a group of the best chromosomes is selected through the selection process. The genetic operators, crossover and mutation, are applied to this 'surviving' population in order to improve the next generation solutions. Crossover is a recombination operator that combines subparts of two parent chromosomes to produce offspring. This operator extracts common features from different chromosomes in order to achieve even better solutions. Mutation is an operator that introduces variations into the chromosome. The modifications can consist of changing one or more values of a chromosome. Through the mutation operator the search space is explored by looking for better points. The process continues until the population converges to the stop criterion.

The most crucial step in applying GA is to choose the objective function that is used to evaluate the fitness of each chromosome. In this paper, the method of tuning PID-type FLC parameters using GA is based on minimizing the Integral of the Squared Error (ISE) used in [18].

5 Flatness and trajectory planning

The flatness approach is used in a discrete-time framework for system (1). Let the studied dynamic linear discrete system described by (28)

$$D_j(q)y_j(k) = N_j(q)v_j(k) \tag{28}$$

q is the forward operator, $v_j(k)$ and $y_j(k)$ are, respectively, the input and the output signals and $D_j(q)$ and $N_j(q)$ the polynomials defined by

$$D_j(q) = q^n + a_{j,n-1}q^{n-1} + \ldots + a_{j,1}q + a_{j,0} \tag{29}$$

$$N_j(q) = b_{j,n-1}q^{n-1} + \ldots + b_{j,1}q + b_{j,0} \tag{30}$$

where the parameters $a_{j,l}$ and $b_{j,l}$ are constants, $l = \{0, 1, ..., n-1\}$.

The system is considered as linear and controllable, consequently it is flat, [16].

The flat output $z_j(k)$, on which depend the output $y_j(k)$ and the control $v_j(k)$, can be seen as being the partial state of a linear system, [19]

$$v_j(k) = D_j(q)z_j(k) \tag{31}$$

$$y_j(k) = N_j(q)z_j(k) \tag{32}$$

The open loop control law can be determined by the following relations, [18]

$$v_j^d(k) = f(z_j^d(k), ..., z_j^{d^{(r+1)}}(k)) \tag{33}$$

$$y_j^d(k) = g(z_j^d(k), ..., z_j^{d^{(r)}}(k)) \tag{34}$$

f and g are vectorial functions and $z_j^d(k)$ is the desired trajectory that must be differentiable at the $(r+1)$ order.

In order to plan the desired flat trajectory $z_j^d(k)$, the polynomial interpolation technique is used.

Let consider the state vector: $Z_j^d(k) = (z_j^d(k)\ z_j{}^{d^{(1)}}(k)\ \ldots\ z_j^{d^{(r+1)}}(k))^T$ containing the desired flat output and its successive derivatives, [18].

The expression of $Z_j^d(k)$ can be given as following where k_0 and k_f are two moments known in advance.

$$Z_j^d(k) = M_{j,1}(k-k_0)c_{j,1}(k_0) + M_{j,2}(k-k_0)c_{j,2}(k_0, k_f) \tag{35}$$

such that

$$M_{j,1} = \begin{pmatrix} 1 & k & \cdots & \frac{k^{n-1}}{(n-1)!} \\ 0 & 1 & \cdots & \frac{k^{(n-2)}}{(n-2)!} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \tag{36}$$

$$M_{j,2} = \begin{pmatrix} \frac{k^n}{n!} & \frac{k^{n+1}}{(n+1)!} & \cdots & \frac{k^{2n-1}}{(2n-1)!} \\ \frac{k^{n-1}}{(n-1)!} & \frac{k^n}{n!} & \cdots & \frac{k^{(n-2)}}{(n-2)!} \\ \vdots & \ddots & \ddots & \vdots \\ k & \cdots & \frac{k^{n-1}}{(n-1)!} & \frac{k^n}{n!} \end{pmatrix} \quad (37)$$

$$c_{j,1} = Z_j^d(k_0) \quad (38)$$

$$c_{j,2} = M_{j,2}^{-1}(k_f - k_0)(Z_{j^d}(k_f) - M_{j,1}(k_f - k_0)Z_j^d(k_0)) \quad (39)$$

After planning a desired flat trajectory in discrete-time framework, the real output $y_j(k)$, to be controlled, follows the desired trajectory $y_j^d(k)$ such that (40), [18].

$$y_j^d(k) = N_j(q)z_j^d(k) \quad (40)$$

6 Application of an electronic throttle valve

6.1 Electronic throttle valve modeling

The proposed approach is applied here for the case of the electronic throttle valve, Figure 3, [15].

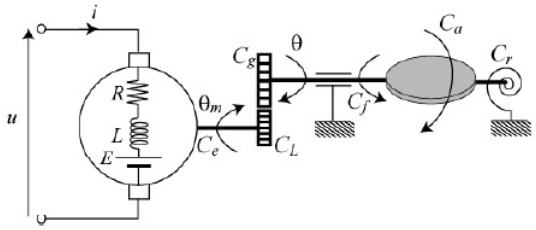


Figure 3: Electronic throttle system

The electrical part of this system is modeled by (41), [20,21].

$$u(t) = Ri(t) + L\frac{d}{dt}i(t) + k_v\omega_m(t) \quad (41)$$

L is the inductance, R the resistance, $u(t)$ and $i(t)$ the voltage and the armature current respectively, k_v a constant electromotive force and $\omega_m(t)$ the motor rotational speed.
The mechanical part of the throttle is modeled by a gear reducer, characterized by its reduction ratio γ such as (42)

$$\gamma = \frac{C_g}{C_L} \quad (42)$$

C_L is the load torque and C_g the gear torque.
The mechanical part is modeled according to (43) and (44), such that, [20,21]

$$J\frac{d}{dt}\omega_m(t) = C_e - C_f - C_r - C_a \quad (43)$$

$$\frac{d}{dt}\theta(t) = (180/\pi/\gamma)\omega_m(t) \quad (44)$$

$\theta(t)$ is the throttle plate angle, C_e the electrical torque, C_f the torque caused by mechanical friction, C_r the spring torque, C_a the torque generated by the airflow and J the overall moment of inertia.
The electrical torque is defined by

$$C_e = k_e i(t) \quad (45)$$

where k_e is a constant.
The electronic throttle valve involves two complex nonlinearities due to C_r and C_f, given by their static characteristics, [15]

- a dead zone in which the control voltage signal has no effect on the nominal position of the valve plate,
- two hysteresis combined with a saturation, due to the valve plate movement, limited by the maximum and the minimum angles.

The static characteristic of the nonlinear spring torque C_r is defined by

$$C_r = \frac{k_r}{\gamma}(\theta - \theta_0) + D\,\mathrm{sgn}(\theta - \theta_0) \quad (46)$$

for $\theta_{\min} \leqslant \theta \leqslant \theta_{\max}$, k_r is the spring constant, D a constant, θ_0 the default position and sgn (.) the following signum function

$$\mathrm{sgn}\,(\theta - \theta_0) = \begin{cases} 1, \text{ if } \theta \geqslant \theta_0 \\ -1, \text{ else} \end{cases} \quad (47)$$

The friction torque function C_f of the angular velocity of the throttle plate can be expressed as

$$C_f = f_v\omega + f_c\,\mathrm{sgn}(\omega) \quad (48)$$

where f_v and f_c are two constants.
By substituting in equation (43), the expressions C_e, C_f and C_r and by neglecting the torque generated by the airflow C_a, the two nonlinearities $\mathrm{sgn}(\theta - \theta_0)$ and $\mathrm{sgn}(\omega)$ and the two constants $\frac{k_r}{\gamma}\theta_0$ and f_v, the studied system is linear then it can be modelized by the following transfer function $H(s)$ (49), [21]

$$H(s) = \frac{(180/\pi/\gamma)k_e}{JLs^3 + JRs^2 + (k_ek_v + Lk_s)s + Rk_s} \quad (49)$$

with $k_s = (180/\pi/\gamma^2)k_r$ and s the Laplace operator; the identified parameters of this system are given in Table 2 at $25°C$, [20].

Parameters	Values	Units
R	2.8	Ω
L	0.0011	H
k_e	0.0183	$N.m/A$
k_v	0.0183	$v/rad/s$
J	4×10^{-6}	$kg.m^2$
γ	16.95	-

Table 2: Model's parameters

Recent work has shown that the ETV can be modeled by two linear models identified from the default position of the throttle plate for two values of the parameter k_s and for the sampling period $T_e = 0.002s$, [21]

- a model representing the position of the plate above the position by default for: $k_s = 1.877 \times 10^{-4} kg.m^2$, then the corresponding discrete-time transfer function $H_1(q^{-1})$ given by (50),

$$H_1(q^{-1}) = \frac{0.007480q^{-1} + 0.01334q^{-2} + 0.0007376q^{-3}}{1 - 1.948q^{-1} + 0.954q^{-2} - 0.006152q^{-3}} \tag{50}$$

- a model representing the position of the plate below the position by default for: $k_s = 1.384 \times 10^{-3} kg.m^2$, then the corresponding discrete-time transfer function $H_2(q^{-1})$ given by (51),

$$H_2(q^{-1}) = \frac{0.007479q^{-1} + 0.01333q^{-2} + 0.0007376q^{-3}}{1 - 1.946q^{-1} + 0.954q^{-2} - 0.006152q^{-3}} \tag{51}$$

6.2 Simulation results

In order to test the proposed fault tolerant control approach, tow models of an ETV in actuator fault and an unknown disturbance case are given as state space formulation as

$$x(k+1) = A_j x(k) + B_j u(k) + E_j d(k) + B_j f_a(k) \tag{52}$$

$$y(k) = C_j x(k)$$

where the parameter matrices and vectors of each model are given as

$$A_1 = \begin{pmatrix} 1.948 & -0.954 & 0.006152 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{53}$$

$$A_2 = \begin{pmatrix} 1.946 & -0.954 & 0.006152 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{54}$$

$$B_1 = B_2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T \tag{55}$$

$$C_1 = \begin{pmatrix} 0.007480 & 0.01334 & 0.0007376 \end{pmatrix} \tag{56}$$

$$C_2 = \begin{pmatrix} 0.007479 & 0.01333 & 0.0007376 \end{pmatrix} \tag{57}$$

$$E_1 = E_2 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T \tag{58}$$

The studied system fulfills the conditions of Lemma 1.

$$\begin{cases} rank(C_1 E_1) = rank(E_1) = 1 \\ rank(C_2 E_2) = rank(E_2) = 1 \end{cases} \tag{59}$$

then $(C_1 A_{11})$ and $(C_2 A_{21})$ are both observable.
By applying Theorem, the system is stable asymptotically for any switching signal for symmetric matrice P definite positive given by (60).

$$P = \text{e - 010} \begin{pmatrix} 0.1429 & -0.0857 & 0.1903 \\ -0.0857 & 0.5085 & 0.0494 \\ 0.1903 & 0.0494 & 0.4251 \end{pmatrix} \tag{60}$$

Then, the rest parameters of UIOs are calculated as below

$$F_1 = \begin{pmatrix} -0.0266 & 0.8566 & -1.5415 \\ -0.0078 & 0.1954 & -0.3515 \\ 0.0165 & -0.4329 & 0.7788 \end{pmatrix} \tag{61}$$

$$F_2 = \begin{pmatrix} -0.0233 & 0.8717 & -1.5349 \\ -0.0071 & 0.1998 & -0.3515 \\ 0.0148 & -0.4384 & 0.7715 \end{pmatrix} \tag{62}$$

$$T_1 = \begin{pmatrix} 0.9657 & -0.6187 & -0.3470 \\ -0.0343 & 0.3813 & -0.3470 \\ -0.0343 & -0.6187 & 0.6530 \end{pmatrix} \tag{63}$$

$$T_2 = \begin{pmatrix} 0.9672 & -0.6196 & -0.3476 \\ -0.0328 & 0.3804 & -0.3476 \\ -0.0328 & -0.6196 & 0.6524 \end{pmatrix} \tag{64}$$

$$K_1 = \begin{pmatrix} 0.0002 \\ 0.0031 \\ -0.0056 \end{pmatrix}, K_2 = \begin{pmatrix} 0.0182 \\ 0.0208 \\ -0.0389 \end{pmatrix} \tag{65}$$

$$H_1 = \begin{pmatrix} 46.3945 \\ 46.3945 \\ 46.3945 \end{pmatrix}, H_2 = \begin{pmatrix} 46.4707 \\ 46.4707 \\ 46.4707 \end{pmatrix} \tag{66}$$

In this study, the method of tuning PID-type FLC parameters using GA is based on minimizing the ISE. If $y_j^d(k)$ is the desired flat trajectory and $\hat{x}_j(k)$ the estimated state, then we have

$$\varepsilon_j(k) = -K_{c,j} \hat{x}_j(k) + y_j^d(k) \tag{67}$$

For the ISE defined by

$$ISE = \int_0^t \varepsilon_j^2(k) dt \tag{68}$$

In this paper, the considered fitness function is taken as inverse of this error, i.e. the following performance index

$$fitness\ value = \frac{1}{ISE} \tag{69}$$

Then, the obtained optimum values of the input/output scaling factors $(K_e, K_d, \alpha, \beta)$, using genetic algorithm are given as follows: $K_e = 0.2006$, $K_d = 0.8071$, $\alpha = 0.1146$ and $\beta = 0.2063$.
The desired discrete time flat trajectory $z_j^d(k)$, with $j = \{1, 2\}$ can be computed according to the following polynomial form

$$z_j^d(k) = \begin{cases} \frac{cst1}{B_j(1)}, \text{ if } 0 \leqslant k \leqslant k_0 \\ Poly_{1,j}(k), \text{ if } k_0 < k \leqslant k_1 \\ \frac{cst2}{B_j(1)}, \text{ if } k_1 < k \leqslant k_2 \\ Poly_{2,j}(k), \text{ if } k_2 < k \leqslant k_3 \\ \frac{cst1}{B_j(1)}, \text{ if } k > k_3 \end{cases} \tag{70}$$

where $cst1$ and $cst2$ are constant parameters, $k_0 = 3s$, $k_1 = 6s$, $k_2 = 10s$ and $k_3 = 15s$ are the instants of transitions, $B_j(1)$ is the static gain between the flat output $z_j(k)$ and the output signal $y_j(k)$ for each model and $Poly_{1,j}(k)$ and $Poly_{2,j}(k)$ are polynomials calculated using the technique of polynomial interpolation.
The desired trajectories $y_1^d(k)$ and $y_2^d(k)$ are then given in Figure 4.

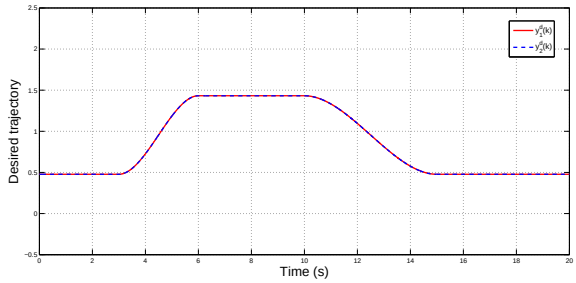


Figure 4: Desired trajectories $y_1^d(k)$ and $y_2^d(k)$

Let us consider the fault vector

$$f_a(k) = \begin{bmatrix} f_{a1}^T(k) \\ f_{a2}^T(k) \end{bmatrix} \tag{71}$$

such as $f_{a2} = 0$ and f_{a1} defined as follows

$$f_{a1}(k) = \begin{cases} 0, & k \in [0, 10] \\ -0.1\sin(0.314k), & k \in]10, 13] \\ 0, & k \in]13, 20] \end{cases} \tag{72}$$

The simulation results illustrated in Figure 5 and Figure 6, show some oscillations for the outputs signals and for the tracking error due to the unknown input. We remark that the system's responses with and without FTC track desired trajectories with disturbances rejection.

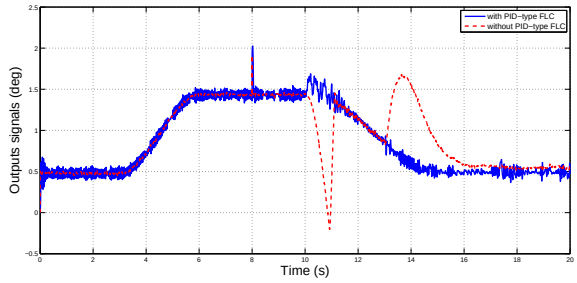


Figure 5: System outputs in actuator fault case with and without PID-type FLC

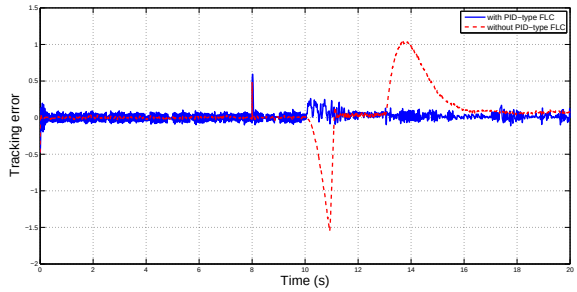


Figure 6: Tracking error in actuator fault case with and without PID-type FLC

At the time of actutor faults occurrence $k = 10s$, the system's behavior was changed. The system without FTC becomes unstable, whereas for the same reference input and by using the FTC, the system remains stable and the tracking error have a small deviation from zero which shows the effectiveness of the proposed FTC approach.

Figure 7, Figure 8 and Figure 9 show the residual values generated using UIOs and the switched signal. The switching between the two models is acheived based on residual values comparision.

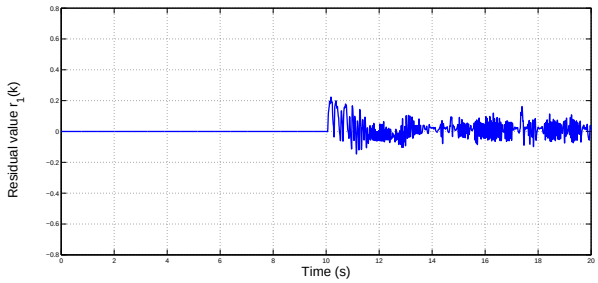


Figure 7: Residual value $r_1(k)$ in actuator fault case with PID-type FLC

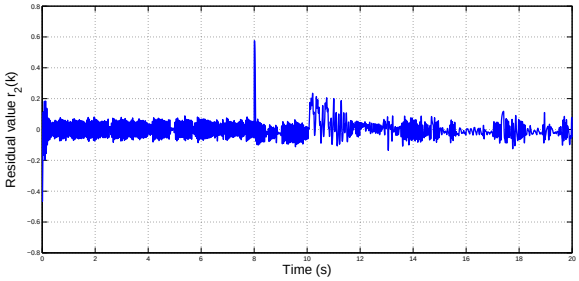


Figure 8: Residual value $r_2(k)$ in actuator fault case with PID-type FLC

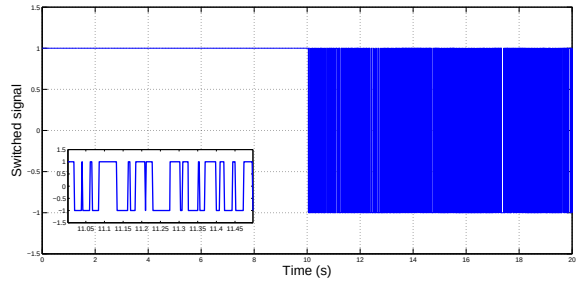


Figure 9: Switched signal in actuator fault case with PID-type FLC

7 Conclusion

In this paper, a new fault tolerant control law based on PID-type FLC is designed for the nonlinear complex system ETV, modelized by a multimodel structure. The approach is based on the use of a reference

model generated using the flatness property. The proposed control law is, then, designed to minimize the error between the desired flat trajectory and the estimated state, generated using UIOs, even in the presence of actuator faults based on minimizing the ISE.

References

1. B.Y. Huo, Y.M. Li and S.C. Tong, Fuzzy adaptive fault-tolerant output feedback control of multi-input and multi-output non-linear systems in strict-feedback form, IET Control Theory and Applications, **6**, 2704–2715, 2012.
2. Q.K. Shen, B. Jiang and V. Cocquempot, Fuzzy logic system-based adaptive fault-tolerant control for near-space vehicle attitude dynamics with actuator faults, IEEE Transactions on Fuzzy Systems, **21**, 289–300, 2013.
3. S.C. Tong, B.Y. Huo and Y.M. Li, Observer-based adaptive decentralized fuzzy fault-tolerant control of nonlinear large-scale systems with actuator failures, IEEE Transactions on Fuzzy Systems, **2**,1–15, 2014.
4. D. Ichalal, B. Marx, J. Ragot and D. Maquin, Fault tolerant control for Takagi-Sugeno systems with unmeasurable premise variables by trajectory tracking, IEEE International Symposium on Industrial Electronics (ISIE), Bari, 2010.
5. M. Blanke, M. Kinnaert, J. Lunze and M. Staroswiecki, Diagnosis and Fault-Tolerant Control, Springer-Verlag Berlin Heidelberg, 2006.
6. R. Nouailletas, D. Koeing and E. Mendes, LMI design of a switched observer with model uncertainty: Application to a hysteresis system, Proc. of the 46th IEEE CDC, New Orleans, 6298–6303, 2007.
7. D. Wang, W. Wang and P. Shi, Robust fault detection for switched systems with state delays, IEEE Trans. on Systems and Cybernetics-B, **39**, 800–805, 2009.
8. H. Yingxue, T. Shaocheng and L. Yongming, Adaptive fuzzy backstepping control for a class of switched nonlinear systems with actuator faults, International Journal of Systems Science, **47**, 3581–3590, 2015.
9. A. Benzaouia, M. Ouladsine, A. Naamane and B. Ananou, Fault detection for uncertain delayed switching discrete-time systems, International Journal of Innovative Computing, Information and Control, **8**, 1–11, 2012.
10. M. Oudghiri, M. Chadli and A. El Hajjaji, Robust observer-based fault tolerant control for vehicle lateral dynamics, International Journal of Vehicle Design, **48**, 173–189, 2008.
11. D. Ichalal, B. Marx, D. Maquin and J. Ragot, Nonlinear observer based sensor fault tolerant control for nonlinear systems, 8th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Mexico, 2012.
12. M. Benrejeb, Stability Study of Two Level Hierarchical Nonlinear Systems, 12th IFAC Symposium on Large Scale Systems: Theory and Applications, **43** (8), 30–41 Villeneuve d'Ascq, 2010.
13. J. Chen, R.J. Patton and H.Y. Zhang, Design of unknown input observers and robust fault detection filters, International Journal of Control, **63**, 85–105, 1996.
14. W. Jiawei, S. Yi and Z. Miao, Fault Detection for Linear Switched Systems Using Unknown Input Observers, Proceedings of the 33rd Chinese Control Conference, Nanjing, 2014.
15. W. Gritli, H. Gharsallaoui and M. Benrejeb, Fault Detection Based on Unknown Input Observers for Switched Discrete-Time Systems, Int. Conf. on Advanced Systems and Electrical Technologies IC-ASET, Hammamet, 2017.
16. S. Bouallègue, M. Ayadi, J. Haggége and M. Benrejeb, A PID type fuzzy controller design with flatness-based planning trajectory for a DC drive, IEEE International Symposium on Industrial Electronics, Cambridge, 1197–1202, 2008.
17. M. Guzelkaya, I. Eksin and E. Yesil, Self-tuning of PID-type fuzzy logic controller coefficients via relative rate observer, Engineering Applications of Artificial Intelligence, **16**, 227–236, 2003.
18. W. Gritli, H. Gharsallaoui and M. Benrejeb, PID-type Fuzzy Scaling Factors Tuning Using Genetic Algorithm and Simulink Design Optimization for Electronic Throttle Valve, International Conference on Control, Decision and Information Technologies CoDIT, Malta, 2016.
19. M. Fliess, J. Lévine, P. Martin and P. Rouchon, On differentially flat non linear systems, In Proc. IFAC-Symposium NOLCOS, pp. 408-413, Bordeaux, 1992.
20. M. Lebbal, H. Chafouk, G. Hoblos and D. Lefebvre, Modelling and Identification of Non-Linear Systems by a Multimodel Approach: Application to a Throttle Valve, International Journal Information and Systems Science, **3**, 67–87, 2007.
21. C. Yang, Model-based analysis and tuning of electronic throttle controllers, Visteon Corporation, SAE 2004 World Congress & Exhibition, 63–67, 2004.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 194-202(2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Review on security issues in RFID systems

Mohamed El Beqqal*, Mostafa Azizi

University Mohamed first, ESTO, 60000, Morocco

ARTICLEINFO

Article history:
Received: 11 November, 2017
Accepted: 01 December, 2017
Online: 14 December, 2017

Keywords:
RFID System
Classification
Security
Privacy
Countermeasures

ABSTRACT

Radio frequency Identification (RFID) is currently considered as one of the most used technologies for an automatic identification of objects or people. Based on a combination of tags and readers, RFID technology has widely been applied in various areas including supply chain, production and traffic control systems. However, despite of its numerous advantages, the technology brings out many challenges and concerns still not being attracting more and more researchers especially the security and privacy issues. In this paper, we review some of the recent research works using RFID solutions and dealing with security and privacy issues, we define our specific parameters and requirements allowing us to classify for each work which part of the RFID system is being secured, the solutions and the techniques used besides the conformity to RFID standards. Finally, we present briefly a solution that consists of combining RFID with smartcard based biometric to enhance security especially in access control scenarios. Hence the result of our study aims to give a clear vision of available solutions and techniques used to prevent and secure the RFID system from specific threats and attacks.

1. Introduction

This paper is an extension of work originally presented in Wireless Technologies, Embedded and Intelligent Systems (WITS) [1]. It concerns the security in Radio Frequency Identification (RFID) systems. The implementation of this technology in many industrial fields has recently attracted a lot of attention. Due to its important role of identifying people as well as objects, and data tracking, RFID is becoming more and more popular as an important topic of research. In particular, with the emergence of the Internet of things (IoT) paradigm that refers to a kind of networks that links physical objects to the internet to exchange data [2] and where RFID plays a key role. Besides, RFID is considered to be the next generation of barcode technology which is the case in many countries [2]. The use of RFID technology in several domains including supply chain, healthcare, transportation, education and many other fields have significant advantages including the flexible way of deployment with low cost, the possibility of integration and cooperation with Wireless Sensor Network nodes [2]. However, despite its numerous advantages, RFID technology still brings out many challenges including technical problems, customer privacy issues, coexistence of heterogeneous RFID standards [3] and more importantly security and privacy concerns. In fact, an RFID system consists of tags that store the data, readers interacting with the tags and transmitting information to backend servers and finally a network canal that manages this communication [4]. This architecture multiplies the sources of security problems and threats.

Several types of attacks can target the RFID system depending on each part is more vulnerable to the attacker. For example, the modification of the data, Cloning technique, and impersonation can be applied at level tags, whereas the reverse engineering, Eavesdropping Side Channel Attacks can affect the communication between backend servers, readers and tags [5].

In the research field, many of solutions and techniques assuring security and privacy come with processing and cryptographic algorithms such as Elliptic Curve Cryptographic (ECC) technique or mutual authentication mode between all RFID system parts. Most of these techniques remain far from being implemented in practice due to the limited resources of storing in tags and the low capacity of processing and executing complex operations. Furthermore, RFID customers require some guaranties that data contained in tags or exchanged remain private and highly protected against tracking and customer identity reconnaissance. A guarantee must be provided to customers aiming to adopt RFID in their own goals by ensuring that their private data stored in the tags will remain private and protected against spying and tracking

*Corresponding Author: Mohamed El Beqqal, elbeqqal.mohamed@gmail.com

This work presents a clear vision about the concerns of security and privacy in an RFID system. Precisely we compare recent works dealing with security and privacy in an RFID system according to which layer the author is aiming to secure. To achieve this goal, we define some parameters and requirements such as CAI (Confidentiality, Availability and Integrity) based on the information suggested in [6]. We present for each work the category and the technique of solution used to deal with a typical defined attack. In addition, we verify for each work the application of RFID standards. Finally, we propose a device added scheme solution that we have used to enhance security in the access control scenario.

This paper is organized in 5 sections. The Section 2 summarizes the basics components of an RFID system, some examples of applications and a summarized review on most challenges and research directions where we concentrate on the problem of security. The section 3 presents our comparative study based on our parameters defining this study. The section 4 illustrates the smart card based biometrics solution to enhance RFID security and finally section 5 concludes this paper.

2. Background on RFID technology

In this part, we present the components of a typical RFID system, besides the related security challenges.

2.1. RFID Basics

An RFID system consists of 6 main components as shown in Figure 1:

- Tag: A tag is the data carrier and normally contains the ID number, and unique EPC code programmed into the tag.
- Reader & antenna: A reader captures the data provided by the tag when tags come in range of the area covered by the reader using its antenna.
- Middleware: The middleware can be a software as well as hardware dedicated to process data captured by the Reader, then dispatch this information to backend servers.
- Backend servers: The backend servers are the last station to make use of the data collected from RFID components. The information collected can be stored in database or sent to other systems for reporting and further analysis.
- Network infrastructure: Stands for the link between all RFID components which represents the main sources of RFID security and privacy threats.

2.2. Application domains

Animal tracking

RFID tagging systems are considered as powerful tools for animal identification management and tracking. Their benefits go beyond controlling the spread of diseases and ensuring quality-assured and safe food for consumers.
As an example of the application of RFID technology for animal tracking, an approach enabling an effective localization and tracking of small-sized laboratory animals was suggested in [7]

Logistics

The implementation of RFID technologies in logistics environments is increasingly attracting. The investment in this technology has shown that RFID is highly promising for identifying an object uniquely and providing the capability for "complete traceability" [8]. Moreover, it has been recognized as an effective solution to eliminate or reduce inventory inaccuracies [9].

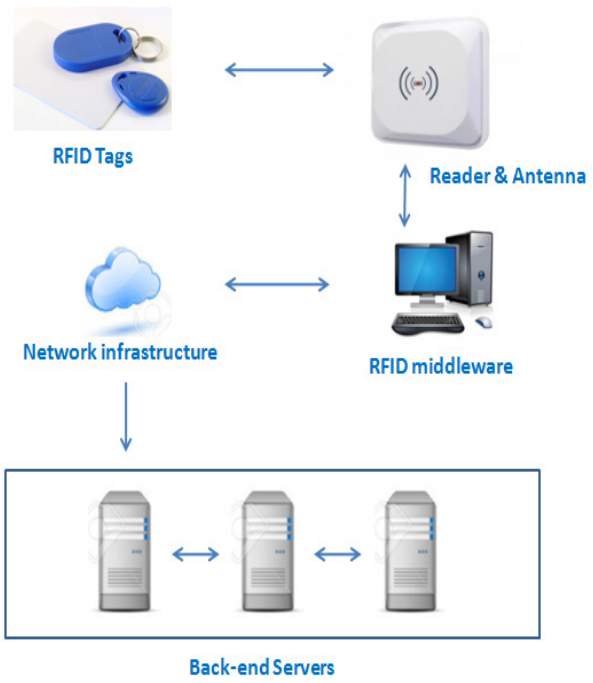


Figure 1. RFID system components

Intelligent transportation

The transport domain has widely benefited from the RFID adaptation due the ease of deployment of intelligence into each vehicle with a low cost approach [11].A vehicle emission inspection and notification system aiming to help daily monitoring of engine emissions through RFID devices was proposed in [12]. Also, a wireless system based on RFID technology was used to manage taxi fleets at the airports, train stations, bus terminals and all the taxis in cities [13]. RFID is also used in tolls collection as shown in Figure 2.

Healthcare

The use of RFID technology in the healthcare domain is making the patient's quality of care better. In fact, besides the collecting of information related to the patient as his presence inside a room and his health but also environment parameters can be detected such as temperature, humidity, and the presence of toxic gas [14].

In the same context, several efforts and research works have been made in this area. Exemplary, authors in [15] suggest a passive RFID platform for monitoring people during the night which deploys a long-range UHF RFID reader, wearable tags

integrated into clothes, and ambient tags dispersed in the environment, as well as a software engine for real-time processing and with warning modules [15].

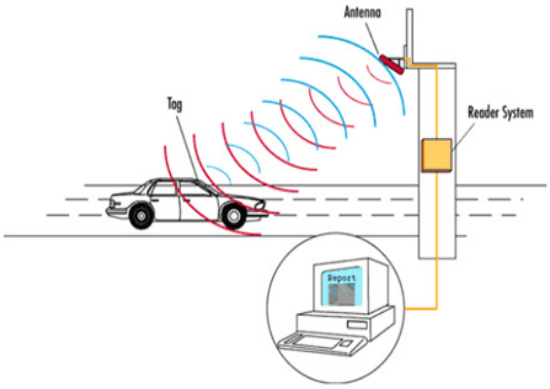


Figure 2. A toll collection using RFID [13]

2.3. RFID major issues

We address in this section some of the prominent issues related to RFID technologies and that still remain challenging topics of interest for a vast number of researchers.

Several works tried to study and gather the issues related the RFID to provide a baseline for researchers to understand the different sources as well as measures that can ensure the safety of sensitive data but more importantly to satisfy customers' needs and requirements in terms of efficiency, security and privacy. Among the most RFID challenges, we found technical problems, customer privacy issues and appliance to RFID standards [3]. In our work, we will focus on the last two problems which were taken into account in our comparative study (Section 3).

Technical problems

Among the biggest challenges during the implementation of an RFID system is studying the physical characteristics and capabilities of its components. In fact, the limited resource of storing information inside the tag, the low data processing and execution of complex operations can be a serious problem of performance and lack of mandatory features. In addition, due to the star architecture of the system, the reader can be the point of crush of the whole system, this problem occurs when a large number of items with RFID tags are energized by an RFID reader at the same time which leads to a confusion at the level of the reader and prevent it from tag scanning. A similar problem occurs when the coverage area of one RFID reader overlaps with another reader which leads to a signal interference and multiple reads of the same tag.

Standardization problems

Several efforts were dedicated to solve the problem of lacking of common RFID standards but the issue is still remaining in the industrial field. In fact, different standards coexist in parallel with divergent interests among which there is ISO 11784 that is utilized in the tracking of RFID cattle, the ISO 11785 for interface protocol, ISO 14443 for application of smart cards in payments and ISO 15693 that is applied to vicinity cards. The development of the electronic product codes (EPC) has been conceived to fill this need.

Obviously, the fact each of the existing standards is associated to a specific part of the RFID system make it different to deploy and protect data within an RFID system where tags standards can not comply with the communication standards for instance. Thus, it is of upmost importance to take into account the compliance to standards when implementing any RFID-based solution.

Security and privacy

Among the advantages of the deployment of RFID system is rendering the identification task automatically also gaining on performance and the ease of execution. All these benefits are not very useful is security and privacy are not guaranteed. For this reason, while designing an RFID system, many parameters should be taken into account to ensure a primary level of security, In fact, the data inside the tag, exchanged with the reader or dispatched to the back-end servers must be confidential, available and unchangeable (CIA characteristics). However, ensuring a full secure system which means keeping the whole system safe and operational, it remains a difficult task due to the several sources of attacks and threats that can coexist in parallel in some scenarios.

Table 1. RFID Typical attacks

Typical attack	The attack principle
Eavesdropping	This attack consists of listening to the network and sometimes recording for example all exchanges between tags and reader. The main goal of this attack is collecting data and taking advantage of the information gathered.
Denial of service	The denial of service consists of rendering equipment unavailable in the RFID system. An example of using this attack is spamming a reader with specific tags frequency requests which will make the reader out of service.
Cloning	The cloning attack is mainly to reproduce RFID tags. Using reverse engineering techniques to extract all the cart properties, precisely, the secrets keys, the tag can be modified and duplicated.
Tracking	The tracking attack consists of associating a tag to a person without his will and track his presence and movement as long as the attacker's reader is in range with the tag

During our research, many papers mentioned in their solutions some typical attacks that can take place in RFID systems, in the Table 1 below, we list some of most known attacks.

Many countermeasures and capabilities are proposed in literature to ensure security of the RFID system such as Pseudo-random based solution, Anonymous-ID scheme, symmetric and asymmetric cryptographic algorithms and others hashing based schemes [2].

Apart from the security aspect, the privacy is considered as a critical issue aiming to protect personal information. The expansion of the technology in several domains is not passing

without risks regarding the privacy the person since his identity information are associated to RFID chips that can be tracked and identified without his knowledge. As explained in table1 and due to the small size if RFID tags, they can be nested in personnel objects and clothes without being detected. The person movement and information are tracked and analyzed each time he come in range of the compatible reader area.

Some countermeasures and solutions were developed to protect personal data and reduce tracking capabilities. The table below illustrates the main techniques proposed in the literature studied for our classification.

Table 2. Some main countermeasures for privacy

Protection method	***How this technique improves privacy protection***
Delegation Tree	The principle of this method is delegating the control of reading tags depending on which privacy policy is assigned to each reader and tag. These specifications can be stored in database for example.
Protocol added *S*chemes	The technique consists of integrating a new coding scheme inside RFID tags and readers in order to create a specific protocol of communication; in this way, external readers cannot access to the network tags.
Tag killing	The tag killing method consists of destroying the content of tag after this one is no longer need to be used. The retailer can use the kill command after entering the right pin code. This technique can be used by the sellers after their products leave the store.
XOR encryption and PRNG	The XOR encryption and PRNG method is based on using a randomized protocol at each communication. This technique is powerful in the way to counter listening attacks such as Eavesdropping which will make the communication with tag a difficult task.
Blocker Tag	The Blocker tag technique aim to create an inductive field by the tag which will block communication between the tag and suspicious readers. Many occurrences of ID tag will be generated in order to hide the real one for the reader.

3. The comparative study

In this section, we will present the goal of our study besides the parameters defined for classification used to compare the selected academic works.

3.1. Motivation and design goal

Several levels of security can be applied to a typical RFID system, besides the nature of its architecture leads us to classify the solutions depending on each part of the system the attacks can take place. Furthermore, we refer to [6] to consider the three security main classes to classify the security threats in the reviewed works: Confidentiality, availability and integrity (CAI).

As defined in [6], the RFID system can be presented as three main layers: *1) RFID Edge Hardware Layer, 2) Communication Layer and 3) Backend Layer.* The possible threats and attacks related to each layer are mentioned based on CAI properties. In our work, besides classifying the attacks related to RFID security, we present also the solutions and techniques used to counter these problems. In [5], most of protection techniques can belong to specific categories as presented in Figure 3:

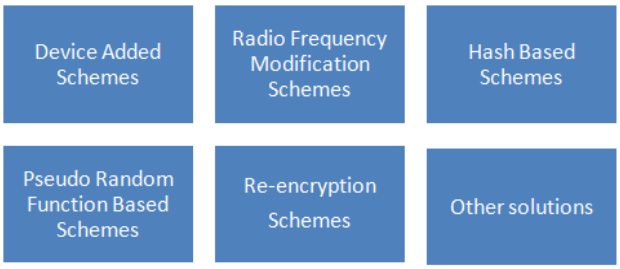


Figure 3. RFID security classes

In our study, we focus on recent works dealing with security and privacy, and we classify them according to the layers previously (attacks and solutions). In addition, we take into consideration other security requirements and parameters rather than CAI (untraceability, non-repudiation, forward secrecy, anonymity …).

Furthermore, we present the type of solutions and techniques dedicated for each layer and the attacks that can be related. At the end, we verify for each studied work the use and conformity to RFID standards. In this way, the following parameters are taken into account in our comparative study:

- *Privacy-focus / Security-focus*: Indicates whether the solution proposed take into account the security or consider privacy as the main concern.
- *Problem description and classification*: this parameter indicated which part of the system the problem can belong to. For this, we refer to the classification explained above.
- *Parameters of Security/Privacy considered:* Here we extract the security/privacy parameters considered by each author. Some papers focus on traditional security goals (confidentiality, integrity and availability) while other solutions take into account additional properties such as authorization, authentication, or forward secrecy.

Clearly, there are other properties to consider when authors focus in privacy as a main concern. These parameters include identity protection, anonymity and untraceability and. It is worth noting that these parameters are whether clearly cited in the paper or deducted.

- *Solution category*: As argued before, we have six mains classes of RFID solutions. For each of the twelve works studied we specify the class solution.
- *Techniques used:* For each class of solutions, we mention the techniques that have been exploited to ensure security and

privacy. Precisely, this includes cryptographic-based and no cryptographic-based methods.

- *Compliance to RFID standards?* Compliance to standards is considered as a primordial issue that need be highly highlighted. In this goal, for each studied work, we check the compliance to RFID or communication standards or if at least this issue is mentioned or an attention has been paid to by authors.
- *Examples of security attacks/ Threats considered:* this parameter indicates the security/ privacy attacks cited and treated in each paper. In the table below, the attack is defined for each part of the system it can occur and the solutions proposed to secure the system form this type of threats.

The objective of our comparative study as shown in Table 3, is not only presenting and classifying the security solutions and techniques regarding each of the recent studied paper but also making visible for researches and companies aiming to adopt an RFID system what are threats and attacks that should be taken into accounts while designing their systems.

As described before, the classification-based layer level helped us to locate the scope of each solution. Furthermore, during the study of the twelve works, we found that additional security / privacy parameters are considered rather than the traditional ones (CAI) in order to identify the several security/privacy threats. For Example, scalability is a primordial while aiming to enhance the system dimensions; some papers such as [24] consider scalability as a first topic while designing their solution. In fact, as mentioned in [22], this parameter remains among the real constraints of the majority of security solutions. Many techniques can be deployed to achieve scalability such as Checks handoff technique. As described in [24], this method was deployed to design a secure protocol supporting IoT.

Concerning techniques used, mutual authentication has taken more focus by authors as indicated in the works [20,16, 24] since, this security method is not only used in server's part but also between the reader and tags.

Concerning the security/privacy focus, many works try to ensure customer's privacy in their secured developed solutions and protocols. For instance, in [17, 20] intractability and an anonymous communication are taking into account to protect users' privacy. Thus, we noticed from the works presented in Table 3 that most of the solutions do not only focus on security or on privacy individually but these two concepts should be both take into consideration.

The second conclusion, we have subtracted from the Table 3 was that several works does not focus only on a single part of the system but they try to cover the whole layers (physical layer, communication layer, Back-end layer) that we have described in the section before.

For instance, authors in [18], propose two algorithms implemented by the Back-End server and a two-party protocol between the Back-End Server and a tag. In this way, two security methods were suggested in [20], the first contribution covers the backend layer by proposing two authentication (mutual and collaborative) schemes for both fixed and mobile readers, and the second contribution consisted at defining a secure channel between the reader and server in the communication layer. The main idea that we try to highlight is to not only paying attention to the data stored into tag or only securing communication between tags and reader but all the whole trajectory of data from data collecting to data storing in backend server.

As presented in Table 3, very few works we have studied focus on the device added schemes solution category. As an example, authors in [19] propose a solution that consist of changing the physical architecture of the RFID reader by dividing it into two different devices, an RF activator and a trusted shield device (TSD) playing an intermediary role between tags and reader for a secure communication [16].

In this way, the authors in [22] used the Physically Unclonable Functions (PUF) to secure data at level of tags (PUF is a security technology that seeks to introduce physical variation into individual devices taking part in cryptographic protocols [26]).

It is of upmost importance to tackle the issue of the RFID standardization. Indeed, and as illustrated in our comparative study, some works take into consideration the conformity of their solutions to RFID standards [12, 18, 19, 20, 20, 23, 24]. Intuitively, before designing any secure RFID solution it is crucial to verify whether this solution respects RFID standards or not.

Table 3: A comparative study of some security/privacysolutions in RFID systems

Characteristics / Proposed Solution	Privacy-focus / Security-focus	Problem description and classification (Layer)	Parameters of Security/Privacy considererd	Solution category	Techniques used	Compliance to RFID standards?	Examples of security attacks/ Threat considered
RFID-Tate: Efficient Security and Privacy Protection for Active RFID over IEEE 802.15.4 [16]	Both	Communication: A protocol for mutually authenticated (tag, reader) communication is suggested.	-Confidentiality -Availability - Authentication	Re-encryption	-Identity-based Encryption (IBE)	Yes : communication is performed over the standard IEEE 802.15.4f	- Cloning attacks. - Tag emulating, - Collision attack. - Spoofing - Replay attack -Sybil attack
A Secure Supply-Chain RFID System that Respects Your Privacy [17]	Both	Commiucation :p-Authors suggest a protocol between the reader and an individual tag	- Confidentiality - Integrity	Re-encryption	Public Key cryptography	Yes: (EPC tags standards)	-Reverse engineering

Secure Tag Search in RFID Systems Using Mobile Readers [18]	Both	Back-End: This paper suggests a search tag protocol; target backeEned Servers	- Confidentiality - Integrity - Anonymity - authorization	Pseudo-random based solution	XOR and 128bit PRNG operations	Yes : (EPC C1G2 compliance)	-Desynchronization, DoS attacks. -Side Channel Attacks: - Eavesdropping
Scalable RFID Security Framework and Protocol Supporting Internet of Things [24]	Security-Focus	Communication: A secure communication protocol and a framework supporting IoT	-Authentication. - Scalability - Confidentiality -Identity pretection -Adaptability	Other	Checks handoff (SCH) technique	Yes: EPCglobal standard	-Malware (Example SQLIA)
A New Security and Privacy Framework for RFID In Cloud Computing [25]	Both	Backend: A security and privacy model for RFID technology integrated to the Cloud computing (Servers) Was suggested.	- Confidentiality - Authentication	Re-encryption	Symmetric-key Cryptography	Not mentioned	No specified attacks was mentionned.
A secure ECC-based RFID authentication scheme integrated with ID-verifier transfer protocol [2]	Both	Both : Back-End: Two algorithms implemeted by the Back-End server are defined. Communication: A two-party protocol between the Back-End Server and a tag is proposed.	-Confidentiality - Availability - Forward security. -Anonymity Authorization, Authentication	Re-encryption	Public Key cryptography Elliptic curve cryptosystem	Mentioned but not applied	-Tag cloning attack. - Eavesdropping. - Replay attack. - Tag masquerade attack. -Denial-of-Service - Location tracking attack. - Server spoofing attack.
A Novel Coding Scheme for Secure Communications in Distributed RFID Systems [19]	Privacy-Focus	Physical Layer	Confidentiality - Integrity Anonymity	Device added scheme	- Reader changes architecture (New) Random Flipping and Random Jamming (RFRJ)	Only mentioned	- Random guessing attack. - Correlation attack. - Ghost-and-leech attack. - Eavesdropping.
A minimum disclosure approach to authentication and privacy in RFID systems [20]	Both	Both BackEnd: Two authentication (mutual and collaborative) schemes for both fixed and mobile reader are suggested. Communication: The channel between the reader and server is protected against attacks.	- Confidentiality - Untraceability Anonymity, Authentication, Forward secrecy	Pseudo-random based solution	Redundancy check (CRC) Pseudo random number generator (PRNG)	Yes (Conforme to EPC Class-1 Gen-2 spec-Ifications).	- Replay attack - Tag impersonation - Desynchronisation attacks.
Secure ownership transfer for multi-tag multi-owner passive RFID environment with individual-owner-privacy [12]	Both	Communication: Authors suggests multi-owner, multi-tag transfer protocol that achieves privacy-among-owners.	Confidentiality - Forward Untraceability Anonymity, Forward secrecy	Pseudo-random based solution	XOR and 128-bit pseudo-random number generators (PRNG)	Yes: complies EPC Global Class-1Gen-2 (C1G2)	-Tracking attacks - Replay Attacks - Denial of Service - Tag/Reader/Server Impersonation.
A Robust Grouping Proof Protocol for RFID EPC C1G2 Tags [21]	Security-focus	Communication: A grouping proof protocol ensuring a high level of security between tags and readers.	-Inegrity -Anonymity -Forward secrecy -Confidentiality	Pseudo-random based solution	(XOR) encryption and 128-bit pseudorandom number generators,	Yes: Complies EPC Global Class-1Gen-2 (C1G2)	-Replay Attacks -Denial of Service (DoS) Reader/Tag/Server impersonation Attack. -Active Attacks
Providing destructive privacy and scalability in RFID systems using PUFs [22]	Privacy-Focus	Communication: A private authentication protocol allowing a reader/tag communication.	- Confidentiality - Integrity - Authentication - Forward secrecy	Other	Physically Unclonable Functions (PUFs)	Not mentioned	- Compromising attack - Side-channel attacks
Computational Cost Analysis on Securing RFID Protocols Conforming to EPC Class-1 Generation-2 Standard [23]	Security-Focus	Both : A mutual authentication protocol seeking to achieve higher security level was proposed	-Confidentiality -Authentication - Forwad secrecy	Pseudo-random based solution	Pseudo-random number generators (PRNG)	Yes : Conform to EPC Class-1 Generation-2 standard (ISO 18000-6).	- DoS attacks - Reply attacks

4. RFID combining solution

4.1. Motivation

The comparative study presented in section 3, has shown many robust solutions and techniques that may counter several attacks, most of the techniques are based on cryptographic schemes and deal with direct attacks on physical layer of the RFID system components such as the tag and reader, whereas other techniques focus on the communication between readers and tags or reader and the backend servers to prevent the system from network attacks such as eavesdropping, man-in-the middle or desynchronization attacks.

Aiming to apply the RFID in the campus environment, we found that some scenarios are not totally secured. So, we propose to combine the RFID technology with other technologies to enhance security in the whole system.

4.2. Smart card based biometrics solution design

The RFID technology can be applied in many areas inside the university campus as shown below, in our work we focus on access control to exams rooms and sensitive areas like laboratories.

- Exams room
- Campus
- Research laboratories
- Library
- Staff offices
- Computer room

For instance, if we take the example of accessing to laboratories, the identity check must be verified before allowing the person access to the area. Based on the solutions proposed in Table 3, a cryptographic technique can be deployed in the inside the tag and the reader to secure the tag reading. However, the two main problems in our case are indicated in Figure 4.

For this reason, our method consists of enhancing the RFID technology by using smart card based biometrics solution. In fact, the RFID technology is seen as the fast-automatic technique of identification, whereas, the smart card check can support complex operations of cryptographic check as indicated in [27]. In addition, the biometric fingerprint will help us to check the true identity of the person aiming to access to the secured area. In Figure 5, we present the proposed system combining RFID with the smart card-based biometrics verification.

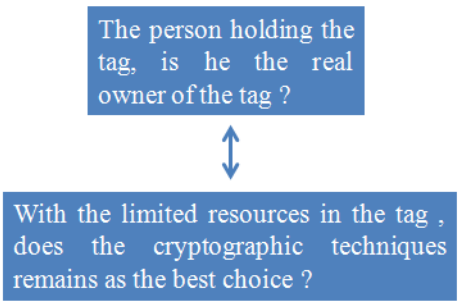


Figure 4. Access control issues in RFID system

As detailed in [27]. The step one consists of reading the RFID tag when it comes in range with the RFID reader. After that, in the step 2, the person's tag is sent to the local database to check if it belongs to the person's tags allowing him access the secure area. Once validated, the person inserts his smartcard and puts his fingerprint for validation. In this case, a first match-on-card is established comparing the user fingerprint with the other one stored in the smartcard. If the check is performed successfully, the fingerprint is sent to the database to verify that no changes were made on the couple (fingerprint, Tag ID). Finally, the door is opened if all the conditions are satisfied.

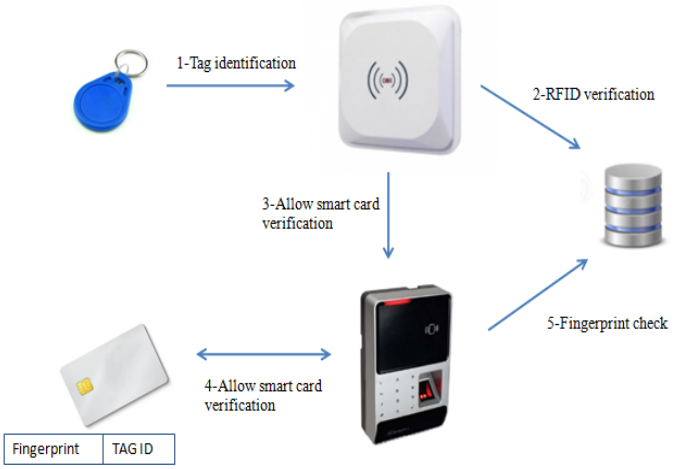


Figure 5. Components of the proposed system

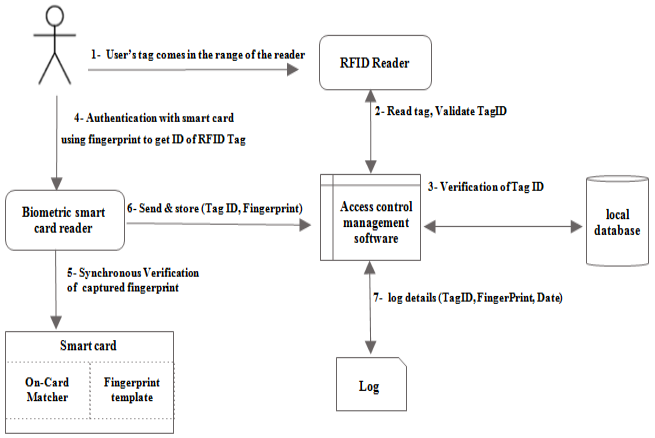


Figure 6. Flowchart of the access control system in sensitive areas [27]

In the proposed solution, we have chosen to combine smart card biometrics solutions with RFID in order to respect the limited resources and capabilities inside the tag for performing complex operations. Furthermore, even if the smart card is falsified or the person ask for authentication with a borrowed RFID tag, the verification explained above based multi-level of security will deny his access to the secured room. The flowchart in Figure 6 summarizes the sequence of actions during the authentication process. The system presented above meets the need presented in Figure 4 which are not guaranteed in the solutions presented in section 3.

5. Conclusion

In this paper, we have presented some recent works aiming to ensure security and Privacy in RFID systems. In particular, we tried to classify these solutions in order to provide a clear understanding of the different threats and risks related to security and privacy.

To achieve this goal, we compared some solutions according to the traditional security and privacy objectives. Also, we mentioned that additional parameters such as scalability and efficiency should be taken into account besides the traditional security and privacy objectives.

In addition, we proposed a combined solution using RFID and smartcard based biometrics to ensure performance and cover some security gaps detected while designing our RFID system. As future work, we plan to develop our device added scheme solution, by studying more the technical possibilities of its implementation inside the campus.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is performed inside the MATSI Lab., ESTO, University Mohammed First, Oujda (Morocco).

References

[1] M. El Beqqal and M. Azizi, "Classification of major security attacks against RFID systems," in 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), 2017. https://doi.org/10.1109/WITS.2017.7934622

[2] Yi-Pin Liao, Chih-Ming Hsiao, "A secure ECC-based RFID authentication scheme integrated with ID-verifier transfer protocol", Journal of Ad Hoc Networks, Volume 18, July 2014, Pages 133–146,2013. https://doi.org/10.1007/978-3-642-35473-1_1

[3] AL-Kassab, J., W.C. Rumsch. 2008. Challenges for RFID cross-industry standardization in the light of diverging industry requirements. IEEE Systems Journal 2(2) 170-177. https://doi.org/10.1109/JSYST.2008.921291

[4] F. Kamoun," RFID System Management:State-of-the Art and Open Research Issues", IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, VOL. 6, NO. 3, SEPTEMBER 2009 . https://doi.org/10.1109/TNSM.2009.03.090305

[5] K. Sabaragamu Koralalage, J. Cheng, "A Comparative Study of RFID Solutions for Security and Privacy: POP vs. Previous Solutions", International Conference on Information Security and Assurance, 2008. https://doi.org/10.1109/ISA.2008.89

[6] A. Mitrokotsa, M. Beye and P. Peris-Lopez, "Classification of RFID Threats based on Security Principles", GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES, 2009

[7] L. Catarinucci, R. Colella, L. Mainetti, L. Patrono, S. Pieretti, Ilaria Sergi, and L. Tarricone, "Smart RFID Antenna System for Indoor Tracking and Behavior Analysis of Small Animalsin Colony Cages", IEEE SENSORS JOURNAL, VOL. 14, NO. 4, APRIL 2014. https://doi.org/10.1109/JSEN.2013.2293594

[8] M. Alamgir Hossain, "Development of an integrated model for RFID extension", BUSINESS PROCESS MANAGEMENT JOURNAL· AUGUST 2014. https://doi.org/10.1108/BPMJ-04-2013-0055

[9] T. Fan, Feng Tao, S. Deng, S. Li, "Impact of RFID technology on supply chain decisions with inventory inaccuracies", Int. J. Production Economics, 2014. https://doi.org/10.1016/j.ijpe.2014.10.004

[10] Oh-Keun Ha Yong-Seok Song Kyung-Yong Chung Kang-Dae Lee · Dongjoo Park, "Relation model describing the effects of introducing RFID in the supply chain: evidence from the food and beverage industry in South Korea",Journal of Pers Ubiquit Comput, 2014. https://doi.org/10.1007/s00779-013-0675-x

[11] Lei Xie, Yafeng Yin, Athanasios V. Vasilakos,Sanglu Lu, "Managing RFID Data: Challenges, Opportunities and Solutions", IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 16, NO. 3, THIRD QUARTER 2014. https://doi.org/10.1109/SURV.2014.022614.00143

[12] Sundaresan, S. Doss, R., Wanlei Zhou, "Secure ownership transfer in multi-tag/multi-owner passive RFID systems", IEEE Global Communications Conference (GLOBECOM), 2013. https://doi.org/10.1109/GLOCOM.2013.6831513

[13] S. Samadi, "Applications and Opportunities for Radio FrequencyIdentification (RFID) Technology in Intelligent Transportation Systems: A Case Study", International Journal of Information and Electronics Engineering, Vol. 3, No. 3, May 2013. https://doi.org/10.7763/IJIEE.2013.V3.330

[14] Sara A, R. Lodato, S. Manzari, C. Occhiuzzi, and G. Marrocco, "RFID Technology for IoT-Based Personal Healthcare in Smart Spaces", IEEE INTERNET OF THINGS JOURNAL, VOL. 1, NO. 2, APRIL 2014. https://doi.org/10.1109/JIOT.2014.2313981

[15] C. Occhiuzz, Carmen Valleseb, S.Amendolab, S. Manzarib," NIGHT-Care: a passive RFID system for remote monitoring and control of overnight living environment", 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014). https://doi.org/10.1016/j.procs.2014.05.414

[16] Sadikin, M.F., Kyas, M., "RFID-tate: Efficient security and privacy protection for active RFID over IEEE 802.15.4, The 5th International Conference on Information, Intelligence, Systems and Applications, IISA 2014.in press. https://doi.org/10.1109/IISA.2014.6878787

[17] Arbit, A., Oren, Y. ; Wool, A, "A Secure Supply-Chain RFID System that Respects Your Privacy", Journal of Pervasive Computing, IEEE (Volume:13 , Issue: 2), 2014. https://doi.org/10.1109/MPRV.2014.22

[18] Sundaresan, S., Doss, R. ; Piramuthu, S. ; Wanlei Zhou, 'Secure Tag Search in RFID Systems Using Mobile Readers', Journal of Dependable and Secure Computing, IEEE Transactions on (Volume:12 , Issue: 2), 2014, in press. https://doi.org/10.1109/TDSC.2014.2302305

[19] Sakai, K, Sun, Min-Te ; Ku, Wei-Shinn ; Lai, T.H., "A Novel Coding Scheme for Secure Communications in Distributed RFID Systems",Computers, IEEE Transactions on (Volume:PP , Issue: 99), 2015, in press. https://doi.org/10.1109/TC.2015.2423671

[20] Robin Doss, , Wanlei Zhou, Saravanan Sundaresan, Shui Yu, Longxiang Gao , 'A minimum disclosure approach to authentication and privacy in RFID systems", journal of Computer Networks, Volume 56, Issue 15, Pages 3401–3416 pp. 544 - 549 , 2012. https://doi.org/10.1016/j.comnet.2012.06.018

[21] Sundaresan, S, Doss, R. ; Piramuthu, S. ; Wanlei Zhou," A Robust Grouping Proof Protocol for RFID EPC C1G2 Tags", Information Forensics and Security, IEEE Transactions on (Volume:9 , Issue: 6), 2014. https://doi.org/10.1109/TIFS.2014.2316338

[22] Mete Akgün, M. Ufuk Çağlayanb, "Providing destructive privacy and scalability in RFID systems using PUFs", Journal of Ad Hoc Networks, Volume 32, Pages 32–42, 2015, in press. https://doi.org/10.1016/j.adhoc.2015.02.001

[23] Avoine, G. Bingol, M.A. ; Carpent, X. ; Yalcin, S.B.O., "Privacy-Friendly Authentication in RFID Systems: On Sublinear Protocols Based on Symmetric-Key Cryptography". obile Computing, IEEE Transactions on (Volume:12 , Issue: 10), 2013. https://doi.org/10.1109/TMC.2012.174

[24] BR Ray, J Abawajy, M CHowdhury, "Scalable RFID security framework and protocol supporting Internet of Things", Computer Networks, vol. 67, pp. 89-103, 2014. https://doi.org/10.1016/j.comnet.2014.03.023

[25] Plos, T. Hutter, M. ; Feldhofer, M. ; Stiglic, M., "Security-Enabled Near-Field Communication Tag With Flexible Architecture Supporting Asymmetric Cryptography", Very Large Scale Integration (VLSI) Systems, IEEE Transactions on (Volume:21 , Issue: 11), 2013. https://doi.org/10.1109/TVLSI.2012.2227849

[26] D. Moriyama1, S. Matsuo1 and M.Yung, "PUF-Based RFID Authentication Secure and Private under Memory Leakage", IACR Cryptology ePrint Archive , 2013. https://doi.org/10.1016/j.adhoc.2015.02.001

[27] M. El Beqqal, M. A. Kasmi, and M. Azizi, "Access control system in campus combining RFID and biometric based smart card technologies," Adv. Intell. Syst. Comput., vol. 520, pp. 559–569, 2017. https://doi.org/10.1007/978-3-319-46568-5_56

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 203-216(2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Novel Analysis of Synchronous and Induction Generators in Parallel Operation Mode in an Isolated Electric System

Vinicius Zimmermann Silva*, Ângelo Jose Junqueira Rezek, Rafael Di Lorenzo Corrêa

Federal University of Itajubá UNIFEI, Energy and Electric System Institute ISEE, 37500903, Brazil

ARTICLEINFO

Article history:
Received: 29 October, 2017
Accepted: 01 December, 2017
Online: 15 December 2017

Keywords:
Induction Generator
Isolated Electric System
Generators in Parallel
Synchronous Generator

ABSTRACT

This paper presents an analysis of a parallel connection of one synchronous generator and one self-excited induction generator that feeds a resistive load and an Induction Motor. The system voltage and frequency are controlled by a voltage control loop and a speed control loop connected to synchronous generator. The induction generator speed is controlled by its primary machine that is fed by an autotransformer and a diodes bridge. Through by voltages applied by an adjustable tap autotransformer connected to induction generator's primary machine, it is possible widen the range of its shaft speed if compared with the shaft speed caused by only field flux variation method. Then, by the autotransformer method is possible to widen the speed and power limits from the induction generator what increases the induction generator contributions and relieves the power supply from synchronous generator. Analysis of generators power supply and its interactions in various operational scenarios are shown. The results enable comparisons of the two methods of induction generator speed control, either by auto-transformer method or by field flux variation method. The first results in larger range of speed and power from the induction generator. Therefore, it has more representative features of actual field conditions.

Nomenclature and Abbreviations

V_{SG}	Synchronous generator voltage
I_{SG}	Synchronous generator current
f_{SG}	Synchronous generator frequency
$I_{fieldSG}$	Synchronous generator field current
n_{SG}	Synchronous generator speed
DCM_{IG}	Direct current motor coupled with induction generator
DCM_{SG}	Direct current motor coupled with synchronous generator
DCG_{SG}	Direct current generator coupled with synchronous generator
I_{aDCMSG}	Armature current of direct current motor coupled with synchronous generator
I_{fDCMSG}	Field current of direct current motor coupled with synchronous generator
V_{IG}	Induction generator voltage
f_{IG}	Induction generator frequency
I_{IG}	Induction generator current
n_{IG}	Induction generator speed
I_{aDCMIG}	Current of direct current motor coupled with induction generator
I_{fDCMIG}	Field current of direct current motor coupled with induction generator
I_c	Capacitor bank current
I_{loadR}	Resistors banks current
Q	Reactive power
pF, cosϕ	Power factor
Xc	Capacitive reactance
SG	Synchronous generator
IG	Induction generator
PIG	Induction generator power
PSG	Synchronous generator power
P_{DCMSG}	DCM power input coupled with synchronous generator
P_{DCMIG}	DCM power input coupled with induction generator
I_{WSG}	SG active current
I_{WIG}	IG active current
Va_{DCMIG}	DCM armature voltage coupled with IG
Va_{DCMSG}	DCM armature voltage coupled with SG
Ra	Armature resistance

* Vinicius Zimmermann Silva, vinicius.zimmermann@yahoo.com.br

C	Conjugate
Xs	Synchronous reactance
E	Back electromotive force
V_{pn}	Phase neutral voltage
FPSO	Floating production storage offloading
IM	Induction Motor
IdwMT	Motor reactive current
IwMT	Motor active current

1. Introduction

This paper is an extension of work originally presented in PEDG conference [1].

The use of induction generators in generation systems is usual in wind plants [2-4] or eventually in small hydroelectric generation units [5]. Otherwise, from studies started in [1], it is shown a new knowledge border which consider the use of IG in parallel of SG in an isolated electric system. The first results and behavior of this kind of generation system were shown in [1] and are followed by this current article with original and deeper analysis.

Then, this work is composed of analysis over an expanded system that includes an induction motor, a resistive load, an autotransformer connected to a diodes bridge to feed the DCM_{IG} armature and then permits to increase manually the output electric power limits if compared to flux variation method (3) presented in [1]. Therefore, this article shows the IG supplying more active power and the induction motor inserted in experiment what enables a novel analysis of the system behavior.

This paper presents results that show additional remarkable characteristics by operation in parallel mode among one SG and one IG. Some characteristics, such as reduced weight and size, easier maintenance, and shorter manufacturing and delivery time are more associated with induction generators and are relevant to oil platforms or offshore installation concerns. Besides, it has absence of dc supply for excitation and better transient performances [5].

One of the potential motivation for this study is to develop a potential alternative capable of optimizing the main electric system currently adopted in oil platforms and floating production storage and offloading (FPSO) ships to become cheaper, simpler, lighter, and more efficient.

The potential application of this study is the replacement of one or two synchronous generators by one or two induction generators in oil offshore platforms or FPSO. A typical offshore electric system uses three or four turbo-generators in 13800 V, 60 Hz, driven by dual fuel (fuel gas or diesel oil) turbines. In typical operation, three main turbo-generators operate and can supply all unit consumers with the fourth in standby. During the peaks of load, the fourth main turbo-generator may be required to meet the demand. Therefore, the whole electric system shall be suitable for this operational condition.

Therefore, the study target is to analyze various aspects of topology and operating involving an induction generator in parallel with a synchronous generator for application in an isolated electric system, and to establish the operational viability, advantages and challenges.

2. Isolated Electric System

The isolated electric system was sized as shown in the following equations and data. The automatic controls of the system consist of a voltage control loop and a speed control loop. Both are inserted in the SG scheme as shown in Figure 5 and Figure 6. The dataplates of the principal equipment are shown in Tables 1 to 6.

The parameterization used in both electronic control boards is indicated in [1], except for the references of voltage and speed for each of the control loops which are respectively 65.5 and 65.2 as shown in Figure 9.

2.1. Dataplate

In tables 1 to 4 are shown the main equipment's dataplates and the tables 5 to 6 are shown the load's dataplates used in the experiments along this work.

Table 1: DMC_{IG} Dataplate

Direct Current Motor $_{IG}$				
220 V	7.72 A	1.7 kW	1500 rpm	600 $mA_{fieldmax}$

Table 2: DMC_{SG} Dataplate

Direct Current Motor $_{SG}$				
220 V	9.1 A	2.0 kW	1800 rpm	600 $mA_{fieldmax}$

Table 3: IG Dataplate

Induction Generator (IG)					
220 V	7.5 A	1.86 kW	1410 rpm	0.8 PF	50 Hz

Table 4: SG Dataplate

Synchronous Generator (SG)					
230 V	5.0 A	2.0 kVA	1800 rpm	0.8 PF	60 Hz
V_{field}: 220 V			$I_{fieldmax}$: 600 mA		

Table 5: Load Dataplate

Resistive Load		
Load 1 (kW)	Load 2 (kW)	Load 3 (kW)
2/3	2/3	2/3

Table 6: Load, Induction Motor Dataplate

Induction Motor			
0.37 kW	1715 rpm	Cos ϕ: 0.71	60 Hz

2.2. Equations - Part I

Follow the main equations for analysis comprehension and to establish a studies base:

- Direct Current Motor

$$n = \frac{E}{k * \emptyset} \quad (1)$$

$$n = \frac{Ua - (\sum Ra) * Ia}{k * \emptyset} \quad (2)$$

$$C_{conjugate} = k * \emptyset * Ia \quad (3)$$

- Synchronous Generator

$$\dot{E} = \dot{U}a + (Ra + jX_s) * \dot{I}_a \quad (4)$$

- Induction Generator

$$S_{slip}\% = \frac{n_s - n_{IG}}{n_s} * 100 \quad (5)$$

$$n_s = \frac{120 * f_{SG}}{P_{number\ of\ poles}} \quad (6)$$

2.3. Capacitor Bank Sizing

As informed at the IG dataplate

$$\cos\phi = 0.8, \text{then } \sin\phi = 0.6$$

The reactive power is calculated to attend the reactive demand of the Induction Machine:

$$Q = \sqrt{3} * V * I * sen\phi \quad (7)$$

$$Q = \sqrt{3} * 220 * 7.5 * 0.6 => 1714.7 \text{ VAr}$$

$$Q_{generated} = Fc * Q \quad (8)$$

Fc =1.2 [6]

For the machine coupled at a resistive load which requires 7.5A, it is necessary for the reactive power generation to be approximately 2057.6 VAr as demonstrated below.

$$Q_{generated} = 1.2 * 1714.7 => 2057.6 \text{ VAr} \quad (9)$$

$$Q = \frac{3*V^2}{X_c} \quad (10)$$

$$Q = \frac{3*220^2}{X_c} => X_c = \frac{3*220^2}{2057.6} = 70.6\ \Omega \quad (11)$$

$$X_c = 70.6\ \Omega$$

$$X_c = \frac{1}{2*\Pi*f*C} = \frac{1}{2*\Pi*50*C} => \quad (12)$$

$$C = \frac{1}{2*\Pi*50*70.6} \cong 40\mu F \text{ per phase} \quad (13)$$

2.4. Resistive Divider Sizing:

- Field Control Loop Resistive Divider as in Figure 5 and Figure 6.

$$(1\ k\Omega + r\_1) \rightarrow 300\ V_{(brigge\ output\ voltage)} \quad (14)$$

$$r_2(1\ k\Omega) \rightarrow 4V_{(MP\ 410T\ voltage\ limit)} \quad (15)$$

$$r_1 = 74\ k\Omega\ and\ r_2 = 1k\Omega \quad (16)$$

- Resistor Power Sizing:

$$P_{power} = {300^2}/{(75 * 10^3)} = 1.2\ W \quad (17)$$

- The resistors that were selected based on the sized resistor, were:

$$r_1 = 79.1\ k\Omega\ and\ r_2 = 947\Omega \quad (18)$$

- Speed Control Loop Resistive Divisor as in Figure 5 and Figure 6.

$$(1\ k\Omega + r_1) \rightarrow 36\ V_{(tacho\ generator\ output\ voltage)} \quad (19)$$

$$r_2(1\ k\Omega) \rightarrow 4V_{(MP\ 410T\ voltage\ limit)} \quad (20)$$

$$r_1 = 8\ k\Omega\ and\ r_2 = 1k\Omega \quad (21)$$

- Resistor Power sizing:

$$P_{power} = {36^2}/{(9 * 10^3)} = 1.14\ W \quad (22)$$

- The resistors chosen, based on the sized resistor, were

$$r_1 = 8{,}3\ k\Omega\ and\ r_2 = 947\Omega$$

2.5. Equations - Port II

Follow the system equations base to calculate the power and efficiencies shown in Table 7 and Table 8 for each generator and entire group of machines.

$$I_{SG}{}^2 = Iw_{SG}{}^2 + Ic_1{}^2 \quad (23)$$

$$I_{IG}{}^2 = Iw_{IG}{}^2 + Ic_2{}^2 \quad (24)$$

$$Ic - IdwMT = Ic_1 + Ic_2 \quad (25)$$

$$I_{wloadR} + IwMT = Iw_{SG} + Iw_{IG} \quad (26)$$

As I_{SG}, I_{IG}, Ic and I_{wloadR} are measured values, shown in Table 7 and Table 8, IdwMT(Motor's reactive current) and IwMT(Motor's active current) are values calculated in appendix 1. Then, the system has 4 four variables and 4-four equations. Then, for each scenario, the four variables Iw_{SG}, Ic_1, Iw_{IG} and Ic_2 were calculated by Matlab software and are shown in Table 7 and Table 8.

The entire group efficiency and the efficiency of each subgroup were calculated based on PIG, PSG, PDCMSG, and PDCMIG, as follows:

$$P_{SG} = \sqrt{3} * V_{SG} * Iw_{SG} \quad (27)$$

$$P_{IG} = \sqrt{3} * V_{IG} * Iw_{IG} \quad (28)$$

$$P_{DCMSG} = Va_{DCMSG} * Ia_{DCMSG} \quad (29)$$

$$P_{DCMIG} = Va_{DCMIG} * Ia_{DCMIG} \quad (30)$$

$$\eta_{group}\% = \frac{P_{SG}+P_{IG}}{P_{DCMSG}+P_{DCMIG}} * 100 \quad (31)$$

$$\eta_{SG}\% = \frac{P_{SG}}{P_{DCMSG}} * 100, \quad (32)$$

$$\eta_{IG}\% = \frac{P_{IG}}{P_{DCMIG}} * 100 \quad (33)$$

2.6. Equations - Part III

The equations e calculus used to determined the IM parameters and IM equivalent circuit as primary, secondary

impedances and currents, including IM currents showed in Table 8 are calculated in item 5 of appendix 1.

2.7. The Experiment and Schemes

The experiment was mounted in the laboratory as shown in Figure 1 and Figure 2. The detailed circuit is shown in Figure 6. Figure 2 shows another experiment view, including the taco generator and its connections. The detailed circuit is shown in Figure 6. Figure 3 shows the typical connections of both circuit board MP 410T which are implemented in this experiment for the voltage loop and speed loop. They are also indicated in Figure 5 and Figure 6.

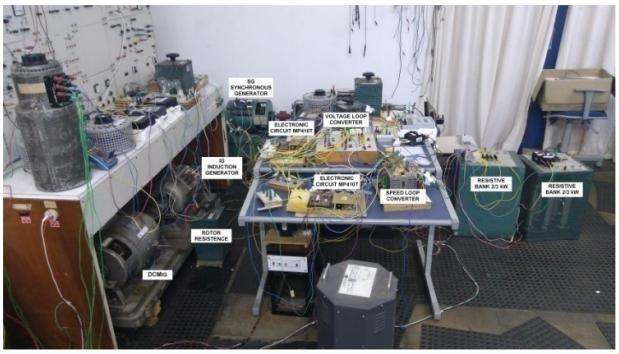


Figure 1 Laboratory assembly

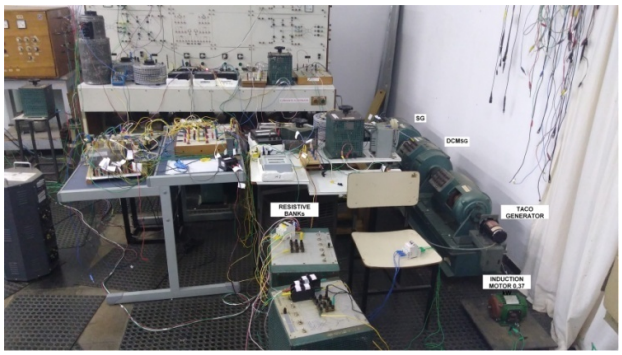


Figure 2 Another view of laboratory assembly

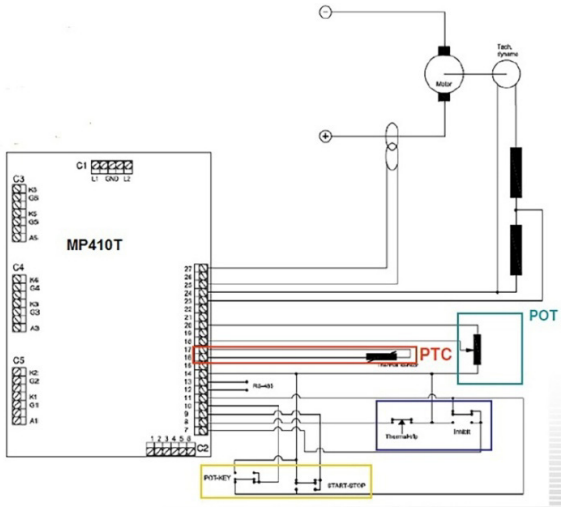


Figure 3. Connections arrangement of MP410T and devices

Figure 4 shows the 1B6C configuration used for the two thyristor converter bridges which were used in the control loops.

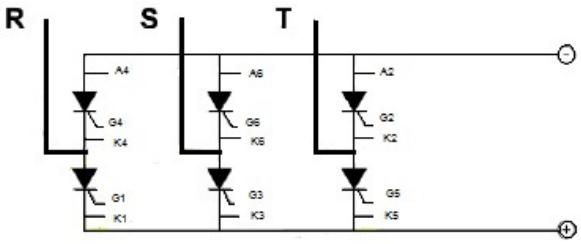


Figure 4. Converters configuration used in voltage and speed loops

As seen in [1], the power from IG, PIG, was limited because the I_{fDCMIG} reached the maximum viable value in accordance with DCM_{IG} current specifications. This limited speed from DCM_{IG} was a challenge because it was not showing a representative contribution from PIG. It was necessary to elevate the DCM_{IG} speed, n_{IG}.

Then, the methodology consists of implementing the comparative analysis of the power and efficiencies, starting with an analysis between the PIG from the scheme in [1] and PIG from the scheme in Figure 5 and Figure 6, covering the scenarios shown in Table 7 related to Figure 5 and Table 8 related to Figure 6. Finally, a comparative efficiency analysis was conducted for two subgroups, one composed of IG and DCM_{IG} and the other composed of SG and DCM_{SG} in scenarios related to the scheme in [1] and the schemes in Figure 5 related to Table 7 and Figure 6 related to Table 8.

Thus, the increase of DCM_{IG} speed, increase of PIG contributions and improvement of IG subgroup efficiencies are shown in results.

Figure 6 shows the scheme implemented in the laboratory to overcome the challenge related to limitation of IG speed [1] and bring the experiment closer to actual offshore conditions, taking into advantages of IG power capacity. It was used as an autotranformer connect to a diodes bridge to vary the voltage applied on the DCM_{IG} armature circuit and obtain a higher speed and PIG.

2.8. Voltage and Speed Control Loops

The Figure 7 and Figure 8 show the closed control loop used to control the SG speed and voltage via circuit boards.

Figure 9 (a) and Figure 9 (b) show the circuit boards with the respective working point or reference points defined during parameterization for the speed control loop, 65.2 and the voltage control loop, 65.5.

As [1], the master and slave behavior were verified by the generators, considering that SG is the master and IG is the slave. The frequency and voltage are determined by SG, while active power consumed at load is controlled by IG.

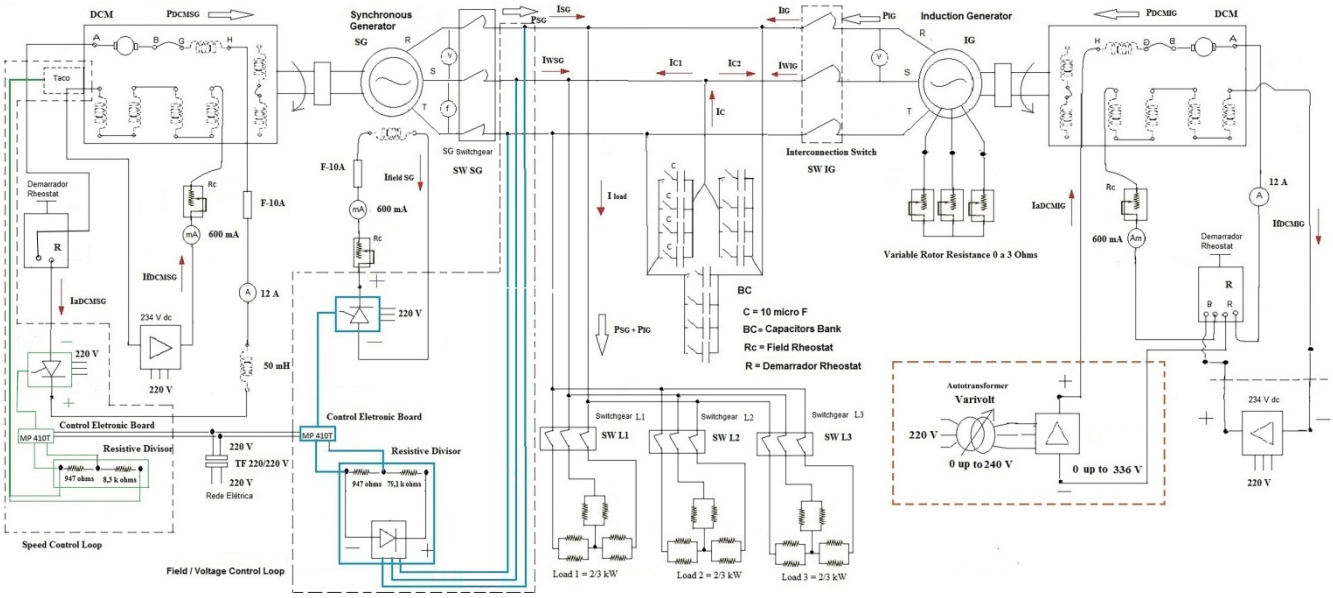


Figure 5 Synchronous generator in parallel with induction generator and its dc motor_ 234 Vdc steady source for DCM_{IG}

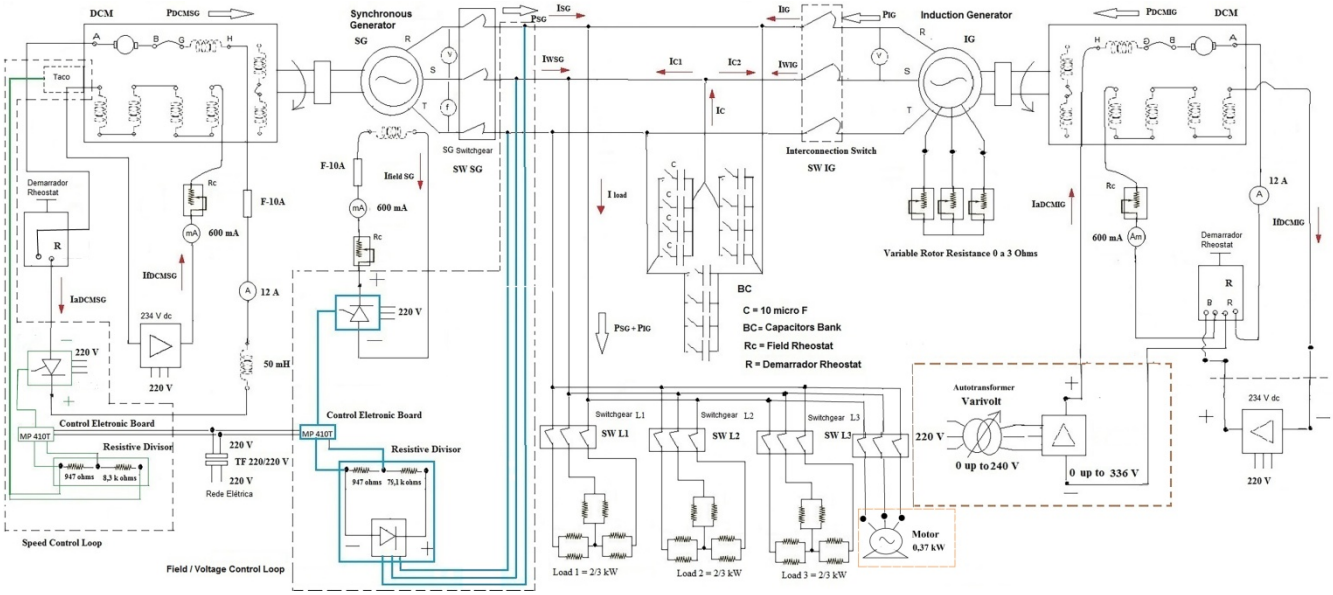


Figure 6: Synchronous generator in parallel with induction generator and its dc Motors _ Vdc variable source for DCM_{IG}

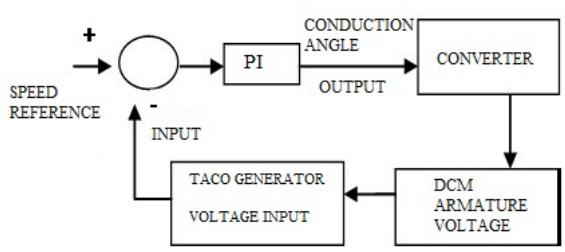


Figure 7. Speed Control Loop

Note: For both regulators, the proportional gain P was experimentally set at 0,010 and, in the same manner, the time constant was set at 0,04 s as in [1].

3. Experiment Data and Results

3.1. Experiment Data

The experimental data obtained in the laboratory are shown in Table 7 and Table 8. The Table 7 is referred to scheme in Figure 5 and Table 8 is referred to scheme in Figure 6. The unique difference is the presence of IM as additional load presented in Figure 6 scheme. Then, the Table 7 and Table 8 shows the data obtained from the scheme shown respectively in Figure 5 and Figure 6 which are similar to scheme shown in [1], except for the presence of an autotransformer in the DCM_{IG} field circuit for Figure 5 and Figure 6 and an IM as additional load in Figure 6.

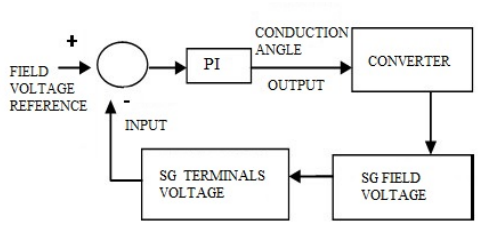
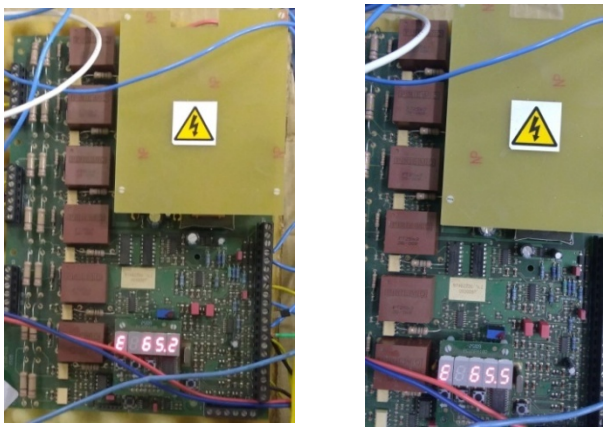


Figure 8. Field voltage control loop



(a) Speed Control Loop (b) Voltage Control Loop

Figure 9. Circuit boards MP410T used to implement the control loops

The autotransformer is responsible for applying a voltage range directly on the DCM_{IG} armature to obtain a larger speed range of the DCM_{IG} and to push more power to the IG, keeping the DCM_{IG} parameters into the rated values. Hence, the elevation of IG speed and power depend on just V_{aDCMIG}, it means this method is different from the scheme presented in [1] that depends on the DCM_{IG} field flux. Both V_{aDCMIG} and the DCM_{IG} field flux are adjusted manually. In summary, the [1] shows the data obtained from the scheme that has the 234Vdc steady source, Table 7 and Table 8 shows the data obtained from the schemes shown in Figure 5 and Figure 6, that has an autotransformer and a diodes bridge as substitute of the 234 Vdc steady source.

The data presented in Table 8 related to Figure 6 system performance results are compared to data obtained from Table 7 related to scheme in Figure 5 and also compared to data obtained from [1]. The scheme in Figure 5 and Table 7 shows the performance of the isolated electric system that considers the voltage application range on the DCM_{IG} armature by an autotransformer and an diodes bridge. The scheme in Figure 6 and Table 8 shows the performance of the isolated electric system that considers the same scheme presented in Figure 5 added by a induction motor, IM, as an additional load.

4. Results

The results and comparative analysis from experiments shown in Figure 5, Figure 6 and the results obtained in [1] are presented in the following.

The results show graphs and analysis of system configurations presented in Figure 5 and Figure 6, data obtained in [1] and analysis of contribution from each generator for each arrange of load, that, in this case, is ether just resistive banks or resistive banks added with an induction motor IM. It is noted that for all graphs obtained from [1] is related to a set system frequency of 55 Hz.

Table 7: Operational Scenarios in Controlled Mode and load being resistors bank

	SG and IG in parallel mode and no load. 40 µF / phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in . 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG alone and load in. 40 µF /phase	SG alone and no load. 30 µF /phase
Results	**Scenario 1B**	**Scenario 2B**	**Scenario 3B**	**Scenario 4B**	**Scenario 5B**	**Scenario 6B**	**Scenario 7B**	**Scenario 8B**	**Scenario 9B**
V_{SG} (V)	220.0	220.0	220.0	220.0	220.0	220.0	220.0	220.0	220.0
f_{SG} (Hz)	60	60	60	60	60	60	60,0	60,0	60,0
$I_{fieldSG}$ (mA)	190	320	300	300	320	350	370	340	100
V_{aDCMSG} (V)	272.9	275.0	276.2	276.8	277.5	278.0	280.7	278.5	270.0
I_{aDCMSG} (A)	2.0	2.5	3.8	4.7	6.2	7.4	9.5	10.6	2.0
I_{fDCMSG} (mA)	510.0	530.0	520.0	520.0	515.0	510.0	515.0	515.0	515.0
V_{IG} (V)	220.0	220.0	220.0	220.0	220.0	220.0	220.0	n.a.	n.a.
f_{IG} (Hz)	60	60	60	60	60	60	60	n.a.	n.a.
n_{IG} (rpm)	1793	1896	1878	1857	1848	1821	1794	n.a.	n.a.
V_{aDCMIG} (V)	291.3	323.6	316.7	308.3	301.3	293.0	284.1	n.a.	n.a.
I_{aDCMIG} (A)	0.5	8.0	7.0	6.0	4.7	3.5	2.0	n.a.	n.a.
I_{fDCMIG} (mA)	330.0	330.0	330.0	330.0	330.0	330.0	330.0	n.a.	n.a.
I_{wloadR} (A)	0.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	0.0
I_{IG} (A)	2.4	5.8	5.1	4.4	3.7	2.9	2.4	0.0	0.0
I_{SG} (A)	2.9	1.5	2.0	2.8	3.6	4.5	5.6	7.4	4.0
I_{C} (A)	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	4.0
I_{WSG} (A)	**0.0**	***1.05***	***1.48***	***1.97***	***2.47***	***3.32***	***4.65***	***5.03***	**0.0**
I_{C1} (A)	***2.94***	***1.13***	***1.60***	***2.13***	***2.66***	***3.03***	***3.13***	***5.40***	**4.0**
I_{C2} (A)	***2.45***	***4.26***	***3.80***	***3.27***	***2.74***	***2.37***	***2.27***	***0.0***	**0.0**
I_{WIG} (A)	***0.0***	***3.95***	***3.51***	***3.03***	***2.53***	***1.68***	***0.35***	***0.0***	**0.0**
PSG (W)	***0.0***	***400.5***	***565.5***	***750.0***	***939.8***	***1266.6***	***1770.1***	***1915.8***	**0.0**
PIG (W)	***0.0***	***1504.8***	***1335.8***	***1155.3***	***965.5***	***638.67***	***135.17***	***0.0***	**0.0**
P_{DCMSG} (W)	***545.8***	***687.5***	***1049.6***	***1301.0***	***1665.0***	***2057.2***	***2666.7***	***2962.7***	**543.4**
P_{DCMIG} (W)	***145.7***	***2588.8***	***2216.9***	***1849.8***	***1416.1***	***1025.5***	***568.2***	***0***	**0.0**
η_{GROUP} (%)	***0.0***	***58.15***	***58.33***	***60.47***	***61.84***	***61.80***	***58.89***	***64.66***	**0.0**
η_{SG} (%)	***0.0***	***58.25***	***53.88***	***57.65***	***56.44***	***61.57***	***66.38***	***64.66***	**0.0**

η_{IG} (%)	***0.0***	***58.12***	***60.43***	***62.45***	***68.18***	***62.28***	***23.79***	***n.a.***	**n.a.**

Obs.: 1- Italic and bold text: calculated values by equations 23 to 33
2- Frequency was chosen based on the average between the SG and IG rated frequency
3- n.a: not applicable

Table 8: Operational Scenarios in Controlled Mode and load being resistors bank and induction motor

	SG and IG in parallel mode and no load. 40 µF / phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG and IG in parallel mode and load in. 40 µF /phase	SG alone and load in. 40 µF /phase	SG alone and no load. 30 µF /phase
Results	**Scenario 1C**	**Scenario 2C**	**Scenario 3C**	**Scenario 4C**	**Scenario 5C**	**Scenario 6C**	**Scenario 7C**	**Scenario 8C**	**Scenario 9C**
V_{SG} (V)	220.0	220.0	220.0	220.0	220.0	220.0	220.0	220	220.0
f_{SG} (Hz)	60	60	60	60	60	60	60	60	60
$I_{fieldSG}$ (mA)	170	420	410	400	400	430	475	390	100
V_{aDCMSG} (V)	278.9	277.0	277.5	280.5	281.0	281.1	282.9	279.2	271.7
I_{aDCMSG} (A)	2.0	2.9	4.0	5.0	6,3	7,6	9,6	10,5	2,0
I_{fDCMSG} (mA)	530.0	530.0	530.0	530.0	530.0	520.0	520.0	520.0	515.0
V_{IG} (V)	220.0	220.0	220.0	220.0	220.0	220.0	220.0	0.0	n.a.
f_{IG} (Hz)	60	60	60	60	60	60	60	0.0	n.a.
n_{IG} (rpm)	1803	1893	1868	1857	1843	1821	1793	0.0	n.a.
V_{aDCMIG} (V)	288.9	330.0	318.0	310.0	302.3	295.2	285.9	0.0	n.a.
I_{aDCMIG} (A)	1.0	8.0	7.0	6.0	4.7	3.5	2.0	0.0	n.a.
I_{fDCMIG} (mA)	330.0	330.0	330.0	330.0	330.0	330.0	330.0	0.0	n.a.
Iw_{loadR} (A)	0.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	0.0
I_{wMT} (A)	0.0	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.0
I_{dwMT} (A)	0.0	1.23	1.23	1.23	1.23	1.23	1.23	1.23	0.0
I_{IG} (A)	2,3	6,0	5,0	4,4	3,6	2,9	2,3	0.0	0.0
I_{SG} (A)	2.8	0.7	1.6	2.3	3.3	4.1	5.3	6.7	4.0
I_C (A)	5.4	5.4	5.4	5.4	5.4	5.4	5.4	5.4	4.0
I_{WSG} (A)	***0.0***	***0.65***	***1.29***	***1.99***	***3.04***	***3.74***	***4.95***	**5.23**	**0.0**
I_{C1} (A)	***2.94***	***0.28***	***1.03***	***1.17***	***1.29***	***1.66***	***1.88***	**4.2**	**4.0**
I_{C2} (A)	***2.46***	***3.91***	***3.14***	***3.00***	***2.88***	***2.51***	***2.29***	**0.0**	**0.0**
I_{WIG} (A)	***0.0***	***4.55***	***3.91***	***3.22***	***2.16***	***1.45***	***0.25***	**0.0**	**0.0**
PSG (W)	***0.0***	***247.97***	***490.34***	***753.80***	***1156.9***	***1428.7***	***1887.8***	**1991.7**	**0.0**
PIG (W)	***0.0***	***1733.5***	***1491.1***	***1227.7***	***824.52***	***552.74***	***93.68***	**0.0**	**0.0**
P_{DCMSG} (W)	***557.8***	***803.30***	***1110.0***	***1402.5***	***1770.3***	***2136.4***	***2715.8***	**2931.6**	**624.5**
P_{DCMIG} (W)	***288.9***	***2640.0***	***2226.0***	***1860.0***	***1420.8***	***1033.2***	***571.8***	**0.0**	**0.0**
η_{GROUP} (%)	***0.0***	***57.55***	***59.40***	***60.73***	***62.09***	***62.52***	***60.27***	**67.59**	**0.0**
η_{SG} (%)	***0.0***	***30.87***	***44.17***	***53.75***	***65.35***	***66.88***	***69.51***	**67.59**	**0.0**
η_{IG} (%)	***0.0***	***65.66***	***66.99***	***66.00***	***58.03***	***53.50***	***16.38***	**0.0**	**n.a.**

Obs.: 1- Italic and bold text: calculated values by equations 23 to 33
2- Frequency was chosen based on the average between the SG and IG rated frequency
3- n.a: not applicable

For all experiments shown in this paper, the system frequency is 60 Hz. Then, the synchronous speed for the experiments in [1] is 1650 rpm (6) and 1800 rpm for the experiments shown in this paper.

To highlight the contributions from autotransformer and diodes bridge in the experiments in relation to the experiments without these devices, it will be presented some analysis based on graphs and experiment results.

Figure 10 shows that the PIG was limited to 1164 W [1] in scenario 2 [1] because the over current of I_{aDCMIG} 9.8 A [1] is a value greater than the rated I_{aDCMIG} of 7.72 A as Table 1. To overcome this barrier, an autotransformer and a diodes bridge were installed to widen the voltage range applied over the DCM armature, as shown in Figure 5 and Figure 6. The armature voltage range with the autotransformer and diodes bridge can vary from 0 to 336 Vdc. This range is greater than the previous V_{aDCMIG} of 234 Vdc. As the DCM_{IG} speed, n_{IG}, is directly proportional to V_{aDCMIG}, as shown in (2), the n_{IG} is elevated proportionally with the V_{aDCMIG} as shown in scenarios 2 to 8, in Table 7 and Table 8. Then, in this latter method, I_{aDCMIG} variation depends on the load conjugate variation only; as long as I_{aDCMIG} decreases, I_{aDCMSG} rises because of the load distribution to the generators, whereas the DCM_{IG} flux, ϕ, remains constant as (3).

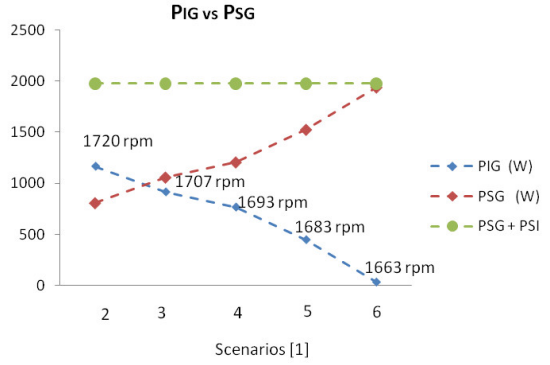


Figure 10. PIG vs PSG with DCM_{IG} field flux variation and V_{aDCM} kept constant

Figure 11 and Figure 12 shows that the power range supplied from IG, PIG, through the autotransformer method, is greater than PIG of 234 Vdc through the steady source method. Consequently, the power complement from SG is

lower than the PSG from 234 Vdc steady source method. Both methods feed the entire resistive load in Figure 10 and Figure 11, and resistive and motor load as shown in Figure 12.

In Figure 13, Figure 14 and Figure 15 shows the comparative results of IG and SG primary machine output power, P_{DCMIG} vs P_{DCMSG}. These performances are similar to PIG and PSG performances shown in Figure 10, Figure 11 and Figure 12. The P_{DCMIG} in Figure 13 elevates because the flux decreases and V_{aDCMIG} is kept constant [1] The P_{DCMIG} in Figure 14 and Figure 15 elevates because of the V_{aDCMIG} increases, which is manually adjusted by the autotransformer shown in Figure 5 and Figure 6.

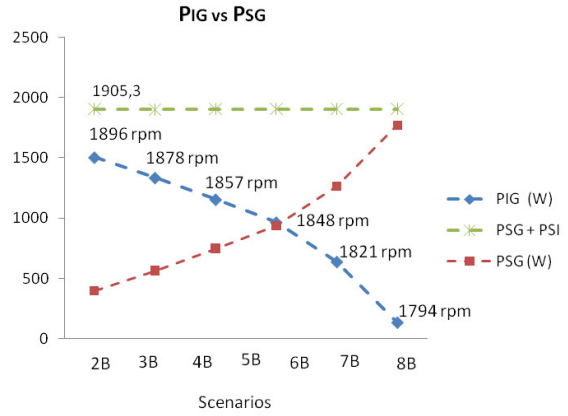


Figure 11. PIG vs PSG with V_{aDCMIG} variation and field flux kept constant (Only resistive load)

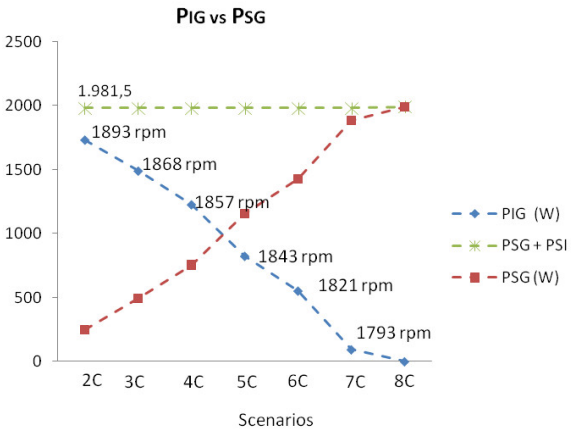


Figure 12. PIG vs PSG with V_{aDCMIG} variation and field flux kept constant (Resistive load plus IM)

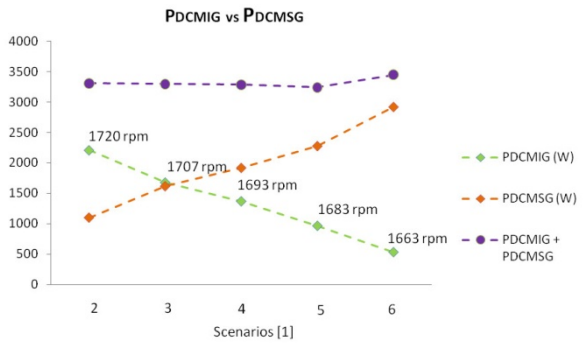


Figure 13. P_{DCMIG} vs P_{DCMSG} with DCM_{IG} field flux variation and V_{aDCMIG} kept constant

The efficiency results shown in Figure 16, Figure 17 and Figure 18 present each subgroup efficiency for each scenario and load conditions. Each subgroup efficiency is resulted from relation between a generator and a DCM as shown in (32) and (33). In summary, the target scenarios are: 2 to 7 (without autotransformer) as in [1], scenarios 2B to 8B (with autotransformer and just resistive load) and scenarios 2C to 8C (with autotransformer, resistive load and induction motor).

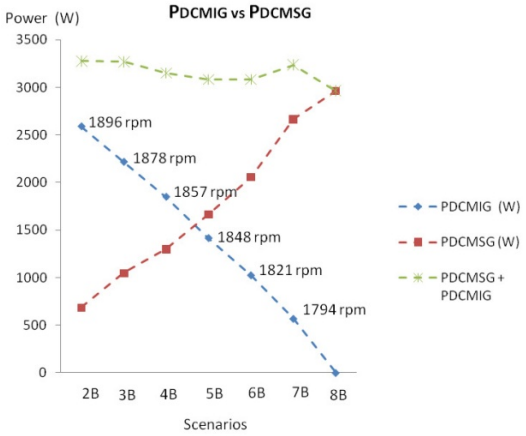


Figure 14. P_{DCMIG} vs P_{DCMSG} with V_{aDCMIG} variation and field flux kept constant (Only resistive load)

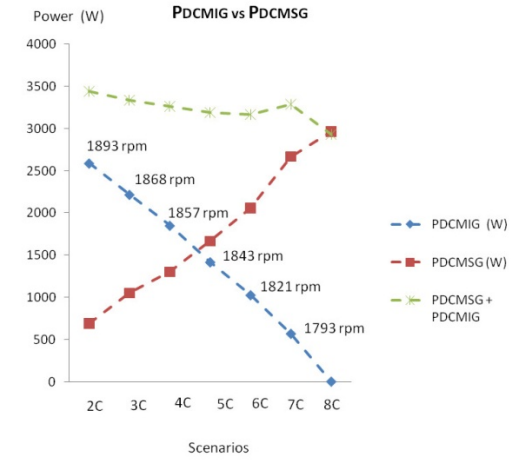


Figure 15. P_{DCMIG} vs P_{DCMSG} with V_{aDCMIG} variation and field flux kept constant (Resistive load plus IM)

The IG subgroup efficiencies shown in Figure 17 and Figure 18 are bigger than IG subgroup efficiencies from similar scenarios in Figure 16. Moreover, the elevation of power from IG due to the increase of DCM_{IG} speed, n_{IG}, is more representative of actual offshore conditions. The turbines can assume whatever speed required from generators in offshore platforms.

The reduction of losses contributed to increased IG subgroup efficiency as shown in Figure 17 and Figure 18. Fig. 16 shows the machine's efficiencies when the I_{fDCMIG} decreases to elevate the speed. The efficiency improves when the V_{aDCMIG} elevation method is applied as shown in Figure 17 and Figure 18.

The Figure 17 and Figure 18 are higher than the efficiency presented in Figure 16 considering that the Figure 16 indicates a DCM_{IG} speed range is more restricted as already clarified in [1].

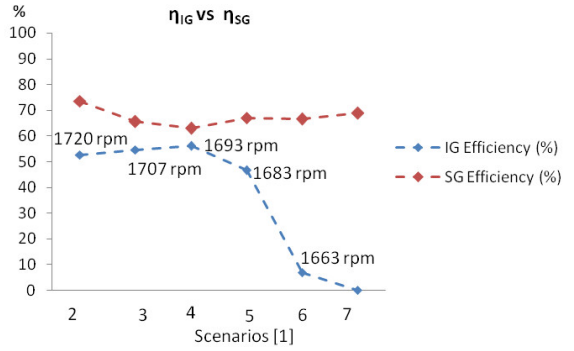


Figure 16. η_{IG} vs η_{SG} with DCM_{IG} field flux variation and V_{aDCMIG} kept constant

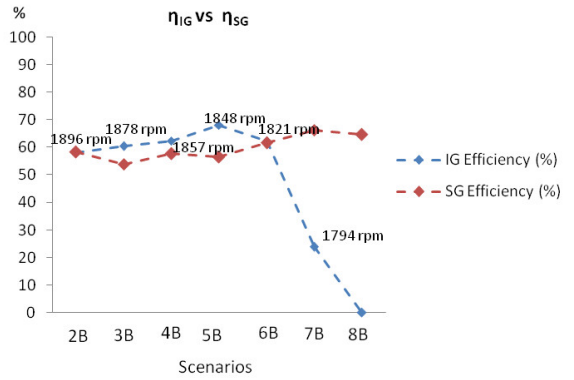


Figure 17. η_{IG} vs η_{SG} with V_{aDCMIG} variation and field flux kept constant (Only resistive load)

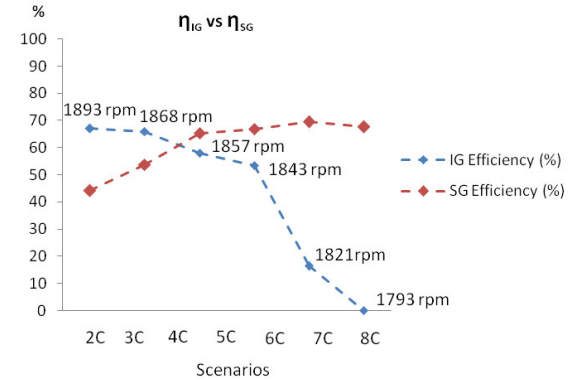


Figure 18. η_{IG} vs η_{SG} with V_{aDCMIG} variation and field flux kept constant (Resistive load plus IM)

5. Conclusion

The cited electric system shows voltage and frequency regulation for each scenario transition as demonstrated by results and respective analysis.

As it happened in [1], it was realized the master and slave behaviors to respectively SG and IG. SG controls the system voltage and frequency and IG follows the voltage and frequency defined by SG. The IG establishes the active power supplied to the system and the SG complements the rest of active power and reactive power required. The PIG depends on the IG shaft speed imposed by DCM_{IG}. The PSG and the system voltage and frequency are controlled by voltage and speed control loops presented in Figure 5 and Figure 6.

It was realized a greater PIG range in the scenarios in Table 7 and in Table 8, than the scenarios from [1] due to V_{aDCMIG} elevation method. The PSG complemented the PIG manually set to attend the load.

The V_{aDCMIG} elevation method resulted in increase of P_{DCMIG} and PIG and higher IG subgroup efficiency than the I_{fDCMIG} reduction method[1]. The increase of P_{DCMIG} and PIG were results of IG speed increase, $n_{IG,}$ that was obtained by elevation of V_{aDCMIG} as shown in Figure 11, Figure 12 and Figure 14. The use of V_{aDCMIG} range instead of Vdc steady source enables to reach the I_{aDCMIG} limit of 8,0 A and 1893 rpm shaft speed, it is 5,16% higher than synchronous speed. It is higher than Vdc steady source method limit that reached just 4,24% higher than synchronous speed, Figure 16 and [1].

The efficiencies from the V_{aDCMIG} elevation method were greater than efficiencies from the scheme with Vdc steady source [1], as shown in Figure 17 and Figure 18. V_{aDCMIG} elevation method does not have the additional current losses of Vdc steady source method. When the field flux ϕ, I_{fDCMIG}, is decreased, I_{aDCMIG} is increased by (3) and IG speed, $n_{IG,}$ is increased by (2). The main losses for the Vdc steady source method are related to higher I_{aDCMIG} elevation, $P(w)=Ra* I_{aDCMIG}^2$. Then, by the Vdc steady source method, besides I_{aDCMIG} be elevated due to the rise of PIG, I_{aDCMIG} is also elevated due to reduction of I_{fDCMIG}. Higher I_{aDCMIG} result in higher losses and lower efficiencies.

As shown in power performance graphs related to the V_{aDCMIG} elevation method, Figure 11, Figure 12, Figure 14 and Figure 15, the PIG and P_{DCMIG} are more representative of actual operation conditions in offshore platforms.

In continuation of this content, a study of load and generation transients has been done and submitted to publication in COBEP, Brazilian Power Electronic Conference that took place in November, 2017.

In the next stage of this work, it will be studied energy regeneration and ballast load as a additional way to the system regulation.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank the UNIFEI and PETROBRAS for the general support along the study.

References

[1] V. Z. Silva, A. J. J. Rezek, and R. L. Corrêa, "Analysis of synchronous and induction generators in parallel operation mode in an isolated electric system", Brazil, 2017, https://doi.org/10.1109/PEDG.2017.7972459

[2] S.P. Singh., B. Singh, M.P. Jain, "Comparative study on the performance of a commercially designed induction generator with induction motors operating as self-excited induction generators", IEE Proc. C, vol. 140, no.5, pp: 374 - 380, 1993, https://doi.org/ 10.1049/ip-c.1993.0055

[3] M. H. Haque, "Self-excited single-phase and three-phase induction generators in remote areas", IEEE conference, pages 38-42, 2008, https://doi.org/ 10.1109/ICECE.2008.4769169

[4] E. Touti, R. Pusca, J. F. Brudny, A. Châari, "Asynchronous generator model for autonomous operating mode", IEEE conference, pages 1-3, 2013, https://doi.org/ 10.1109/ECAI.2013.6636208

[5] D. M. Macedo, "The use of pumps operating as turbines and induction generators in electric power generation" (in Portuguese), Master in Science Dissertation, Federal University of Itajubá(UNIFEI), Itajubá, 2004.

[6] J. M. Chapallaz, P. Eichenberger, and G. Fischer, *Manual on Pumps used as Turbines.* MHPG Series, v. 11, Friedr. Vieweg&SohnVerlagsgesellschaftmb, Germany, 1992.

[7] J. R. Cogo, J.C. Oliveira, and J. P. G. Abreu, "Tests in induction machines", editor EFEI, Itajubá-MG, Brazil, 1983.

[8] A. P. Rizzi, " Electrical Measurements: power, energy, power factor, demand", in Portuguese, Livros Técnicos e Científicos Editora S.A., Rio de Janeiro-RJ, 1980.

Appendix 1_ Induction Motor Parameters

The follow calculus were developed on base in [7].

1. No Load and Locked Rotor Tests

The Figure 1 shows the circuit implemented in laboratory to no load test and locked rotor test.

The temperature during the test= 26ºC.

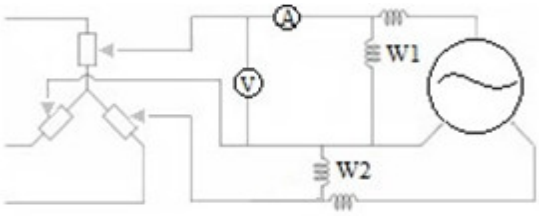
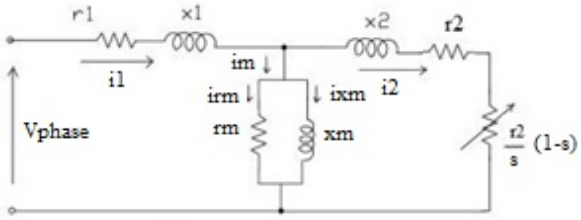


Figure 1: No load test and locked rotor test circuit

Figure 2: Induction Motor Equivalent Circuit



The Figure 2 shows the induction motor IM equivalent circuit.

5.1. No Load Test

In this section, it will be presented the no load test and locked rotor test to calculus the IM parameters such as impedances, currents, losses and efficiency

The Table 1 presented the no load test measurements.

Table 1: No load Test

Vapply (V)	W1 (W)	W2 (W)	V (V)	A (A)	W1+W2(W)
240	-120.0	+220.0	240.2	1.50	100.0
220	-100.0	+170.0	220.8	1.35	70.0
200	-70.0	+130.0	200.3	1.15	60.0
180	-60.0	+110.0	180.3	1.00	50.0
160	-40.0	+80.0	160.0	0.90	40.0
140	-30.0	+60.0	140.0	0.75	30.0
120	-20.0	+45.0	120.0	0.65	25.0
100	-15.0	+30.0	100.1	0.55	15.0
80	-10.0	+20.0	80.0	0.50	10.0
60	-4.0	+12.5	60.0	0.25	8.5
40	0.0	+5.0	40.0	0.15	5.0
20	0.0	+2.0	20.1	0.05	2.0
0	0.0	0.0	1.99	0.0	0.0

Figure 3 shows the curve of voltage applied and power measured by wattmeter (w1) and (w2). Note that the cross of curve and vertical axe result in estimated attrition and ventilation losses (*Pav*). In this case *Pav* is 1 W as demonstrated.

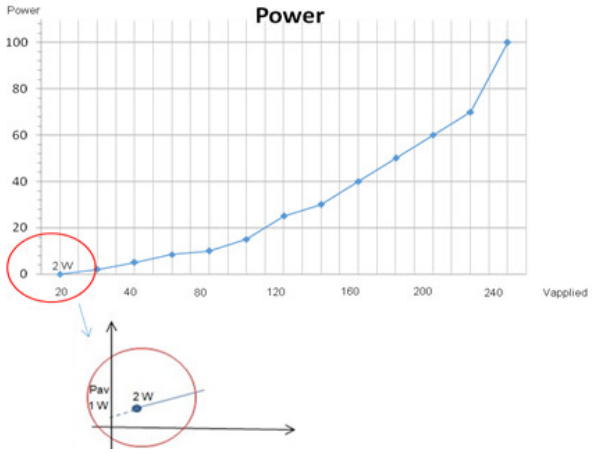


Figure 3: Results of no load test: W1+W2 vs Voltage applied

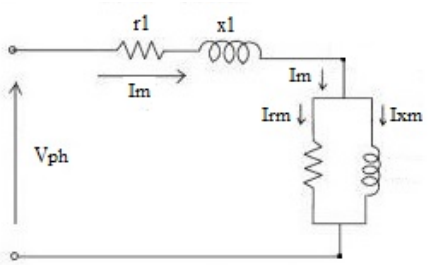


Figure 4: Equivalent circuit - No load test

- Magnetization Current

Im = Magnetization Current

Im= 1.35 A as shown in Table 1.

5.2. Locked Rotor Test

The Table 2 shows the locked rotor test measurements.

Table 2
Locked Rotor Test

W1 (W)	W2 (W)	V (V)	In (A)	W1+W2 (W)
20	65	39.34	1.80	85

- Rated Current

In = Rated Current

$$in = {}^{P}/_{\sqrt{3} * Vpp * cos\emptyset * \eta} \quad (1)$$

$$in = {}^{370}/_{\sqrt{3} * 220 * 0.71 * 0.76}$$

$$in = 1.80\ A \quad (2)$$

The Figure 3 shows the induction motor IM equivalent circuit during locked rotor test

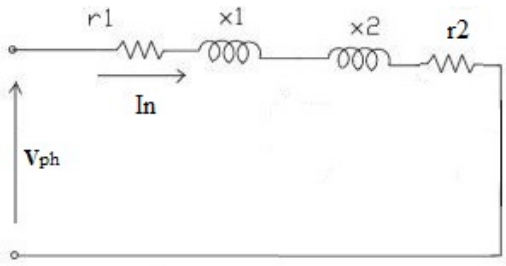


Figure 5: Equivalent circuit _ Locked rotor test

5.3. Magnetization Branch Parameters

Rm= Motor Average Measured Resistance
ry= Phase resistance in star connection
rd= Phase resistance in delta connection

Table 3:
Average Resistance

$R_{AA'}$	$R_{BB'}$	$R_{CC'}$
10.1	10.3	10.16

$$R_m = {(10.1 + 10.3 + 10.16)}/{3} = 10.16\ \Omega \quad (3)$$

- r1 Calculus

Follow the calculus memory of stator winding resistor r1.

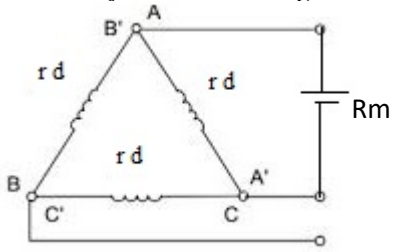


Figure 6: IM winding connection

$$R_m = {rd * 2rd}/{3rd} \quad (4)$$

$$R_m = {2}/{3}\,rd \quad (5)$$

$$rd = {3}/{2}\,rm \quad (6)$$

$$ry = {1}/{3}\,rd \quad (7)$$

(5) in (6)

$$ry = {Rm}/{2} = {10.16}/{2}$$

$$r1 = 5.08\ \Omega \quad (8)$$

- r2 Calculus

From Table XI, $W1 + W2 = 85\ W$

$$W1 + W2 = 3 * (\mathrm{r1} + \mathrm{r2}) * \mathrm{in}^2 \quad (9)$$

$${85}/{(3 * 1.8^2)} = 5.08 + r2$$

$$r2 = 3.66\ \Omega \quad (10)$$

- Reactive Power

$$Q = \sqrt{3} * Vpp * IdwMT * sen\emptyset = \sqrt{3} * (W2 - W1) \quad (11)$$

From Table XI, W2-W1=45

$$Q = \sqrt{3} * (65 - 20) = 77.94\ Var \quad (12)$$

$$Q = 3 * (x1 + x2) * In^2 \quad (13)$$

(13)and (2) in (14)

$$77{,}94 = 3 * (x1 + x2) * 1.8^2 \quad (14)$$

$$(x1 + x2) = 8.01\ \Omega \quad (15)$$

$$\begin{cases} x1 \rightarrow x2 \\ r1 \rightarrow r2 \end{cases} \quad [8] \quad (16)$$

(8) and (10) in (16)

$$\begin{cases} x1 \rightarrow x2 \\ 5{,}08 \rightarrow 3.66 \end{cases}$$

$$x1 = {r1 * x2}/{r2} \quad (17)$$

$$x1 = {5{,}08 * x2}/{3.66} \quad (18)$$

From (15) and (18), result in

$$5{,}08 * x2 + x2 * 3.66 = 29.32 \quad (19)$$

$$x2 * 8{,}74 = 29.32 \quad (20)$$

$$x2 = 3.35\ \Omega \quad (21)$$

(21) in (18), result in

$$x1 = {5{,}08 * 3.35}/{3.66} \quad (22)$$

$$x1 = 4.65\ \Omega \quad (23)$$

2. Power and Losses Calculus

The Pav was defined in Figure 3 and consist in 1 W and Pnl (no load losses) is no load test for 220 V, it resulted in 70 W (W1 +W2).

- Stator Joules Losses

$$\text{Pjstator26º= Stator Joule Losses} = 3 * r1 * im^2 \quad (24)$$

$$\text{Pjstator26º} = 3 * 5.08 * 1.35^2 \quad (25)$$

$$\text{Pjstator26º} = 27.77\ \text{W} \quad (26)$$

- Phf_ histerese and foucault losses

$$P_{hf} = (W1 + W2) - 3 * r1 * im^2 - Pav \quad (27)$$

$$P_{hf} = 70 - 3 * 5.08 * 1.35^2 - 1 \quad (28)$$

$$P_{hf} = 41.23\ W \quad (29)$$

5.4. No Load Reactive Power

$$Q_o = \sqrt{3} * (W2 - W1) \quad (30)$$

From Table 1, $W2 - W1 = 270W$

$$Q_o = \sqrt{3} * 270\ \text{W}$$

$$Q_o = 467.65\ \text{Var} \quad (31)$$

5.5. Magnetization Branch

- No load Reactive Power in Magnetization Branch

$$Q_m = Qo - 3 * X1 * io^{\,2} \quad (32)$$

$$Q_m = 467.65 - 3 * 4.66 * 1.35^{\,2} \quad (33)$$

$$Q_m = 442.23\ Var \quad (34)$$

- Magnetization Branch Power Factor Angle

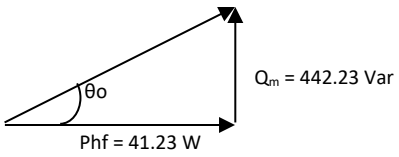


Figure 7: Magnetization branch power factor

$$\theta o = arctg\ {}^{442.23}/_{41.23} \quad (35)$$

$$\theta o = 84.67^{o} \quad (36)$$

- Magnetization Branch Currents

$$\text{Irm} = \text{Im} * \cos\theta o \quad (37)$$

$$\text{Irm} = 1.35 * \cos 84.67^{o} \quad (38)$$

$$\text{Irm} = 0.1254\ A \quad (39)$$

$$\text{Ixm} = \text{Im} * \text{sen}\theta o \quad (42)$$

$$\text{Ixm} = 1.35 * \text{sen}84.67^{o} \quad (40)$$

$$\text{Ixm} = 1.34\ \text{A} \quad (41)$$

- Magnetization Branch Resistance

$$\text{rm} = {}^{\text{Phf}}/_{(3 * \text{Irm}^2)} \quad (42)$$

$$\text{rm} = {}^{41.23}/_{(3 * 0.1254^2)} \quad (43)$$

$$\text{rm} = 873.97\ \Omega \quad (44)$$

- Magnetization Branch Reactance

$$\text{xm} = {}^{\text{Qm}}/_{(3 * \text{Ixm}^2)} \quad (45)$$

$$\text{xm} = {}^{442.23}/_{(3 * 1.34^2)} \quad (46)$$

$$\text{xm} = 82.10\ \Omega \quad (47)$$

3. Parameters of Induction Motor

Figure 8 shows all IM parameters calculated as resistors, reactance and currents as data obtained during experiment, without temperature correction because the motor operated no load.

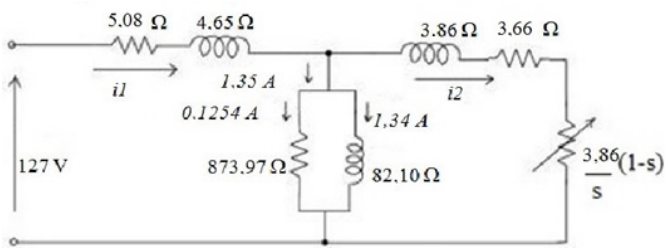


Figure 8: Parameters of Induction Motor _IM

4. Induction Motor Efficiency

It will be calculated the temperature correction factor, corrected resistences to 40ºC, corrected winding losses, Pwind40ºC, and total losses, Ptl, before induction motor efficiency calculus.

5.6. Temperature Correction Factor and Resistence Correction

- Temperature Correction Factor Kt 40ºC

$$\text{Kt } 40^{o} = {}^{234.5 + 40}/_{(234.5 + 26)} \quad (48)$$

$$\text{Kt } 40^{o} = 1.054 \text{ to r1 and r2} \quad (49)$$

- Resistence Correction

$$\text{r1 } 40^{o} = \text{r1 } 26^{o} * 1.054 \quad (50)$$

$$r1\ 40^{o} = 5.08 * 1.054 \quad (51)$$

$$r1\ 40^{o} = 5.35\ \Omega \quad (52)$$

$$r2\ 40^{o} = r2'26^{o} * 1.054 \quad (53)$$

$$r2\ 40^{o} = 3.66 * 1.054 \quad (54)$$

$$r2\ 40^{o} = 3.86\ \Omega \quad (55)$$

5.7. Corrected Winding Losses and Total Losses

- Winding Losses

The winding losses, Pwind, measured in locked rotor test, Table 2, need to be corrected to 40ºC. Follow the corrected winding losses.

$$Pwind40^{o} = kt\ 40^{o} * Pwind26^{o}\ (\text{locked rotor test}) \quad (56)$$

$$Pwind40^{o} = 1.054 * 85W \quad (57)$$

$$Pwind40^{o} = 89.59\ W \quad (58)$$

- Total Losses

The total losses, Ptl, calculus:

$$Ptl = Phf + Pav + Pwind40^{o} \quad (59)$$

$$Ptl = 41.23W + 1W + 89.59W \quad (60)$$

$$Ptl = 131.82\ W \quad (61)$$

5.8. Rated Motor Efficiency Estimate

Pout = Motor output power (Motor dataplate_Table 6)
Pin=Motor Input power

$$\eta = {Pout}/{Pin} \quad (62)$$

$$Pin = Pout + losses \quad (63)$$

$$Pin = 370\ W + 131.82\ W \quad (64)$$

$$Pin = 501.82\ W \quad (65)$$

$$\eta = {370}/{501.82} \quad (66)$$

$$\eta = 73.73\ \% \quad (67)$$

5. Induction Motor Currents

Follow the calculus of IM reactive and active currents (IdwMT and IwMT). This values are shown in Table 8 for all scenarios tested.

- Reactive Current

$$Q = \sqrt{3} * 220 * IdwMT = \sqrt{3} * (W2 - W1) \quad (68)$$

$$\sqrt{3} * 220 * IdwMT = 467.65 \quad (69)$$

$$IdwMT = 1.23\ A\ ** \quad (70)$$

** The difference between IdwMT and Ixm is perfectly acceptable. The difference close to 0,1 A between expected and found values is due to small measurements errors and instruments precision and accuracy errors. It does not impact the presented modeling. Therefore, the results and modeling continue valid.

- Active Current

$$P = (W1 + W2) = 70.0\ W \quad (71)$$

$$Q = \sqrt{3}(W2 - W1) = 467.6\ W \quad (72)$$

$$S = \sqrt{70.0^2 + 467.6^2} = 472.86\ VA \quad (73)$$

$$\cos\emptyset = {P}/{S} = {70}/{472.86} \quad (74)$$

$$\cos\emptyset = 0.148 \quad (75)$$

$$IwMT = Io * \cos\emptyset = 1.35 * 0.148 \quad (76)$$

$$IwMT = 0.20\ A \quad (77)$$

In Figure 9 are shown resistor banks current and no load IM current.

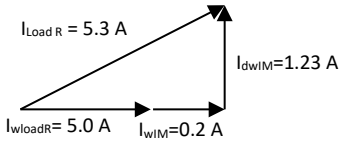


Figure 9: Resistors bank and IM currents in scenarios C

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 217-224 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Measurement of the Bridge Surface Deflections Using Near-Field Amplitude of Secondary Radiators System

Oleksandr Poliarus[*,1], Yevhen Poliakov[1], Andrii Lebedynskyi[1], Viktor Ivanov[2], Ruslan Pashchenko[2]

[1]*Department of metrology and life safety, Kharkiv National Automobile and Highway University, 61002, Ukraine*

[2]*O. Ya. Usikov Institute for Radiophysics and Electronics of the National Academy of Sciences of Ukraine, 61085, Ukraine*

ARTICLE INFO

Article history:
Received: 01 November, 2017
Accepted: 17 November, 2017
Online: 15 December 2017

Keywords :
Remote measurement
Radiating element
Antenna directivity pattern
Field amplitude measurement
Surface deflection

ABSTRACT

In remote measurements on large-sized objects main focus is on measuring deformations of bridge surfaces and determination of the dynamic amplification factor in particular. The proposed method estimates deflections of entire lower surface of a bridge due to placement of secondary radiators on this surface. Antenna array directivity pattern (ADP) is distorted during deflection of the bridge surface. ADP distortions contain the information about deflections of the individual points of the surface. Minimization of the distance in a functional space with a quadratic metric (the functional) between a distorted measured ADP and theoretical one allows us to determine these deflections. The main problem is measurement of the secondary radiators system ADP in far zone which often cannot be achieved for real bridges in practical situations. Therefore, a new method for determining surface deflections based only on field amplitude measurements in near-field zone of a secondary radiators system is considered in the article. This amplitude of field strength depends on the bridge surface deflection. The modulus of difference between the normalized amplitude distribution of the measured field at the outputs of the receiving array elements, which is created by the radiators system of an unloaded bridge, and the similar amplitude field distribution of a loaded bridge is minimized and deflection of the bridge at the points of the radiating elements placement is calculated.

1. Introduction

This paper is an extension of the work originally presented in the International Conference on Antenna Theory and Techniques [1].

In engineering practice, an important place is taken by measurements of surface deflections of large-sized objects. The dynamics of these deflections can serve as a basis for diagnosing the objects' state. Deflection measurements for lower part of the surface are one of the most important actions in maintenance of bridges. For this purpose, contact sensors are usually mounted directly on the lower part of the bridge. The installation of contact sensors and transducers interrupts traffic on the bridges and can be time-consuming. Non-contact measurements on bridges which are carried out remotely have recently become an alternative to the contact ones. The analysis of remote method for estimating bridge surface deflections was carried out in [1]. It is based on measurement of the antenna directivity pattern (ADP) distortion of the radiators which are installed on the lower surface of a bridge.

2. Modern Methods of Bridges Surface Deflections Measurement

Let us consider the methods of deflection measurements on bridge structures. The number of deflection measurement methods has increased essentially in the last decade. Bridge safety monitoring is based, in a great extent, on measurements of deflections that become indispensable for this aim. New technologies include intelligent robots that feature a high speed of measurement and precision, laser image deflection measurement, GPS satellite-surveying systems, electro-optical imaging deflection, wireless inclination sensor systems [2, 3] and techniques of digital image correlation between images of undeformed and deformed configurations [4]. In [5], for measuring deflection and vibration of bridges a ground-based interferometric radar is suggested. The maximum measurement range is equal to 1000 m and the errors of displacement radial component can reach

[*]Oleksandr Poliarus, Email : poliarus.kharkov@ukr.net

0.1 mm. Most of the proposed methods are expensive or very expensive.

The main characteristics of any bridge are dynamic amplification factor (DAF), bridge natural frequency and bridge acceleration [6]. The dynamic amplification factor can be evaluated as the maximum total load effect divided by the maximum static effect [7]

$$DAF = \frac{\varepsilon_{dyn}}{\varepsilon_{stat}}, \quad (1)$$

where ε_{dyn} – maximum dynamic response which is usually determined as vertical deflection of the bridge surface due to dynamic loads (for example, moving vehicles), ε_{stat} – maximum deflection under static loads (for example, immovable vehicles).

In [8], an optimal scheme for measuring DAF of a bridge using radar is proposed. This radar measures the phase of reflected electromagnetic wave in noisy environments. In its turn, the phase depends on deflections of the bridge surface.

Increased measurement accuracy can be provided by using measuring devices which operate on the different physical principles. An optimal scheme for determining bridge surface deflection by measuring it simultaneously with a mechanical sensor and radar has been developed in [9]. Practically, operation quality of such scheme does not depend on noise or interference. Since radar-based methods use phase measurements that are sensitive to moving both bridge and earth surfaces, seismic fluctuations of the earth surface reduce the accuracy of the deflections measurement due to traffic, for example. In addition, DAF estimation (1) involves measuring static and dynamic deflections that are separated in time due to technological reasons. This fact imposes strict requirements to stability of radar frequency that increases the cost of a measuring system.

Consequently, all modern measuring technologies for bridge structures can be divided into three groups. The first group includes the use of radar systems of millimeter and optical ranges of electromagnetic waves [10-15] for estimation of deformations and deflections of bridge elements, including stay-cables which support the buildings. For this purpose, in particular, phase interferometers with a high-resolution range are used. These systems measure dynamic characteristics of bridge structures at local points. The second group includes the use of opto-mechanical systems of technical vision [16-18] followed by image processing based on developed algorithms, which allows to analyze the dynamics of bridge structures in distributed areas. These methods do not determine the DAF. The third group is based on application of the algorithms for processing GPS data with Kalman filtration to obtain the value of the bridge DAF [19].

DAF measurement is performed at one point of the bridge structure. To measure it at another point, it is necessary to change the direction of the main maximum of the transmitting and receiving antennas directivity pattern. The existing measuring methods are not able to estimate the dynamics of deflection changes in the entire lower surface of the bridge. In order to fix that, a system for deflection measurements on the lower surface of a bridge by creating, on the surface, an equidistant linear antenna array that re-radiates the incoming electromagnetic wave was proposed in [1] (Figure 1). When the surface is bent due to the process of bridge loading, the radiators also move together with it and the ADP is distorted.

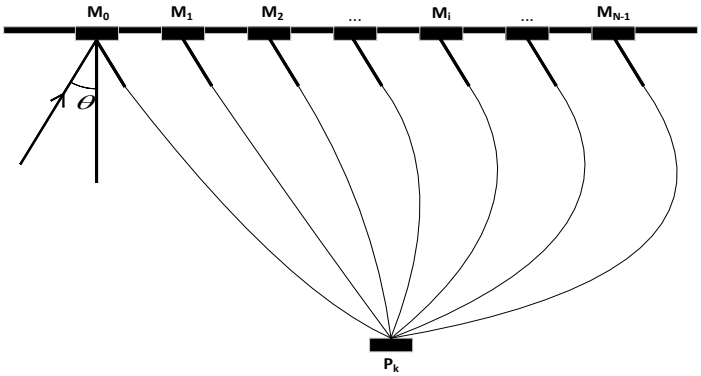


Figure 1: The system of secondary radiators on the lower part of a bridge surface

The antenna directivity pattern is measured in far zone of the antenna and then it is compared with the undisturbed one which was previously measured in the absence of deflections on the unloaded bridge. The integral of square of the difference between the two ADPs creates a functional which depends on the displacements of the secondary radiators at the deformed surface. Minimization of this functional by the method of global random search (using a genetic algorithm) leads to estimation of all the displacements of radiators and, consequently, the deviation of the points on the bridge surface. A similar method is used in [20] for antennas synthesis and even earlier in solving the inverse problem of measurement in linear inertial systems or systems with memory [21].

Let us consider this method. In [1], the authors have developed the functional,

$$J(s_0, s_1, ..s_{N-1}) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(f_m(\theta) - \sum_{i=0}^{N-1} A_i \cdot e^{j\left[\phi_i + \frac{2\pi}{\lambda} l_i \cos\theta - \frac{2\pi}{\lambda} s_i \sin\theta\right]} \right)^2 d\theta, \quad (2)$$

in which s_i, $i = 1, 2, ... N-1$ is a deflection of the surface at the point of the secondary radiator placing, $f_m(\theta)$ – the measured ADP of the radiators system from which the theoretical ADP of the same system, that is described in (2), is subtracted.

For an unloaded bridge $s_i = 0$

$$J(s_0, s_1, ..s_{N-1}) \approx 0$$

$s_i \neq 0$, there may be a difference between the two ADPs of the radiators system. Minimization of the functional (2) using a genetic algorithm results in the determination of unknown deflections s_i. It should be known, however, that formula (2) is the only one and has N unknown deflections. Since the dependence of the functional (2) on the deflections in the N

space is complicated, it is possible to find a non-global, but a local minimum of (2) which leads to incorrect ("phantom") solutions of the deflections determination. One or more of them may have an abnormal meaning. Since the deflections were fixed during the simulation, the "phantom" solutions could easily be distinguished from the correct ones. As shown in [1], when the ratio of the distance d between the radiators to the wavelength λ exceeds 0.5, a certain probability of "phantom" solutions begins to appear. It may reach a value of 0.6 with an increase in the ratio $\frac{d}{\lambda}$ to one and more. The reason for this phenomenon is the appearance of array diffraction maxima and increased number of the local extrema in the search space of genetic algorithm. Consequently, "phantom" solutions can be almost completely eliminated by rational choice of the ratio $\frac{d}{\lambda}$ of the radiating antenna array. In addition, the use of a priori information about surface deflections reduces their probability significantly, even when the ratio $\frac{d}{\lambda}$ is large. This information is often known from previous bridge research or can be obtained by other methods, for example, using the measurements made by mechanical sensors.

It should be noted that instead of passive secondary reflectors, one can use the antenna array of active radiators located at the same points. However, the cost of such system is probably higher.

3. Results of Experimental Research

For experimental verification of this effect, a "Kredo" coherent-pulsed radar with an electromagnetic wave length of 1.8 cm and a pulse duration of 0.4 μs (Figure 2) was used. The experimental setup also included a digital oscilloscope DSO-2250, USB (analogue-digital converter (ADC) with a sampling frequency of 100 MHz) and a laptop. In Figure 2a, a photo of a radar sounding path is shown, in Figure 2b – the placement of the angle reflectors against the background of a reinforced concrete wall, and in Figure 2c – the experimental installation. At a distance of 0.7 m from a reinforced concrete wall and at a height of 1.5 m, the array of five secondary radiators in the form of triangular horns (Figure 2b) was constructed; the distance between horns was equal to 1.5 m. The length of the horn side was 30 cm, and its depth was 15 cm.

The analog radar signals reflected from the horns were received by a mirror antenna, converted to the intermediate frequency and arrived at the input of the amplitude detector. The signal from the output of this amplitude detector entered the ADC of the oscilloscope, was converted into a digital code, and then into a text file (file with extension .txt), which is available in various computer mathematical packages.

To estimate the amplitude of reflected signals, an informative signal (Figure 3) consisting of 200 points was cut out from a long signal (5000 points). It corresponded to the distances at which the angle reflectors were located. In Figure 3, a blanking signal (long, up to a distance of approximately 80 m) which corresponds to the duration of the sounding signal is displayed initially and then a signal (of a shorter duration), that describes the amplitude of the signal reflected from the horns, appears.

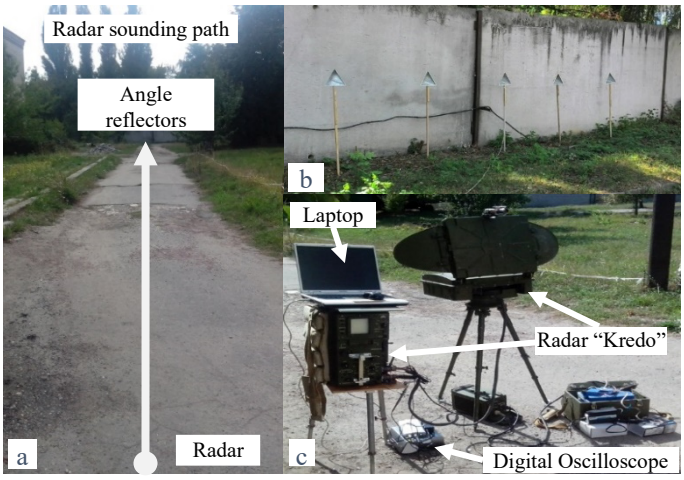


Figure 2: Photographic scheme of the experiment: radar sounding path (a); placement of the angle reflectors (b); general view of the experimental installation (c)

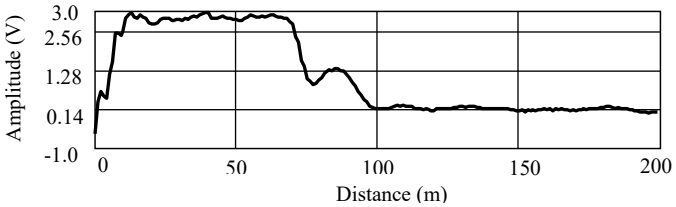


Figure 3: Distance distribution of amplitude of the signal received from angle reflectors

During the experiment the following positions of the angle reflectors were considered: a reinforced concrete wall without reflectors (variant 1); five reflectors on the same line (Figure 2b) (variant 2); one central reflector that was shifted forward by 20 cm (variant 3); two reflectors (central and next to the right) were shifted forward by 20 cm (variant 4); three reflectors (central and two next to the right) were shifted forward by 20 cm (variant 5); one reflector (rightmost) has been moved forward by 10 cm, two reflectors (central and next to it to the right) were shifted by 20 cm forward (variant 6); two rightmost reflectors were shifted forward by 10 cm, the central reflector was shifted by 20 cm forward (variant 7); three reflectors (central and two next to it to the right) were shifted forward by 10 cm (variant 8).

The reflected signals for the above-mentioned variants of the angle reflectors positions were recorded 10 times and then their values were averaged. The obtained experimental data were used to estimate the amplitude of the signals that were reflected from the angle reflectors with different placements relative to each other. The total field amplitude of secondary radiation of the horns at a range of 85 m was measured using the radar. The field amplitude at the radar receiver output was changed if any radiator was shifted in the radial direction (toward the radar location), that shows its sensitivity to movements of the radiators. For example, in the above-described variants, the experimental values in volt of the field amplitude were as follows: 1.16, 1.30, 1.39, 1.51, 1.39, 1.48, 1.40, 1.34. The shape of ADP of the secondary radiators system in experiments was not measured.

4. New Method for Estimating Bridges Surface Deflections

The problem of determination of the bridge surface deflections, described above, is technically complicated. The practical measurement of ADP for the secondary radiators system causes the greatest difficulty. First, we need to measure the field amplitude dependence on the angular coordinate, and it is not easy to do near a real bridge. Secondly, the measurements should be made in the far-zone of antenna at distance r_{fz}, calculated as two squares of antenna size L divided by the wavelength λ. For example, if $L = 30$ m and $\lambda = 30$ cm, then $r_{fz} \geq 6$ km and this value of far-zone distance is impossible for practice. Of course, with increasing wavelength and surface dimensions, whose deformation is measured, there are other problems associated with, for example, the efficiency of radiation of the passive elements of array. Consequently, practical application of the method presented in [1] is limited by the conditions of bridge sounding.

This problem can be solved on the basis of the amplitude-phase $(A_i,\ \phi_i)$ distribution measurement in the near-field zone of the antenna and obtaining the ADP of the radiators system on the basis of these measurements. Such problems were solved by many authors, for example, [22]. In practice, high-precision measurements of the amplitude and phase of the field in real weather and other conditions are complicated and expensive which reduces the value of this approach. Consequently, simpler methods of remote measurements of bridge surface deflections are required. This forces us to give up high-precision measurements of the amplitude and phase in the near-field zone of antenna.

The main idea of this article is to use only amplitude measurements in the near-field zone of the radiators system. The obtained data do not allow to determine the ADP but our task is to estimate only the bridge surface deflections. In this case, there is no need for complex operations of ADP measuring. Instead of these operations, we use a linear receiving antenna consisting of Q receiving elements, for example, horns located at points $P_q, (q = 1, 2, .., Q)$ (Figure 4) at the same distance δ from each other.

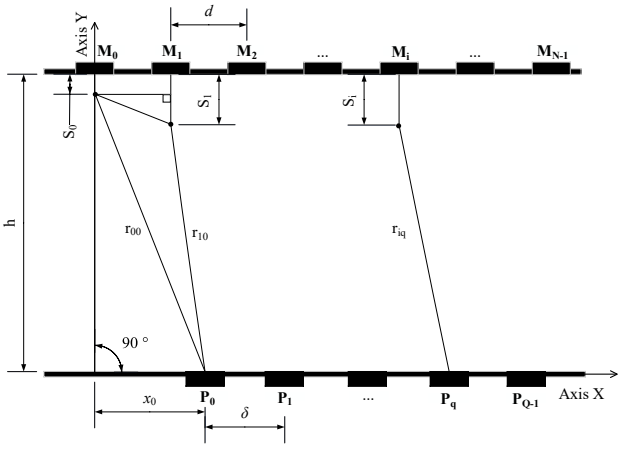


Figure 4: The layout of the receiving antenna array elements (points P_q) relative to the radiating elements position (points M_i)

The receiving element P_q is located from the transmitting element M_i at a distance which is easily determined from the geometry of Figure 4:

$$r_{iq} = \sqrt{(h - S_0)^2 + (x_0 + \delta \cdot q)^2 + (d \cdot i)^2 + (S_i - S_0)^2 - 2\Lambda_{iq} \cdot \Upsilon_{iq}}, \quad (3)$$

where $\Lambda_{iq} = \sqrt{(h - S_0)^2 + (x_0 + q \cdot \delta)^2}\sqrt{(d \cdot i)^2 + (S_i - S_0)^2}$,

$$\Upsilon_{iq} = \cos\left(\frac{\pi}{2} - \operatorname{arctg}\left(\frac{x_0 + q \cdot \delta}{h - S_0}\right) - \operatorname{arctg}\left(\frac{S_i - S_0}{d \cdot i}\right)\right),$$

h – distance between the lower surface of a bridge and array of receiving elements placed in parallel to the transmitting array;

S_i – amplitude of a bridge surface deflection at the points M_i, i. e. in locations of radiators;

x_0 – distance between the origin of the coordinates and the first in order of the receiving element P_0.

The complex amplitude of field strength at the output of each q receiving element is the sum of the fields created by all transmitting elements and is determined by the known formula [23] which for this case has the following form

$$\dot{E}_q = \sum_{i=0}^{N-1} \frac{\dot{A}_i}{r_{iq}} e^{-j\frac{2\pi}{\lambda} r_{iq}}, \quad (4)$$

where $\dot{A}_i = A_i e^{j\phi_i}$; A_i, ϕ_i – amplitude and phase of the field, which is created by the i radiator (further, for simplicity, initial phase field is taken equal to zero, i. e. $\phi_i = 0$).

We believe that all secondary radiators ADPs are equal and wide so that the extreme elements of the transmitting array irradiate all elements of the receiving array by energetically equivalent waves. The amplitude A_i of the field in the simulation was the same. Thus, the distance (3) makes the largest contribution to the total amplitude distribution of field strength. The amplitude of the field created by the secondary radiators at an unloaded bridge is determined similarly (4)

$$\dot{E}_q^0 = \sum_{i=0}^{N-1} \frac{\dot{A}_i}{r_{iq}^0} e^{-j\frac{2\pi}{\lambda} r_{iq}^0}, \quad (5)$$

where the upper index "0" indicates variables describing the position of radiators on an unloaded bridge, the surface shape of which is not necessarily a plane. These variables are easy determined before carrying out measurements. Due to the specificity of the problem geometry (Figure 4), the amplitudes of field strength will be different at the output of every receiving element. After the bridge loading the deflections s_i

λ

elements E_q depend only on the distances r_{iq} containing the deflections. The simulation was carried out in several stages:

- Determination of the field amplitudes E_q^0 at the outputs of the receiving elements of the array for an unloaded bridge, provided that the distances r_{iq} are given;
- Measurement of deflections s_i by mechanical sensors. In this article the typical values of bridge deflections were set by authors;
- Determination of the field amplitudes E_q at the outputs of the array receiving elements for a loaded bridge, provided that the distances r_{iq} and the deflections s_i are specified.

Thus, amplitudes of the field at the outputs of receiving elements of the array for an unloaded bridge and a loaded one were obtained in the simulation process for different conditions. It was further assumed that the deflections S_i, which were previously given, are now unknown. To determine these deflections and compare them with the given ones or for this case with reference deflections, a functional was composed:

$$J(s_0, s_1, ..s_{N-1}) = \left| E_q^0 - E_q \right| \Rightarrow \min \qquad (6)$$

It is minimized with the help of a genetic algorithm by variations on S_i. As a result, the "unknown" values of the bridge surface deflections were determined.

5. The Simulation Results

In Figure 5, 6 normalized field amplitudes at the outputs of the receiving antenna array elements located at distances $h = 10$ m (Figure 5) and $h = 50$ m (Figure 6) from the lower surface of a bridge are presented. The length of the electromagnetic wave was $\lambda = 20$ cm. Similar figures for the wavelength $\lambda = 5$ cm are shown in Figures 7, 8.

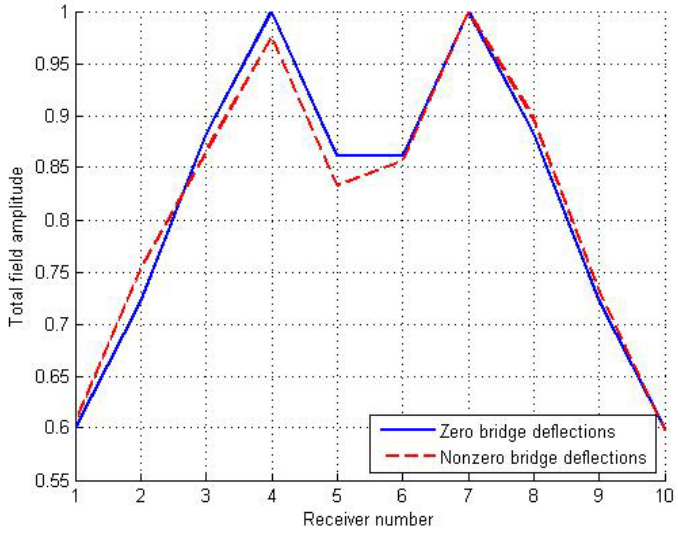


Figure 5: Normalized field amplitude at the outputs of the array receiving elements for an even (solid line) and deformed (dotted line) bridge surfaces for $\lambda = 20$ cm, h=10 m

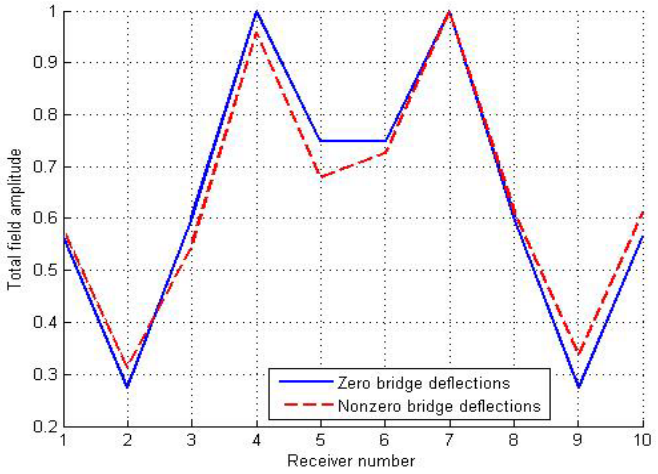


Figure 6: Normalized field amplitude at the outputs of the array receiving elements for an even (solid line) and deformed (dotted line) bridge surfaces for $\lambda = 20$ cm, h=50 m

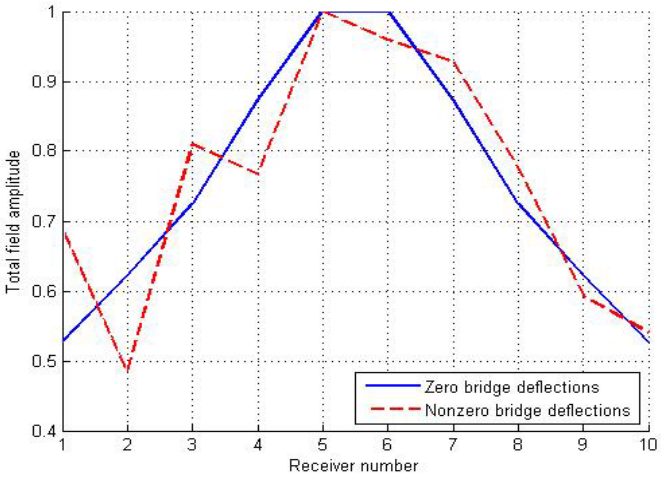


Figure 7: Normalized field amplitude at the outputs of the array receiving elements for an even (solid line) and deformed (dotted line) bridge surfaces for $\lambda = 5$ cm, h=10 m

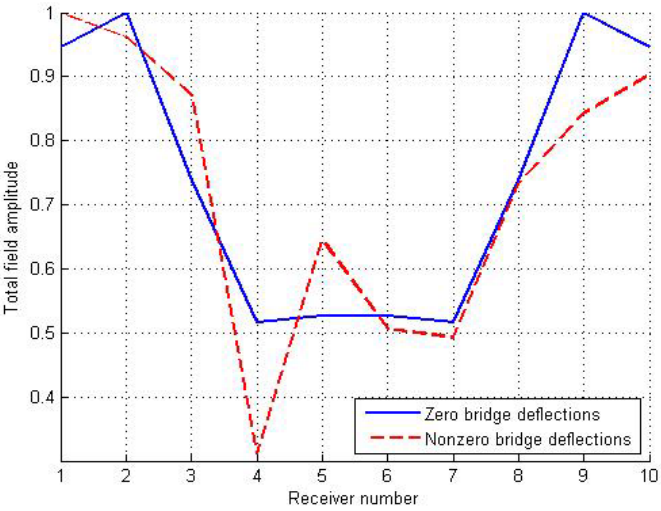


Figure 8: Normalized field amplitude at the outputs of the array receiving elements for an even (solid line) and deformed (dotted line) bridge surfaces for $\lambda = 5$ cm, h=50 m

The genetic algorithm in the process of its work "deforms" the dotted line in such a way that it takes a place of the continuous line. Since the latter is obtained for an unloaded bridge, the algorithm actually reduces existing deflections s_i almost to zero and at the same time determines their absolute values. In this case, the errors of these deflections determination arise. For modeling, the bridge surface deflections in millimeters were selected by authors as follows: $s_0 = 0$, $s_1 = 5$, $s_2 = 4$, $s_3 = 7$, $s_4 = 6$, $s_5 = 4$, $s_6 = 2$, $s_7 = 3$, $s_8 = 4$, $s_9 = 2$. The profiles of the deformed bridge surfaces and the profiles of the reconstructed surfaces that were determined by the genetic algorithm, are shown in Figure 9. Between them, there is some discrepancy which was interpreted as an error of deformed surface restoration.

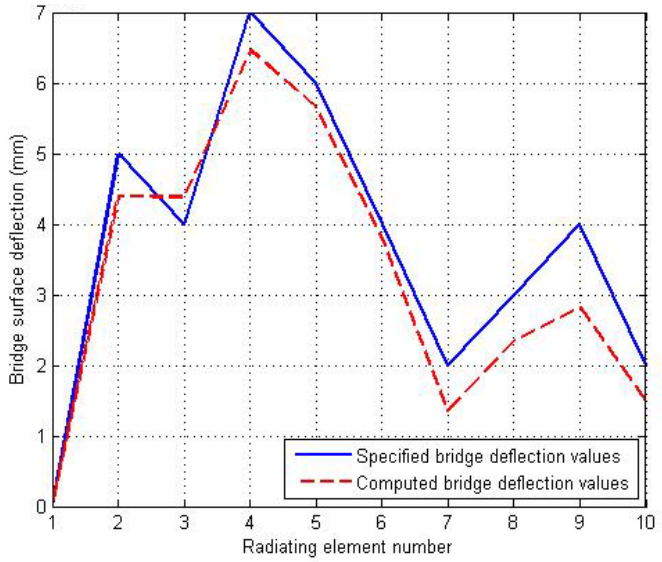


Figure 9: The profile of the deformed bridge surface (solid line) and the profile of the surface that was determined by genetic algorithm (dotted line)

The dependence of relative errors of surface profile restoration on the number of receiving elements in array for different wavelengths is shown in Figure 10.

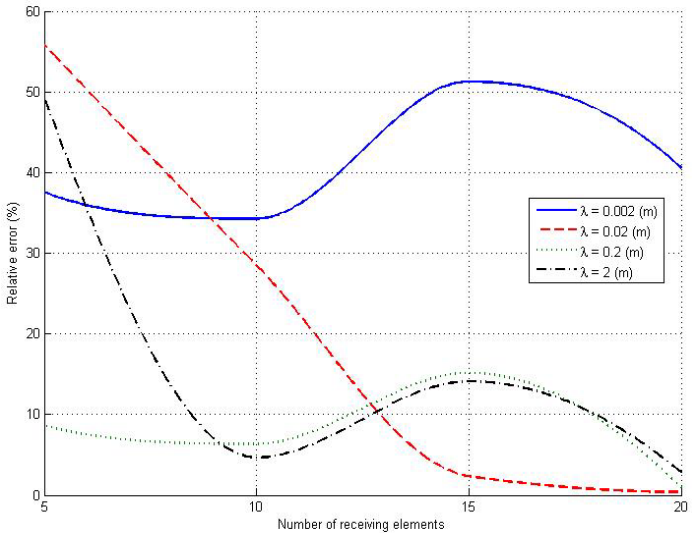


Figure 10: Dependences of the relative errors of the bridge deformed surface profile restoration on the number of receiving elements in array for different wavelengths

In the centimeter and decimeter range of wavelengths the relative errors of bridge surface profile determination are decreased with the increase in array receiving elements number and when the number 10 of receiving elements is exceeded, these errors are less than 10%. This quantity is limited in practice by reception conditions, because the method requires that all elements must receive the radiation from all transmitting elements. Relative error also depends on the wavelength which is used for radiation (Figure 11). The number of receiving elements is equal to 10.

For each bridge the dependence (Figure 11) will be different but there always is a wavelength range which provides the minimum error of the bridge surface deflection determination.

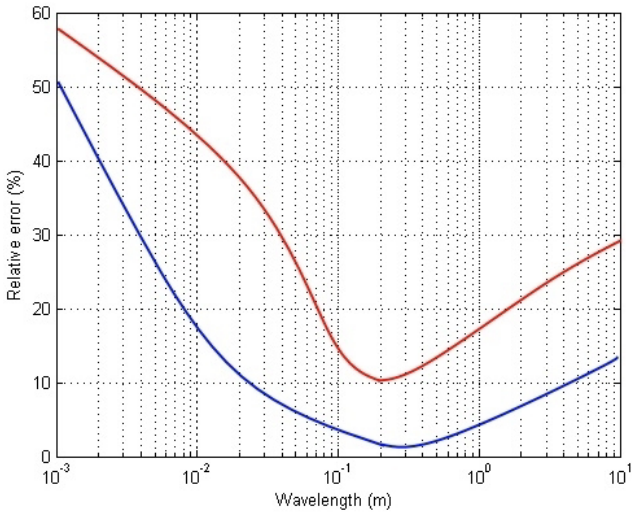


Figure 11: Dependences of the relative errors of the bridge deformed surface profile restoration on the wavelength: blue line – $h = 50$ m, $\delta = 3$ m; red line – $h = 100$ m, $\delta = 2$ m

Dependence of the relative errors of the bridge surface profile restoration on signal / noise ratio is represented in Figure 12. A deterministic signal describing the field amplitude and a white noise of different intensity were added in simulation. In this case the work of the genetic algorithm for finding the global minimum became more complicated.

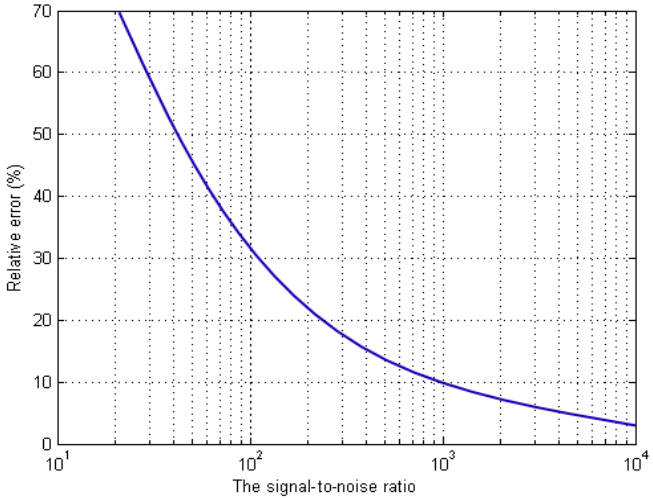


Figure 12: Dependence of the relative error of the bridge deformed surface restoration on the signal-to-noise ratio

The dependence of the "phantom" solutions probability on the ratio $\frac{d}{\lambda}$, that is, on the electrical distance between the radiators, is shown in Figure 13. Figure 14 demonstrates similar dependence on the number of receiving elements. The wavelength was 20 cm, and the distance between the lower surface of the bridge and the receivers was 50 m.

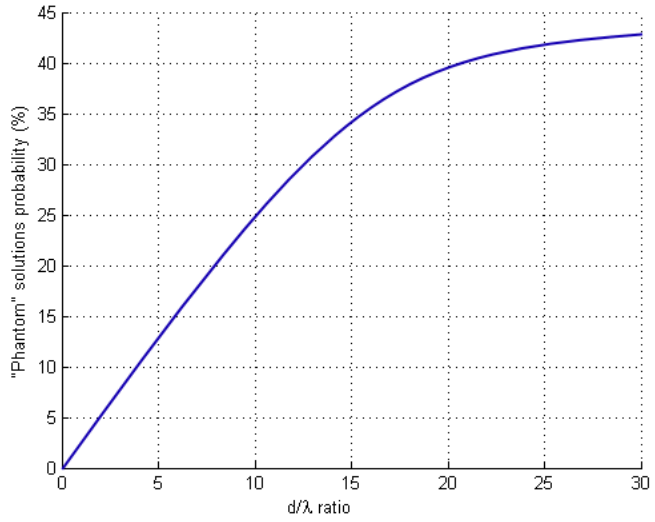


Figure 13: Dependence of "phantom solutions" probability on the ratio $\frac{d}{\lambda}$

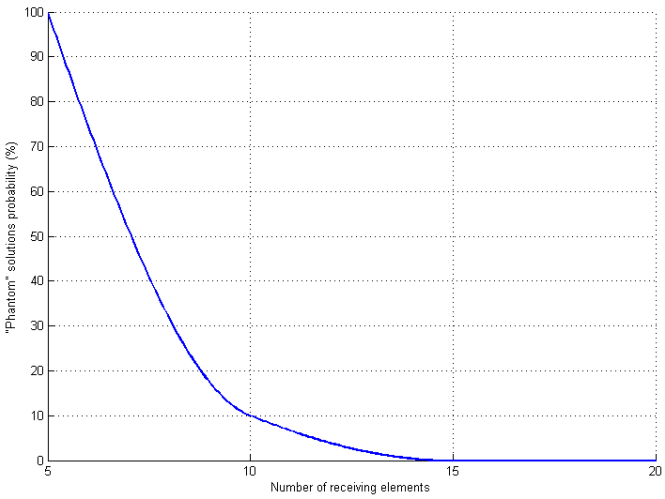


Figure 14: Dependence of "phantom solutions" probability on the number of receivers

The solution was considered as "phantom" one if the error of the bridge deformed surface restoration exceeded 10%. Figures 13 and 14 impose requirements for the number of radiators and receiving elements of the appropriate antenna arrays. Increased number of secondary radiators not only improves the quality of the curved surface profile measurement, but also significantly reduces the probability of the "phantom" solutions.

In the article, the problem of bridge surface deflections determination was solved for simplified conditions in the absence of radio waves reflection from the concrete part of the bridge surface and other parts, neglecting the influence of the steel elements of the reinforced concrete structure, etc. However, one should expect that the approach proposed in the article will not change. Indeed, in the process of the field amplitude measuring at the outputs of receiving elements of the array at an unloaded bridge, all the influencing factors which remain at a loaded bridge, will be taken into account. The difference between these fields appears only in the presence of bridge deflections in the second case. Hence the functional (6) will probably remain the same, that is, the method is expedient to implement in real conditions, too. However, the errors estimation for bridge surface deflections restoration in such conditions remains a challenging problem.

6. Conclusions

New methods for measuring surface deflection of bridges and other large objects were discussed in the paper. It is based on the processing of measurement information about the electromagnetic field amplitude which can be obtained from systems of secondary (passive) radiators (reradiators) or active radiators installed on the lower part of the bridge. The measuring information is obtained at the output of the receiving elements of a linear antenna array which can be installed at any convenient distance from the lower part of the bridge. Minimization of the functional developed in the study allows us to obtain numerical data on the movement of radiators, as well as the deflection of the entire surface of the bridge respectively. The paper defines the conditions under which the errors of deflections restoration in the process of the functional minimizing do not exceed the permissible ones. They depend on the electromagnetic wave length, the geometry of placement of the radiators and receiving elements, as well as on the level of noise in the measurement location. The application of a global random search of the functional extremum (using genetic algorithms) for the global minimum determination in certain situations leads to appearance of wrong ("phantom") solutions when the genetic algorithm finds the minimums that are not global. Recommendations to limit the probability of obtaining "phantom" solutions to a value of zero were also discussed. A mathematical model of electromagnetic waves scattering was made according to a simplified scheme, which did not take into account the reflection from the concrete surface of the bridge and other elements. The consideration of these factors will not make any significant changes to the principle of surface deflection determination, since in any case surface deflections certainly lead to a change in the field amplitude in the observation point, and these changes are the basis for determining deflections. The sensitivity of the field amplitude to the displacement of secondary radiators is proved in the paper experimentally. The proposed method can be used for determination of deflections of other large objects.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Poliarus O. V., Poliakov E. O., Brovko Ja. S. "Measurement method of the object surface deflections using antennas theory". - 2017 XI International Conference on Antenna Theory and Techniques ICATT'2017, May 24-27, 2017. - Kyiv, Ukraine, pp. 416-418.

[2] Yan Yu, Hang Liu, Dongsheng Li, Xingquan Mao, Jinping Ou. "Bridge Deflection Measurement using Wireless MEMS Inclination Sensor Systems". – International Journal on Smart Sensing and Intelligent Systems, 2013, vol. 6, №1, pp. 38…57.

[3] Xuefeng Zhao, Hao Liu, Yan Yu, Xiaodong Xu, Weitong Hu, Mingchu Li. "Bridge Displacement Monitoring Method Based on Laser Projection". – Sensing Technology. – Sensors, 2015, 15, pp. 8444…8463.

[4] Satoru Yoneyama and Hiroki Ueda. "Bridge Deflection Measurements Using Digital Image Correlation with Camera Movement Correction". – Material Transactions, 2012, vol. 53, №2, pp. 285…290.

[5] P. Kuras, T. Owerko, Ł. Ortyl, R. Kocierz, P. Kohut, K. Holak, K. Krupiński. Comparison of methods for measuring deflection and vibration of bridges. - Proceeding of the Joint international symposium on Deformation monitoring: 2–4 November 2011, Hong Kong, China, 2011, pp. 1..8.

[6] Pacglite, A. Pacglitis. "The Dynamic Amplification Factor of the Bridges in Latwia". – Journal Procedia Engineering, 2013, vol. 10, pp. 851…858.

[7] Brady S. P., O'Brien J. O., Znidaris A. "Effect of vehicle velocity on the dynamic amplification of a vehicle crossing a simply supported bridge". – Journal of Bridge Engineering, 11 (2), pp. 241…249.

[8] Poliarus O. V., Barchan V. V. "Radio engineering system for estimating of the dynamic amplification factor of bridge structure". – Kharkiv: Applied electronics, 2011, vol. 10, №3, pp. 368…371. (in original: Полярус О. В., Барчан В. В. Радіотехнічна система оцінки коефіцієнта динамічності мостової конструкції. – Харьков: Прикладная радиоэлектроника, 2011, т. 10, №3.)

[9] Poliarus O. V., Barchan V. V., Poliakov E. O. "Method of accuracy increasing of vibrational displacements measurements for a machine on the basis of measuring system integration". – Kharkiv: Technology of instrument making, 2012, №1, c. 25…29. (in original: Полярус О. В., Барчан В. В., Поляков Є. О. Методика підвищення точності вимірювань вібропереміщень машини на основі комплексування вимірювачів. – Харьков: Технология приборостроения, 2012, №1, с. 25…29)

[10] Carmelo Gentile. "Application of Microwave Remote Sensing to Dynamic Testing of Stay-Cables". - Remote Sensing, 2010, vol. 2, pp. 36-51.

[11] Patrick Paultre, Omar Chaallal, Jean Proulx. "Bridge dynamics and dynamic amplification factors – a review of analytical and experimental findings". – Can. J. Civ. Eng., 1992, №19, pp. 260-278.

[12] Cheol-Hwan Kim, Byung-Wan and Jin-Taek. "Application of Laser Vibrometer to the Measurement and Control of Cable Tensile Forces in Cable-Stayed Bridges". - International Journal of Distributed Sensor Networks, 2012, vol. ID 810682, 7 pages.

[13] Charles R. Farrar, Timothy W. Darling, Albert Migliori and William E. Baker. Microwave Interferometers for non-contact vibration measurements on large structures. - Los Alamos National Laboratory, 2000. – 14 p.

[14] Larry D. Olson, P. E. Olson. Recent advances in NDE and SHM of bridge superstructure with sonic and radar methods. – FHWA 2010, Nondestructive Evaluations of Bridges, 2010, pp. 1-12.

[15] Mayer, L., Yanev, B., Olson, L.D. and Smyth, A.W. "Monitoring of the Manhattan Bridge for Vertical and Torsional Performance with GPS and Interferometric Radar Systems", Proceedings CD, Transportation Research Board 89th Annual Meeting, 2010.

[16] Piotr Olaszek. "Investigation of the dynamic characteristic of bridge structures using a computer vision method".- Measurement, 1999, vol. 25, pp. 227–236.

[17] Shien Ri, Tatsuro Numayama, Masumi Saka, Kenichi Nanbara and Daisuke Kobayashi. "Noncontact Deflection Distribution Measurement for Large-Scale Structures by Advanced Image Processing Technique" - Materials Transactions, Vol. 53, No. 2, 2012, pp. 323…329.

[18] "Displacement Measurement of Burr Arch-Truss Under Dynamic Loading Based on Image Processing Technology". - 6th International Conference on Advances in Experimental Structural Engineering 11th International Workshop on Advanced Smart Materials and Smart Structures Technology August 1-2, 2015, University of Illinois, Urbana-Champaign, United States.

[19] R. Kaloop Mosbeh, Li Hui. "Analysis of bridge movements and damage based on GPS monitoring technique". – Интерэкспо Гео-Сибирь, 2010, вып. 1, т. 1, с.1-5.

[20] Poliarus O. V., Lebedynskiy A. V. "Methods of studying the antennas synthesis problem in practical lessons". – X Anniversary International Conference on Antenna Theory and Techniques ICATT'2015, April 21-24, 2015.-Kharkiv, Ukraine. – pp. 87-89.

[21] Полярус О. В., Поляков Є. О. "Наближене розв'язання оберненої задачі вимірювань та його метрологічне забезпечення": монографія . - Х. : Видавництво «Лідер», 2014. – 120 с. (in Ukrainian)

[22] Hirokazu Kobayashi, Yoshio Yamaguchi, Yi Cu. "Simple Near-field to Far-field Transformation Method Using Antenna Array-factor". - Journal of Wireless Networking and Communications, 2012, №2(4), pp. 43-48.

[23] Harish A. R., Sachidananda M. "Antennas and Wave Propagation". – Oxford University Press, 2007. – 411 p.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 225-235 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

A Taxonomy for Enhancing Usability, Flexibility, and Security of User Authentication

Susan Gottschlich*

Raytheon Co., 1001 Boston Post Road East, Marlborough, MA, 01752, 508-490-2339

ARTICLE INFO

Article history:
Received: 29 August, 2017
Accepted: 03 December, 2017
Online: 15 December, 2017

Keywords:
Continuous authentication
Multifactor authentication
Internet of Things
Software Defined
Wearable devices
Cybersecurity
Biometrics

ABSTRACT

Two technology trends - a move toward software defined capabilities and toward networked devices – support both unprecedented innovations and requirements for security. A fundamental aspect of security is user authentication, which allows devices and software applications to establish their user's identity and identity is in turn used to establish which of its capabilities the user is authorized to access. While multiple authentication steps, known as multifactor authentication, are being used more widely throughout the military, government, businesses, and consumer sectors, the selection and implementation of which authentication factors to require is typically defined by security policy. Security policy is in turn typically established by a security organization that may have no formal metrics or means to guide its selection of authentication factors. This paper will present a taxonomy for describing authentication factors including important attributes that characterize authentication robustness to aid in the selection of factors that are consistent with the user's mission. One particular authentication factor that I have developed will be discussed in the context of this taxonomy to motivate the need to broaden current definitions and security policies. The ultimate goal of this paper is to inspire the development of standards for authentication technologies to both support mission aware authentication innovation and to inform decision making about security policies concerning user authentication and authorization. Further, this paper aims to demonstrate that such an approach will fundamentally enhance both security and usability of increasingly networked, software-defined devices, equipment and software applications.

1. Introduction

Mobile devices, such as cell phones and tablets, can be extremely useful for military personnel, security personnel, employees, and consumers in performing an entire gambit of tasks from routine day to day tasks to time critical emergency response. Current cybersecurity best practices encourage the use of user authentication in order for these devices to be accessed. Increasingly, equipment that was once thought of as hardware, is becoming essentially software defined. This includes military radios, medical devices, and autonomous cars. This trend supports the possibility of requiring users to authenticate with their cars, medical devices, and radios, for example, before gaining access to their functionality, thereby enhancing security and discouraging theft and/or hacking. Further, Internet of Things (IoT) devices are becoming pervasive throughout homes and industries, and security policies for IoT device access have not consistently been established and implemented.

This paper is an extension of work originally presented in [1], where a method for continuous secondary factor authentication for military or security personnel required to perform missions while in harm's way was motivated and described. Specifically, this paper will use the Concept of Operation (ConOps) and method discussed in [1] and related ConOps to motivate the need for mission aware authentication factors and in turn to describe and validate authentication factors so that appropriate ones can be selected. This paper will further argue that standards and/or specifications and a means of validating compliance is crucial to support innovation and transition to productization.

In [2], the National Institute of Standards and Technology (NIST) defines multifactor authentication as "Authentication

* Corresponding Author: Susan Gottschlich, susan.gottschlich@raytheon.com

using two or more different factors to achieve authentication. Factors include: (i) something you know (e.g., password/PIN); (ii) something you have (e.g., cryptographic identification device, token); or (iii) something you are (e.g., biometric)." When only two factors are used, this is referred to two factor or dual factor authentication. Dual factor authentication is currently a requirement for NIST's Risk Management Framework (RMF). RMF compliance is a requirement being levied on all military information systems.

Continuous authentication is an emerging authentication factor methodology that implements authentication as a process rather than a simple event such as entering a user name and password. The idea is that once a user establishes his or her identity through another authentication factor or as an initialization step in a continuous authentication factor, the user is continuously monitored by some means so that their identity is continuously established.

Continuous authentication can enhance usability by alleviating the need for a user to re-authenticate, for instance, if they allow their authenticated session to become idle for longer than some inactivity period prescribed by security policy. Continuous authentication can also enhance security by potentially detecting instantaneously when a user may have lost control of his or her device or software application session, eliminating the need for security policies to establish arbitrarily the length of inactivity periods allowed before re-authentication is required.

This paper argues that in order to enhance both security and usability, user authentication should be pervasively implemented on software-defined networked or stand-alone devices or software applications, but that the factors used in authentication should be selected to be consistent with the user's underlying mission and/or need for using the device or application. Thus, the factors employed may be far different than current approaches that typically involve entering a user name and password and/or pin along with a secondary factor (something the user is or has).

Section II will present background on current work on authentication factors and on biometrics which are increasingly being used to support 'something you are' authentication factors.

2. Background

Until recently, user authentication has pervasively been implemented via the use of a password or pin as a primary factor (something you know). If a secondary factor is warranted, several options for 'something you have' have been implemented including an RSA key, a CAC card, an email account, or a token. Alternatively biometric factors (something you are) such as finger prints, iris scans or voice spectrograms have been used. Section 4 will review ConOps examples where such approaches might not be practical given the user's mission.

Various implementations of an emerging authentication approach – continuous authentication – are beginning to emerge. Some of these implementations are described in industry publications [3, 4], and in conference proceedings [5]. Continuous authentication is the idea that a user is monitored in such a way that they are being continuously authenticated. An approach presented in [5], for instance, uses a combination of color information of users' clothing and face information in order to robustly monitor a user who may not always be facing the computer being used. Other continuous authentication approaches presented in [3, 4] use behavior recognition, such as keystroke patterns (e.g. typing rhythm, mouse movement) potentially used in combination with other biometric factors such as iris patterns.

Continuous authentication can bring about more secure authentication that is also more usable than current pervasive security policies implement. For instance, if a user who is not being continuously authenticated walks away from his or her computer briefly, a nefarious agent can assume this user's computer session. Locking computer sessions during inactivity periods is meant to counter this possibility but if the inactivity period is too short, it can become extremely counterproductive for a user who must frequently re-authenticate.

Alternatively, allowing for a lengthy inactivity period negatively impacts the security posture of the computer session. Continuous authentication is theoretically more secure because it tracks the user continuously, thus no timeout or inactivity period need be implemented. It may also be considered more user friendly because users do not need to re-authenticate.

An important thing to note about continuous authentication implementations is that they are generally tied to the type of device being used. For instance, key stroke and mouse behavior monitoring applies to computers but not necessarily cell phones or IoT devices. Further, they are often dependent on a user's mission or role. Tracking clothing, irises, and keyboard behavior are certainly appropriate for office and home environments. They may be less appropriate in environments where users are wearing uniforms (clothing is always the same color across users and days) and protective eyewear, using their devices occasionally while they perform other challenging tasks (e.g. driving, patrolling), or put into high stress situations that may alter their behavior or voice substantially.

In [1] I presented an approach that I developed that uses a fitness tracking device paired with a mobile device to continuously monitor a user performing potentially hazardous missions. This approach locks the mobile device if conditions suggest that the user is under severe duress (possible captured), dangerously wounded or killed. In other words, this approach monitors the user and when it detects that the user may have lost control of his or her device, it fails the user's authentication with the device. Otherwise, it allows a user persistent uninhibited access to the mobile device.

The novelty of the approach is based on the recognition that all of the conditions that suggest a user has lost control of his or her device will bring about a sudden change in the user's biometrics as monitored by the fitness tracking device. Because this is the case, the various biometric sensors can be monitored

using Kalman filters which can detect jump (or sudden) changes in a biometric variable.

In [1] I describe experiments conducted with this continuous monitoring approach. The experimental results, which were derived from live experiments with 'normal' conditions but simulated experiments with dangerous conditions, confirm the approach's ability to provide strong authentication under the assumed user conditions.

As far as I am aware, this is only continuous authentication approach published that considers how to detect and lock down a device when conditions suggest that the user has lost control of his or her mobile device.

While this is generally not a concern for office workers or personal device users, it may be an important concern for military and security users. This differentiation further suggests the need for selecting mission-aware authentication factors.

When contrasted against NIST's definition of authentication factors, the fitness tracking device might be considered both something you have and a means to identifying something you are. In addition to tracking biometric readings, the fitness tracker's API also allows the authentication software to monitor band contact, so it is possible to detect when the user is no longer wearing the device or the tracker is no longer communicating. This paper will argue that combination factors and other innovations will be encouraged if security organizations are given the tools to evaluation their effectiveness.

Further, I submit that the emergence and commercialization of many new technologies including biometric feature sensing and recognition, wearable technologies, and machine learning (ML) will support more innovative and specialized approaches to user authentication.

There is substantial related literature on the biometric readings that might be drawn from in order to help assess the strength and risks associated with a biometric factor. In his Ph.D. dissertation, Gari D. Clifford discusses heart rate variability (HRV) and its causes [6]. Clifford shows that a person's heart rate is variably by nature. For instance, Circadian rhythms will drive variability throughout the day and factors such as general stress level and caffeine intake will cause day to day variations.

Several publications, including [7, 8, 9] consider heart rate variability, Galvanic skin response (which can be measured with some currently available fitness trackers), and other factors in analyzing what happens during exercise, duress, and the differences between the two. Based upon the research presented in [6-9], I developed rules presented in [1] to differentiate conditions that result in an increased heart rate.

Note that all of the biometric factors used in the algorithm in [1] can be measured by the Microsoft® Band 2, which was used in the experiments that were described. Skin conductance delta (also known as Galvanic skin response), for instance, supports the estimate of duress, but other factors can be used in differentiation. My research indicates that conditions resulting in a sudden decrease in heart rate (e.g. extreme cold, cardiac arrest) generally indicate the user is severely stressed by health or environmental issues, rather than an explicit threat.

Research presented in [10] performed a case study on students under stress due to a university examination. In [11] the use of wearable sensors is compared to electrocardiograms and concluded that the former is sufficient for detecting stress conditions.

Much of the research on heart rate measurements are focused on detecting an abnormal condition that may occur during a medical procedure such as surgery. The approaches discussed in such research generally involve preprocessing the measurement signal and performing machine learning (ML) techniques.

The algorithmic approach presented in [1] rejected the application of a pure ML techniques because measurable biometrics, such as heart rate, will vary from day to day due to many factors such as time of day, intake of caffeine or other drugs, and the user's current activity, and it is not practical to control for these factors.

Rather, the algorithm focuses on detecting jump changes in biometric measurements so that normal variances in measurements are implicitly handled.

The work in [12] describes the use of Electroencephalogram (EEG) for authentication. When mature, this technology may be trusted to uniquely authenticate an individual as a single factor (e.g. as opposed to dual factor or multifactor authentication). An unanswered question is how well these approaches work under high stress and maybe duress situation.

The work in [13] addresses biometric recognition techniques, including facial recognition, voice recognition, finger print recognition and iris recognition, being applied to authentication. It further describes how to assess biometric characteristics in terms of five qualities: robustness, distinctiveness, availability, accessibility and acceptability. The work further developed a taxonomy of uses. The focus of this work, however, is in 'snapshot' authentication – authenticating a user at a particular point in time using a snapshot reading (finger print, facial image, etc.) as opposed to the continuous streams of readings I proposed using in [1].

Nevertheless the discussion in [13] can easily be extended into the continuous monitoring domain if it can be assumed that there are mechanisms to continuously take snapshots of the discussed biometric factors.

In the following I will first propose a taxonomy for authentication factors that expands on the work presented in [13] in Section III. Specifically, I decompose the qualities presented in [13] into measurable and verifiable attributes. Further, these attributes are intended to support a broad array of potential authentication methodologies.

In Section IV, I present three ConOps to consider with regards to the taxonomy. I will conclude with a discussion of next steps.

3. Authentication Factor Taxonomy

The purpose of this section is to present a draft taxonomy for authentication factors and the user's that they are designed to identify and authorize. The goal is to motivate the development of industry and government standards organization that develops and maintains taxonomies and related standards that can be utilized by manufacturers of authentication devices and/or algorithms as well as security organizations or home users who must ultimately determine authentication policy. This draft has been developed informally as a means to motivate a formal taxonomy following a research methodology, for instance, as described in [14, 15].

The draft ontology is decomposed into three related groups of attribute/value pairs for describing authentication factors, user missions, and security policy.

Table I provides a list of key attribute/value pairs that can be used to provide specifications for an authentication factor. It is anticipated to be used by suppliers of authentication factors to provide an unambiguous description of the level of security a given authentication factor offering is able to support. It is also anticipated to be used by security organization to specify policy for the conditions under which a given methodology can be used.

This taxonomy can be seen as an extension of the characterization described in [13] because it encompasses the full range of authentication factors and it including attributes that might be used in a risk analysis (e.g. some connection protocols may be considered to be less secure than others). It further attempts to decompose the qualities into attributes that are measurable and verifiable.

This taxonomy strives to address current and emerging authentication methodologies. It is understood that future innovations may bring about methodologies that will require a further extension in the taxonomy. Nevertheless it is quite possible that future innovations will not change the attributes in the taxonomy but instead extend the range of potential values associated with each attribute.

It is the only taxonomy that I am aware of that is meant to model emerging authentication factor technology including continuous authentication and combination factors (e.g. something you have and something you are). I believe that research aimed at characterizing authentication factors is essential to promote innovation in authentication factor offerings. Without such a mechanism, innovators may have a hard time determining whether or not a proposed authentication concept will, if developed, be acceptable by security organizations and users, and thus less likely to expend resources developing innovative authentication approaches.

The first attribute, category, provides a high level simple characterization. Security policies may use this characterization to structure the definition of acceptable authentication factors. In turn, users may use it to determine the best factor(s) to implement their ConOps.

Table 1: Authentication factor definition

ATTRIBUTE	VALUE
Category (structure which provides high level distinctions to support policy definition)	-Discrete, continuous, or continuous with initial discrete authentication factor -Something you know, something you are, something you have, or a combination factor.
Strength (structure that varies depending on the underlying factor)	-Something you know: number of characters, special characters (for single entry), number of questions (for security question knowledge based authentication) -Something you have: restrictions on how the device is obtained, if/when it expires, policies for how lost, stolen, or damaged devices are replaced - Something you are: frequency of interrogation
False Acceptance Rate	Expressed as a percentage. May be calculable based on strength settings.
False Rejection Rate	Expressed as a percentage. May be calculable based on strength settings.
Validation Approach	The approach used to validate strength, false acceptance rate and false rejection rate established.
Connections required (something you have, something you are)	Wired interfaces including: USB, DS-101. Wireless interfaces including: WiFi, Bluetooth, ZigBee (IoT), Proximity card, Near-Field Communications (NFC)
Acceptable user conditions	A list of well-defined conditions including: Continuous use, continuous use with occasional interruptions, wearable-occasional use, pocketed-occasional use, protective face equipment, protective eye equipment, protective hand equipment, protective finger equipment, exertion-tolerant, stress-tolerant, voice use, covert use.
Authenticated devices and applications	A list of well-defined devices, including computer (keyboard, mouse, camera), tablet, phone, software defined equipment, IoT device, etc.as well as applications that the factor can support.
Acceptable devices for authentication factor (if required)	Varies depending on the factor (e.g. USB token, wearable device, software algorithm to run on device for use).
Additional events detected	A list of well-defined events including: duress, medical issue, fatigue, theft, tamper.

The next three attributes – strength, false positive rate and false negative rate – aim to support objective measurable specifications that address the 'robustness' and 'distinctiveness' qualities described in [13]. The strength attribute represents methods by which an authentication factor can be configured (from a supplier's point of view) or must be configured (from a security policy point of view). The strength attribute is particularly important for something that you know factors (e.g. passwords). For continuous factor authentication, strength may be described in terms of the sampling rate (how frequently classification/recognition is performed). For knowledge-based authentication (e.g. answering security questions), strength may be described in terms of the number of questions required to authenticate. The strength attribute is usually configurable.

False acceptance rate and false rejection rate are important for biometric and continuous authentication factors. They are specialized from the universally accepted receiver operating characteristic (ROC) curve mechanism for statistically characterizing the sensitivity and specificity of a binary classifier (for instance see [16]).

In the realm of user authentication, these measures are intended to capture statistically how likely a false user successfully authenticates or a true user fails to authenticate. This occurs because biometric recognition technologies (face recognition, voice recognition, etc.) are not perfect. Specifically, distinctiveness is addressed by false acceptance rate. A low false acceptance rate suggests that the underlying algorithm is able to distinguish a true user from false users.

These attributes are extremely important in helping to select which of perhaps many similar technologies is appropriate for a given application. They may not, however, be meaningful for more traditional authentication factors (e.g. something that you know) such as passwords as there is virtually no way for these factors to be misrecognized.

False acceptance and rejection rates are usually not directly configurable. However, it is frequently possible to tune them. For instance, lower false acceptance rates may be achievable by tuning algorithmic parameters in such a way that the false rejection rate increases as a result. Other factors, possible tied to strength, might also be tuned to decrease these rates. For instance, acceptance rates may increase or decrease depending on the frequency of classification/recognition.

The validation approach attribute is important in specifying how strength and acceptance rate specifications are validated. False acceptance rates and false rejection rates quotations may be very low if, for instance, statistical validation is performed using a small homogeneous sample set. Further, the validation approach may help a security organization determine under which conditions an authentication factor may be useful. For instance, voice recognition technologies may be robust in quiet lab environments but less so in outdoor public environments. This attributes addresses, in part, the acceptability and accessibility qualities from [13].

The connections requirement is particularly important for security organizations that need to perform a vulnerability analysis. Any connection interface is potentially vulnerable to exploitation, spoofing and/or denial. Further, organization may have implemented mechanisms for minimizing vulnerability for given connection types and thus prefer some technologies over others. Note that this attribute may not be relevant to 'something you know' authentication. This attribute addresses, at some level, the availability, accessibility, and acceptability qualities from [13].

The acceptable user conditions attribute supports the acceptability and accessibility qualities. Emerging authentication factors will likely be specialized for a given user condition or situation in order to support appropriate robustness while also maximizing convenience. For instance, continuous authentication based on facial recognition is only appropriate in situations where the user is expected to always be facing the face image capture component of the authentication factor. This may be appropriate for continuous use of a computer in an office environment, but may be less appropriate for occasional use devices that are frequently pocketed.

The next two attributes Authenticated devices and applications and Acceptable devices for authentication factor are simply characterizations of the equipment/software that a factor is intended to authenticate (e.g. computer, bank account session) and equipment used to perform the authentication. These attributes anticipate that suppliers may want to provide software-based solutions that can run on a variety of equipment. For instance, a factor that involves the use of a fitness tracker could potentially run on a variety of commercially available trackers. Thus, this attribute addresses the 'accessibility' quality discussed in [13].

Finally, the additional events detected attribute formalizes the concept that many potential authentication factors may have applications well beyond that of simple authentication. In particular, continuous biometric authentication factors may be able to detect events not directly related to authentication or security such as health events or fatigue. This attribute addresses the 'acceptability' quality from [13]. In particular, an organization may be more willing to invest in an authentication factor approach if it may bring additional benefits beyond robust authentication.

It should be noted that Table I is used to describe 'opt in' technologies where the user has opted to use an authentication factor in order to identify themselves. In biometric recognition and related areas, 'opt out' technologies are used to identify individuals without their direct compliance. While opt out technologies define an important research area, I do not believe that they are appropriate for authentication factors as most current uses for authentication implicitly involve users opting in in order to utilize the equipment or service being secured.

To achieve security policies that are appropriately flexible while maintaining a strong security posture, it is important to take into account a user's mission. In many situations today, security policies only allow a single combination of authentication factors. The result is that users may not use equipment or services that is too onerous to access, may seek

mechanisms to subvert the security policies decreasing the security posture, or may put up with the inconveniences reducing their overall effectiveness and safety. Future security policies might allow a user community to select their own factors and instead provide guidance on how the selected factors must be configured and used in order to be allowable.

Toward this end, Table II provides a draft taxonomy for users in the context of the mission they are performing and for which they need devices and applications authenticated. It is anticipated to be used by a mission leader. The first attribute, mission class, is intended to be a high level characterization to simplify selection and guidance of authentication factors.

Table 2: User (Mission) Definition

ATTRIBUTE	VALUE
Mission Classes	A list of well-defined mission classes including: Desk worker, Office worker not at desk, Factory worker, Home indoor user, home outdoor user, tactical user in hazardous situations, tactical user not likely in hazardous situations, tactical user in extreme environments, etc.
Anticipated session duration	Varies but used to help determine how often a user might be required to re-authenticate
False Acceptance Rate tolerance	Taken together these might be used to select appropriate factors and/or their configurations. For instance, a factor may be configured for extremely low acceptance rate while allowing higher rejections rates (possibly forcing user re-authentication)
False Rejection Rate tolerance	
Acceptable user anomalies	A list of well-defined conditions that are acceptable for a user to exhibit during a valid authentication session including: Stress, strenuous activity, covert use, extreme environmental conditions, etc.
Mission restrictions and conditions	A list of well-defined restrictions and conditions including: outdoor use, wired and wireless restrictions, etc.
Devices and/or applications requiring authentication	A list of devices or applications including computer, tablet, phone, software defined equipment, IoT devices and web site account sessions.
Authentication factor limitations	A list of well-defined restrictions including: no-keyboard requirement, no carry requirement

The second attribute, expected duration, is important in selecting technologies and methods. For instance, more rigorous initialization and/or calibration steps may be more acceptable for long duration uses. If the duration is anticipated to span multiple days then authentication approaches must deal with users sleeping and bathing. For some technologies, expected battery life on equipment is an important consideration.

False acceptance and rejection rate tolerances are important in helping mission leaders specify acceptable risks. For instance, a mission leader may need to consider the risk of an authenticated device falling into an adversary's hands. To the extent that recognition algorithms are tunable, a mission leader may insist on a tuning that is either very conservative for false acceptance rate or very conservative for false rejection rate.

Mission restricts and conditions are important for helping users consider what authentication technologies may be acceptable. Outdoor usage suggests weatherproofing and often battery operation. Wireless transmissions are ultimately observable and/or disruptable by an adversary, but some are less so than others.

Listing all of the equipment or services that a user might authenticate with in order to carry out their mission may spur the usage of authentication factors that can be applied to multiple devices simultaneously. For instance, a wearable device may potentially interface with all of the equipment a user may need to use during a given mission so that they are less encumbered by individual specific authentication processes.

Authentication device restrictions may be used to reject authentication factors. Something you have factors such as tokens or cards are problematic if a user damages, loses or forgets to bring them, particularly where there may be no viable back up options. This case may suggest a 'no carry' restriction. Users wearing protective hand gear may reject any approach requiring a keyboard or keypad.

Table III is intended to allow a mission leader to work with a security organization to develop appropriate security policies and to allow the mission leader to determine how best to perform his or her mission given acceptable risks. A security policy may allow a variety of primary factors to be used as long as they meet requirements spelled out in the security policy. The same should be possible for secondary factors. A supplier offering an authentication factor that is intended as a secondary factor might provide appropriate interfaces to a variety of primary factors in order to facilitate integration.

As wearable devices and biometric recognition algorithms become more capable, I believe that it should be possible to develop an authentication factor that can be used to 'unlock' a number of devices or services. In particular, software defined equipment could interface with other equipment to obtain authenticated user credentials providing a stronger security posture without placing additional requirements on the user.

Thus, future security policies should define to what degree they will allow multiple simultaneous devices and/or applications to be authenticated by a given authentication factor or combination of factors. Today, a similar concept is in common use, referred to as single sign on, where once a user authenticates with their computer, for instances, their credentials

are made available to software applications (avoiding the need to re-authenticate).

Table 3: Authentication Policy definition

ATTRIBUTE	VALUE
Primary factor requirements	Restrictions on the primary authentication factor. These requirements should specify restrictions on strength and mechanisms for re-authentication if 'something you know' is forgotten or 'something you have' is misplaced, forgotten, or damaged. It may further provide restrictions on secondary factors paired with it.
Secondary factor(s) requirements	Restrictions on the secondary factor(s) (can be more than one). Includes re-authentication mechanisms, restrictions on strength
Acceptable devices or services for use	A list of devices or services that can be simultaneously authenticated via a single (multifactor) authentication.
Anomaly policy	A list of well-defined actions to be taken including logging, notification, etc.
Rejection policy	A description of what happens when user authentication is rejected including logging, notification, locking session, etc.
Recovery policy	A description of how a valid user 'recovers' from authentication factor issue including forgetting information, losing, damaging or forgetting tokens, biometric measurement issues, etc.

Anomaly policy provides guidance on what actions are taken if anomalies are detected. For instance, an anomaly on a user's tablet might trigger a notification to the mission leader that might help the leader decide whether or not to take action.

Finally, rejection policies in the future may be more flexible than they are today. If a user loses control of their device, it may make sense to lock or even zeroize the device. However, if the possibility exists for them to regain control, a policy that might allow partial access to emergency features (e.g. ability to send emergency requests to a mission leader) or re-authentication features.

Note that Table I-III are not meant to be completely exhaustive. Rather their intent is to trigger discussions about how to completely define authentication technologies vis a vis user missions and security policies so that standards and/or specifications can be developed that are inclusive of a wide range of authentication technologies.

While the scope of situations for which this taxonomy is intended to address is extremely wide, Table IV attempts to outline a potential approach to using the three part taxonomy presented in this section to select an appropriate authentication methodology for a given mission.

4. ConOps and Authentication Limitations

To better motivate the need for innovation and flexibility in authentication factors and the security policies used to provide restrictions and requirements for them, this section will describe 2 ConOps for which currently acceptable assumptions made about authentication implementations may not hold and 1 for which they do.

The first ConOps was originally presented in [1]. Here a tactical user, such as a soldier, Marine, law enforcement officer, guard or emergency responder, uses a network-enabled cell phone or tablet to support their operations. The user's device might be of critical importance in helping them identify and locate persons of interests or threats, send out requests for back up or exfiltration, or report back relevant intelligence and observations.

The challenge here is that user authentication is extremely important to insure that if the user's device or they themselves are captured by a red force actor (enemy combatant, criminal, or other threat) that the red force actor is not able to exploit the device to gain access to critical information and/or intelligence. On the other hand, the user authentication should allow the tactical user uninhibited access to the device up until the moment that the tactical user loses control of it.

A fundamental assumption that we made in pursuing this work is that the target user community can be characterized by the following constraints:

- The user needs to access their device even when they are in harm's way and there is a possibility that the device may fall into the hands of adversaries including insider threats.
- The user will need to access their device while performing other tasks that will take precedence over device interaction.
- These may include driving, surveilling or monitoring, or physically interacting with civilians, criminals or other red force actors.
- Thus, authentication factors tied to inactivity periods are not practical as the tactical user may be inactive, vis a vis the device, throughout the majority of their tactical mission.
- The user may want to be able to access sensitive data and applications covertly (for instance, if they are undercover).
- This implies the use of commonly available devices, not specialized restricted equipment.

Table 4: Selecting Authentication Methodology for a Given Mission

CONSIDERATION	TAXONOMY RELEATIONSHIP	COMMENTS
What devices and/or applications are required for the mission and do they need to be protected?	Table II – Mission class, expected duration, devices and/or applications requiring authentication	Include the following: • Software defined equipment (e .g. radios, networks, sensors) • Applications running on computing devices such as computers, tablets, phones Consider if it is possible to implement authentication directly into the device or application: • Can authentication be added to or credentials shared with software defined equipment? • Can authentication be added to applications? Consider if alternative mechanism to protecting devices is possible: • Physical device not require protection but cloud based applications running on them do need to be protected. Is there an effective mechanism for authentication on the cloud? • Software defined equipment does not need to be protected as long as the infrastructure that it integrates with is able to provide protections
Risk of false acceptance	Table II - False Acceptance Rate tolerance	If a red force actor gains access to mission devices and/or applications, what are the risks (critical information access, ability to reengineer underlying technology, ability to perform malicious actions with capture devices and/or applications)? Are there other mechanisms to mitigate the risks?
Risk of false rejection	Table II - False Rejection Rate tolerance	If a user can no longer access their devices and/or applications, what are the issues (inability to communicate with teammates, inability to navigate mission area, loss of knowledge access critical to perform mission, inability to perform actions required for mission). Are there other mechanisms to mitigate the loss of access?
Restrictions on authentication factors	Table II - Authentication factor limitations	Consider if a user can use their fingers at all times that device interaction is required. Consider if a user can wear a device and where (wrist, finger, around neck). Consider if a user may forget something they know. Consider if a user may loss, damage, or forget something they have. Consider if a biometric feature will be available to measure (iris scan, finger prints).
What category(s) of authentication methodology are acceptable	Table I – Categories	Work with Security organization to find categories for which policies exist or develop new policies for given category.
What is the best authentication implementation?	Table I – Remaining attributes	Work with Security organization to select authentication factors and configurations to meet mission requirements and provide appropriate protection.
How to implement security policy?	Table III all attributes	Work with Security organization to determine how best to implement policy to provide appropriate protection and mission flexibility.

- The user's organization does not wish to waive cybersecurity controls associated with identification and authentication because the cybersecurity risk of the waiver is considered higher than is acceptable.

To support this ConOps, we developed a continuous authentication algorithm for monitoring the availability of a fitness tracker, worn by the user, and connected to the device via a consented Bluetooth wireless connection. The algorithm further estimating a user's health status based on biometric sensor readings read from the fitness tracker. In particular, we implemented the algorithm using both a Microsoft® Band 2 and Microsoft® Band 1 fitness tracker connected to a Microsoft® Surface tablet.

While the ultimate use of the proposed algorithm must be determined by the security policies of the organization responsible for managing the mobile devices, the proposed use of the algorithm is to provide a continuous stream of user health status estimations to be used as a secondary authentication factor. Certain statuses, when they are identified, will cause the device to lock and/or wipe sensitive data depending on organizationally defined procedures. In particular, the following situations may indicate authentication failure or partial failure limiting device access, especially to sensitive data:

- The fitness tracker loses contact with the wearer or the connection with the device implying that the user may no longer have his or her device and/or fitness tracker.
- The tactical user is under extreme duress implying that the user may no longer be in full control of the device or may be being manipulated by the threat.
- The tactical user may be dead or their health may be extremely compromised indicating that the user is unable to maintain control of the device.

In [1] we presented a jump Kalman filter (JKF) that continuously monitors the fitness tracker sensor readings looking for 'jump' discontinuities that might indicate one of the above three situations. Jump discontinuities are required to differentiate between, for example, duress and extreme physical exertion, the latter being an acceptable and even expected situation for the tactical user.

The benefit of using JKFs over explicit ML techniques is that JKFs are always using recent history to predict a user's next biometric reading. As long as the biometric readings do not change suddenly, the JKF's estimates will be very close to the actual readings. At times when the variability of the readings is high, the estimated covariance is high, so the allowable difference between the estimates and actual readings is increased. It is only when a true discontinuity in readings is observed that a jump is identified. This means that regardless of what a user's health state is after the user is initially authenticated by whatever means, as long as it does not change suddenly during the continuous monitoring, the JKF does not recognize a duress or other disqualifying condition. See [17,18] for other work related to JKF algorithms.

ML techniques can be very powerful, but many will require an initial training period. This may require retraining during every authentication session when using sensor readings or outputs that change a lot over time, e.g. from day to day. Of course, if the degree of difficulty involved in retraining is low, ML techniques might be very useful.

The implementation of the JKF algorithm and its configuration can be used to illustrate important considerations in user authentication. Ultimately, for a security organization to trust such an algorithm, they should require objective metrics on its performance in all relevant conditions. These include probability of false acceptance and probability of false rejection. Ideally such metrics might be computed by a supplier of this kind of authentication factor but then an important question is how these metrics are collected. Ultimately, it may be difficult for a commercial organization to thoroughly vet such an algorithm with actual users because it might require putting users into, at least convincingly, dangerous situations.

Further, this algorithm performs best when calibrated. Kalman filters normally are calibrated by simply collecting a small amount of history at the beginning of a session. In [1] we used 10 seconds of history to calibrate the JKFs. Jumps were recognized by the difference between actual and predicted readings exceeding the filter's covariance by greater than a predetermined constant factor:

$$|\boldsymbol{x_k} - \boldsymbol{z_k}| > C\ {\sigma_z}^2 \quad (1)$$

where $\boldsymbol{z_k}$ is the measurement, $\boldsymbol{x_k}$ the prediction, ${\sigma_z}^2$ the covariance, and C is a predetermined constant factor. One potential use of ML in conjunction with a JKF is in establishing the value of C, which may or may not be a constant. For instance, it is possible to imagine that the value of C may be best expressed as a function of other variables including the covariance. Ultimately, rigorous calibration mechanisms that are not unduly onerous for the user community are needed and security organizations need confidence in these mechanisms.

It should be noted that in addition to supporting user authentication, the approach presented in [1] supports a host of auxiliary requirements and features, including enhanced situation awareness of tactical users' status, early indicators of conflict or environmental hazards, and possibly early warning of health related issues (e.g. experiencing duress when not in a hazardous situation). It is likely that many of the other approaches being invented for user authentication might also be capable of supporting ancillary features.

A second ConOps is that of a tactical user, such as an Airman conducting an airborne mission. In such missions, physical access to mission aircraft is heavily restricted by armed guards. Because of this, in the past, authentication requirements may have been waived. In recent times, however, concern over insider threats, which may be inadvertent, accidental or deliberate, have forced tactical users to perform dual factor authentication. User authentication both limits access to the aircraft systems to trusted users and enforces nonrepudiation – the concept that an authenticated user cannot perform an action and later deny it.

An issue with requiring 'something you have' such as a token as a secondary authentication factor in such situations, is that if a user forgets, losses, or damages the card it may be necessary to scrub the mission.

Biometric factors may be complicated by the requirement to wear personal protective equipment. A further concern with biometric factors, such as finger prints, is that a red force actor can physically force a user (conscious or unconscious) to provide a reading to the authentication system.

One authentication approach for the tactical user that I believe will support strong security while alleviating concerns over authentication factor requirements involves a continuous authentication factor. Similar to the approach presented in [1], a continuous authentication factor tied to something that a tactical user may normally be wearing, such as a fitness tracker, may be used to establish and maintain the user's identity.

Further, this wearable device may be used to establish connections with all of the devices, such as radios and mission computers that the user needs to interact with in order to conduct their mission. The user could be automatically authenticated to all devices as long as the continuous authentication and connection to the devices remains viable. The continuous

authentication algorithm used to establish the user's identity and continually monitor the user may be changed to reflect different aspects of the mission than the one described in [1].

Wireless technologies are frequently disallowed in classified environments which may include mission aircraft. In order to use a continuous monitoring approach such as the one suggested, either a wired device would be required to monitor the user, or a wireless protocol that is not considered vulnerable must be utilized. Bluetooth, ZigBee (utilized in IoT devices), or Near Field communications (NFC) protocols are probably the most practical for such situations but they may require further encryption or other restrictions before they are allowed in a classified environment.

A third ConOps to consider is one that is implicitly addressed in much of the literature involving continuous authentication. It is that involving a knowledge worker who is usually using a computer or tablet. During the course of a normal day's work the worker would not be expected to undergo severe stress or duress. Any such situation would be considered abnormal and would warrant locking the worker out of their session until he or she reestablishes it via more traditional means such as requiring the entry of a password or PIN or calling a help desk. In this situation, video analytics using the computer or tablet camera, keystroke behavior, or a variety of other factors are likely practical to support strong and usable security.

5. Next Steps

This paper has sought to demonstrate that an increasingly networked software-defined technology footprint in our work and leisure environments is expanding the need and/or opportunity for user authentication before allowing access to computers, phones, IoT devices, services, and general equipment. At the same time, technology advances in wearable devices, biometric sensors, behavior analytics and a host of other fields is supporting the potential for more secure and less onerous authentication methodologies. Developing innovative new authentication factors to the point of productization and mass production, however, will be accelerated if inventors, investors, and manufacturers understand that security organizations and even home users will accept alternative authentication factor technologies to the relatively simplistic ones commonly utilized today.

Toward this end, this paper has attempted to motivate the formation or expansion of a standards body concerning user authentication technologies and policies. The taxonomy presented in Section III is primarily meant to provide early talking points in the formulation of a standards or specification body.

In one somewhat related example of how standards bodies can spur innovation and technology development, the U.S. National Security Agency (NSA) developed and published the Commercial Solutions for Classified Program (CSfC) specification [19]. Manufacturers and solution providers have used this specification to develop capability sets that they in turn have been able to have registered with NSA after providing evidence of compliance with the CSfC specification. This allows third parties to purchase and integrate compliant capability sets into their environments. In the end, the overall process of getting a classified environment certified by NSA has become far more flexible, affordable, and expedient because of the specification and associated registration process.

A similar approach applied to user authentication may ultimately support stronger and less onerous user authentication, and it may further support the expansion of user authentication and access restriction's use within an overall cybersecurity plan further enhancing security and usability at an enterprise level. For instance, if all equipment in a factory, office, or military platform required user authentication to read and/or modify equipment settings and/or to use the equipment, it may be much more difficult for the equipment to be 'hacked' because network penetration alone does not provide access to the equipment and software running on it.

While such an approach may seem unacceptable today, acceptance of emerging authentication methodologies will increase within security organizations as their strength and acceptance rates are proven and their risks are understood and mitigated. Further, acceptance within user communities will increase as authentication processes become simpler and less prone to error.

Similarly, user authentication techniques that are easy to understand and utilize might be used to enhance home security in a consistent and predictable fashion reducing deliberate and/or inadvertent actions causing harm to occupants and/or damage to or loss of property.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

This work was funded in part by Raytheon Internal Research and Development (IRAD) funds.

References

[1] S. Gottschlich, "Incorporating health monitoring and duress detection into mobile device authentication" in 2017 IEEE International Symposium on Technologies for Homeland Security

[2] Joint Task Force and Transformation Initiative, "Security and privacy controls for federal information systems and organizations" NIST Special Publication 800 (2013): 53

[3] B. Violino, "Continuous authentication: Why it's getting attention and what you need to know," in CSO, Mar 14, 2017

[4] R. Walters, "Continuous Authentication: The future of Identity and Access Management (IAM)" in Network World, Sep 16, 2016

[5] K. Niinuma and A. K. Jain. "Continuous user authentication using temporal information" in Proceedings SPIE, 7667(1), 2010.

[6] G. D. Clifford, "Signal processing methods for heart rate variability," Diss. Department of Engineering Science, University of Oxford, 2002.

[7] http://scienceline.org/2007/06/ask-hsu-fightorflight/.

[8] S. Boettger, V.K. Yeragani, L. Donath, H. J. Müller, H. H. Gabriel, K. J. Bär, "Heart rate variability, QT variability, and electrodermal activity during exercise" Med Science Sports Exercise 42(3), 443-448, 2010

[9] D. K. Spierer, E. Griffiths, T. Sterland, “Fight or flight,” The Tactical Edge, Summer 2009.

[10] P. Melillo, M. Bracale, L. Pecchia, “Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination” BioMedical Engineering OnLine, 10(1), 96, 2011

[11] J. Choi, R. Gutierrez-Osuna, "Using heart rate monitors to detect mental stress" in IEEE Sixth International Workshop on Wearable and Implantable Body Sensor Networks, 219-223, 2009

[12] R. J. Rodriguez. “An Electroencephalogram (EEG) Based Biometrics Investigation for Authentication: A Human-Computer Interaction (HCI) Approach,” Doctoral dissertation, College of Engineering and Computing, Nova Southeastern University, 2015.

[13] J. Wayman, A. Jain, D. Maltoni , D. Maio, “An Introduction to Biometric Authentication Systems,” in J. Wayman, A. Jain, D. Maltoni, D. Maio, (eds), Biometric Systems. Springer, 2005.

[14] R. C.Nickerson, U.Varshney, J. Muntermann, “A method for taxonomy development and its application in information systems,” European Journal of Information Systems, 22(3), 336-359, 2013

[15] A. Hevner, S. Chatterjee, Design research in information systems: theory and practice (Vol. 22), Springer Science & Business Media, 2010

[16] T. A. Lasko, J. G. Bhagwat, K. H. Zou, and L. Ohno-Machado, "The use of receiver operating characteristic curves in biomedical informatics" Journal of biomedical informatics, 38(5), 404-415, 2005

[17] A. K. Fletcher, S. Rangan, and V. K. Goyal, "Estimation from lossy sensor data: Jump linear modeling and Kalman filtering" in Proceedings of the 3rd international symposium on Information processing in sensor networks, 251-258, 2004.

[18] L. Galleani, Lorenzo, P. Tavella, "Detection of atomic clock frequency jumps with the kalman filter" IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 59(3), 504-509, 2012

[19] National Security Agency, “Commercial Solutions for Classified,” available: https://www.nsa.gov/resources/everyone/csfc

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 2, No. 6, 236-241 (2017)
www.astesj.com

ASTES Journal
ISSN: 2415-6698

A Study on Isomorphic Properties of Circulant Graphs

V. Vilfred Kamalappan *

Department of Mathematics, Central University of Kerala Periye, Kasaragod, Kerala, 671316 India

ARTICLEINFO

Article history:
Received: 13 October, 2017
Accepted: 06 November, 2017
Online: 20 December, 2017

Keywords:
Self-Complementary circulant graphs
Type-1 and Type-2 isomorphic circulant graphs
Cartesian product and factorization
prime and composite circulant graphs
Factorization theorem, Fundamental theorem

ABSTRACT

$C_n(R)$ denotes circulant graph $C_n(r_1, r_2, \ldots, r_k)$ of order n for a set $R = \{r_1, r_2, \ldots, r_k\}$ where $1 \leq r_1 < r_2 < \ldots < r_k \leq \left\lfloor \frac{n}{2} \right\rfloor$. Circulant graph $C_n(R)$ is said to have the Cayley Isomorphism (CI) property if whenever $C_n(S)$ is isomorphic to $C_n(R)$, there is some $a \in \mathbb{Z}_n^*$ for which $S = aR$. In this paper, isomorphic properties of circulant graphs that includes (i) Self-complementary circulant graphs; (ii) Type-2 isomorphism, a new type of isomorphism other than already known Adam's isomorphism of circulant graphs and (iii) Cartesian product and factorization of circulant graphs similar to the theory of product and factorization of natural numbers are studied. New abelian groups are obtained from these isomorphic circulant graphs. Type-2 isomorphic circulant graphs have the property that they are isomorphic graphs without Cayley Isomorphism (CI) property and thereby new families.

1 Introduction

This paper is an extension of work originally presented in ICMSAO2017 [1] and covers the author's study on a few isomorphic properties of circulant graphs that includes (i) Existence of self-complementary circulant graphs; (ii) Type-2 isomorphism, a new type of isomorphism other than already known Adam's isomorphism of circulant graphs that helps to obtain graphs without CI-property and abelian groups and (iii) Cartesian product and factorization of circulant graphs similar to the theory of product and factorization of natural numbers.

Beauty comes out of symmetry as well as asymmetry. Investigation of symmetries/asymmetries of structures yield powerful results in Mathematics. Circulant graphs form a class of highly symmetric mathematical (graphical) structures. In 1846 Catalan (cf. [2]) introduced circulant matrices and properties of circulant graphs have been investigated by many authors [1-20]. An excellent account of circulant matrices can be found in the book by Davis [2] and circulant graphs in the article [11].

If a graph G is circulant, then its adjacency matrix $A(G)$ is circulant. It follows that if the first row of the adjacency matrix of a circulant graph is $[a_1, a_2, \ldots, a_n]$, then $a_1 = 0$ and $a_i = a_{n-i+2}$, $2 \leq i \leq n$ [15, 17].

Through-out this paper, for a set $R = \{r_1, r_2, \ldots, r_k\}$, $C_n(R)$ denotes circulant graph $C_n(r_1, r_2, \ldots, r_k)$ where $1 \leq r_1 < r_2 < \ldots < r_k \leq \left\lfloor \frac{n}{2} \right\rfloor$. Only connected circulant graphs of finite order are considered, $V(C_n(R)) = \{v_0, v_1, v_2, \ldots, v_{n-1}\}$ with v_i adjacent to v_{i+r} for each $r \in R$, subscript addition taken modulo n and all cycles have length at least 3, unless otherwise specified, $0 \leq i \leq n-1$. However when $\frac{n}{2} \in R$, edge $v_i v_{i+\frac{n}{2}}$ is taken as a single edge for considering the degree of the vertex v_i or $v_{i+\frac{n}{2}}$ and as a double edge while counting the number of edges or cycles in $C_n(R)$, $0 \leq i \leq n-1$. Generally, write C_n for $C_n(1)$ and $C_n(1, 2, \ldots, \left\lfloor \frac{n}{2} \right\rfloor)$ for K_n. We will often assume, with-out further comment, that the vertices are the corners of a regular $n-gon$, labeled clockwise. Circulant graphs $C_{16}(1,2,7)$, $C_{16}(2,3,5)$, $C_{27}(1,3,8,10)$, $C_{27}(3,4,5,13)$ and $C_{27}(2,3,7,11)$ are shown in Figures 1 - 5. Now, let us consider the following definitions and results that are useful in the subsequent sections.

Definition 1.1. [17] *Let n and r be positive integers with $n \geq 4$ and $r < \frac{n}{2}$. Then, clearly, $C_n(r)$ consists of a collection of cycles $(v_0 v_r v_{2r} \ldots v_0)$, $(v_1 v_{1+r} v_{1+2r} \ldots v_1)$, $\ldots$, $(v_{r-1} v_{2r-1} v_{3r-1} \ldots v_{r-1})$. If $d = gcd(n, r)$, then there are d such disjoint cycles, each of length $\frac{n}{d}$. We say that each of these cycles is of period r, length $\frac{n}{d}$ and rotation $\frac{r}{d}$.*

If $r = \frac{n}{2}$, then obviously $C_n(r)$ is just a $1-factor$. Since $C_n(R)$ is just the union of the cycles $C_n(r)$ for $r \in R$, we have a decomposition of $C_n(R)$.

*Corresponding Author: V. Vilfred Kamalappan Email: vilfredkamalv@cukerala.ac.in, Contact No. 0091 9443577502

https://dx.doi.org/10.25046/aj020628

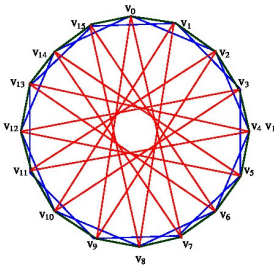


Fig. 1. $C_{16}(1,2,7)$

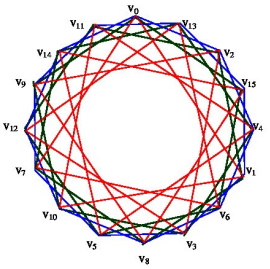


Fig. 2. $C_{16}(2,3,5)$

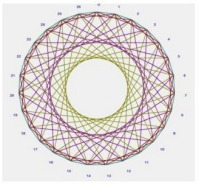


Fig. 3. $C_{27}(1,3,8,10)$

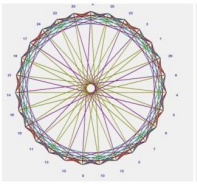


Fig. 4. $C_{27}(3,4,5,13)$

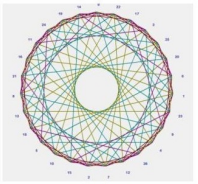


Fig. 5. $C_{27}(2,3,7,11)$

Theorem 1.2. [17] *Let $r \in R$. Then, in $C_n(R)$, the length of a cycle of period r is $\frac{n}{gcd(n,r)}$ and the number of disjoint periodic cycles of period r is $gcd(n,r)$.* □

Corollary 1.3. [17] *In $C_n(R)$, the length of a cycle of period r is n if and only if $gcd(n,r) = 1$, $r \in R$.* □

Remark 1.4. [17] *Let $|R| = k$. Then, the circulant graph $C_n(R)$ for a set $R = \{r_1, r_2, \ldots, r_k\}$ is $(2k-1)-regular$ if $\frac{n}{2} \in R$ and $2k-regular$ otherwise.* □

The following Lemmas are useful to obtain one-to-one mappings.

Lemma 1.5. [15] *Let A and B be two non-empty sets. Let $f : A \to B$ be a mapping. Then, f is one-to-one if and only if $f/_{A'}$ is one-to-one for every non-empty subset A' of A.* □

Lemma 1.6. [15] *Let A and B be non-empty sets and $A_1, A_2, \ldots, A_k$ be a partition of A (each A_i being non-empty, $i = 1, 2, \ldots, k$). Let $f : A \to B$ be a mapping. Then f is one-to-one if and only if $f/_{A_i}$ is one-to-one for every i, $i = 1, 2, \ldots, k$.* □

2 On self-complementary circulant graphs

In 1962, Horst Sachs [14] proved that the sufficient condition for the existence of a self-complementary circulant graph of order n is that every prime factor p of n should satisfy $p \equiv 1 \pmod 4$. He also conjectured that the sufficient condition is a necessary one. We have proved that the self-complementary circulant graph on n vertices does not exist if n has any prime factor which is not of the form $4m+1$, $m \in \mathbb{N}$. Thereby, we establish that the sufficient condition is also a necessary one. The proof is based on counting the number of disjoint cycles of a particular length in K_n and is given in this section [4, 17]. Graphs given in Figures 6 and 7 are self-complementary but not of circulant whereas graphs given in Figures 8 and 9 are self-complementary circulant graphs.


Fig.6. G

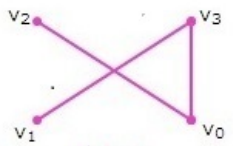


Fig.7. G

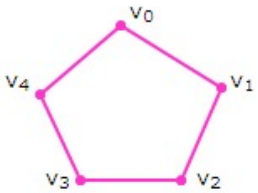


Fig.8. $C_5(1)$

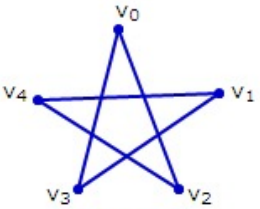


Fig.9. $C_5(2)$

Theorem 2.1. [17] *If $C_n(R) \cong C_n(S)$, then there is a bijection f from R to S so that for all $r \in R$, $gcd(n,r) = gcd(n, f(r))$.*

Proof. The proof is by induction on the order of R. □

Theorem 2.2. [9] *If graph G of order n is self-complementary, then $n \equiv 0, 1 \pmod 4$.* □

Theorem 2.3. [17] *If $C_n(R)$ for a set $R = \{r_1, r_2, \ldots, r_k\}$ is self-complementary, then $n \equiv 1 \pmod 4$.*

Proof. Using Theorem 2.2, we get $n \equiv 0, 1 \pmod 4$. When n is even, a circulant graph is of even degree if and only if its complement is of odd degree since any circulant graph is a regular graph. Thus, self-complementary circulant graph of even order doesn't exist. Hence, we get the result. □

Theorem 2.4. [17] *If $C_n(R)$ for a set $R = \{r_1, r_2, \ldots, r_k\}$ is self-complementary, then $n = 4m+1$, $k = m$ and $|S_i|$ is even where $S_i = \{j : gcd(n,j) = gcd(n,i),\ 1 \le j \le \left[\frac{n}{2}\right]\}$, for all i, $i = 1, 2, \ldots, \left[\frac{n}{2}\right]$.*

Proof. Using Theorem 2.3, we get, $n = 4m+1$. This implies that the degree (of each vertex) of a self-complementary circulant graph of order $4m+1$ is $2m$ which implies $k = m$.

Let $C_{4m+1}(R)$ be a self-complementary circulant graph for $R = \{r_1, r_2, \ldots, r_m\}$. If it contains a cycle of period r (and of length $\frac{n}{gcd(n,r)}$, using Theorem 1.2), then its complement also contains a periodic cycle of period, say, s such that $\frac{n}{gcd(n,s)} = \frac{n}{gcd(n,r)}$, $1 \le r, s \le \left[\frac{n}{2}\right]$, using Theorem 1.2. This implies that $gcd(n,s) = gcd(n,r)$. Here we consider that the cycles of periods r_i and $n - r_i$ are the same in $C_n(r_1, r_2, \ldots, r_k)$, $1 \le i \le k$. Combining the above arguments, we get the result. □

Remark 2.5. *The above theorem states that if a self-complementary circulant graph of order n exists, then $n = 4m+1$ and so it is $2m$-regular and the number of periodic cycles of length i in K_n is always even for each i, $1 \leq i \leq \left[\frac{n}{2}\right]$.*

Theorem 2.6. [17] *Self-complementary circulant graph of order n doesn't exist when n has any prime factor of the form $2(2m-1)+1$, $m \in \mathbb{N}$.*

Proof. Using Theorem 2.3, $n = 4a+1$, $a+1 \in \mathbb{N}$.

Let $n = p_1^{n_1} p_2^{n_2} \dots p_j^{n_j}$ where $p_1, p_2, \dots, p_j$ are the (odd) prime factors of n. Let $p_i \neq 4m+1$ for at least one i, $1 \leq i \leq j$ and for any $m \in \mathbb{N}$. This implies, $p_i = 2(2q-1)+1$ for some $q \in \mathbb{N}$. Consider the circulant graph $C_{4a+1}(1, 2, \dots, 2a) \cong K_{4a+1} = K_n$. Let r be the natural number such that $r = \frac{n}{p_i}$.

Aim To find out all natural numbers lying between 1 and n such that g.c.d. of each one of them with n is exactly r.

Let $gcd(n, pr+s) = r$ where $p, s+1 \in \mathbb{N}$ such that $0 \leq s < r$ and $1 \leq pr+s \leq n = rp_i$. This implies that $s = 0$ and so $gcd(n, rp+s) = gcd(n, rp) = gcd(rp_i, rp) = r$ where $1 \leq pr \leq p_i r = n$. Thus we get $gcd(p, p_i) = 1$ where $p \leq p_i$ which is greater than 1. Therefore the possible values of p are $1, 2, \dots, p_i - 1$.

Thus $r, 2r, 3r, \dots, (p_i-1)r$ are the only numbers lying between 1 and n such that g.c.d. of n with each one of them is exactly r. This implies $r, 2r, 3r, \dots, (p_i-1)r$ are the possible periods of cycles of length p_i each, in $C_n(1, 2, \dots, \left[\frac{n}{2}\right])$ since the length of a cycle of period rp in $C_n(1, 2, \dots, \left[\frac{n}{2}\right])$ is $\frac{n}{gcd(n,rp)} = \frac{rp_i}{gcd(rp_i,rp)} = \frac{p_i}{gcd(p_i,p)} = p_i$ for $p = 1, 2, \dots, p_i - 1$, using Theorem 1.2.

In $C_n(1, 2, \dots, \left[\frac{n}{2}\right])$,

the cycles of period r and $n - r(= (p_i-1)r)$ are the same,

the cycles of period $2r$ and $(p_i-2)r$ are the same,

...

the cycles of period $\frac{(p_i-1)r}{2}$ and $\frac{(p_i+1)r}{2}$ are the same.

Thus, there are $\frac{p_i-1}{2}$ number of possible distinct periodic cycles of (periods $r, 2r, \dots, \frac{p_i-1}{2}$ and) length p_i, each in $C_n(1, 2, \dots, \left[\frac{n}{2}\right])$.

Now, $\frac{p_i-1}{2} = 2q-1$ is an odd number. This implies, any circulant graph of order n and its complementary circulant graph contain unequal number of periodic cycles of length p_i, each. This implies that self-complementary circulant graph of order n does n't exist when n contains any prime factor of the form $2(2q-1)+1$, $q \in \mathbb{N}$, by Remark 2.5. □

Thus, when $n = 9, 21, 33, 49, 57, 69, 77, 81, 93$, etc., self-complementary circulant graph doesn't exist on n vertices, by Theorem 2.6, even though in each case, $n \equiv 1 \pmod 4$.

Now, by combining Theorem 2.6 and the sufficient condition for the existence of a self-complementary circulant graph of order n, we get the following result.

Theorem 2.7. [17] *The necessary and sufficient condition for the existence of a self-complementary circulant graph of order n is that each prime factor p of n should satisfy $p \equiv 1 (mod\ 4)$.* □

3 On Isomorphism of Circulant Graphs

In this section, Type-2 isomorphism, a new type of isomorphism different from already known Adam's isomorphism of circulant graphs, main results related to it and families of new abelian groups obtained from isomorphic circulant graphs are presented. Type-2 isomorphic circulant graphs have the property that they are isomorphic graphs without Cayley Isomorphism (CI) property.

Definition 3.1. [12] *A circulant graph $C_n(R)$ is said to have the CI-property if whenever $C_n(S)$ is isomorphic to $C_n(R)$, there is some $a \in \mathbb{Z}_n^*$ for which $S = aR$.*

Lemma 3.2. [16] *Let S be a non-empty subset of $\mathbb{Z}_n$ and $x \in \mathbb{Z}_n$. Define a mapping $\Phi_{n,x} : S \to \mathbb{Z}_n$ such that $\Phi_{n,x}(s) = xs$ for every $s \in S$ under multiplication modulo n. Then, $\Phi_{n,x}$ is bijective if and only if $S = \mathbb{Z}_n$ and $gcd(n, x) = 1$.* □

Definition 3.3. [3] *Circulant graphs $C_n(R)$ and $C_n(S)$ for $R = \{r_1, r_2, \dots, r_k\}$ and $S = \{s_1, s_2, \dots, s_k\}$ are Adam's isomorphic if there exists a positive integer x relatively prime to n with $S = \{xr_1, xr_2, \dots, xr_k\}_n^*$ where $< r_i >_n^*$, the reflexive modular reduction of a sequence $< r_i >$ is the sequence obtained by reducing each r_i modulo n to yield r_i' and then replacing all resulting terms r_i' which are larger than $\frac{n}{2}$ by $n - r_i'$.*

Lemma 3.4. [16] *Let $m, r, t \in \mathbb{Z}_n \in gcd(n, r) = m > 1$ and $0 \leq t \leq \frac{n}{m} - 1$. Then the mapping $\Theta_{n,r,t} : \mathbb{Z}_n \to \mathbb{Z}_n$ defined by $\Theta_{n,r,t}(x) = x + jtm$ for every $x \in \mathbb{Z}_n$ under arithmetic modulo n is bijective where $x = j + qm$, $0 \leq j \leq m-1$, $0 \leq q \leq \frac{n}{m} - 1$ and $j, q \in \mathbb{Z}_n$.* □

Theorem 3.5. [16] *Let $V(C_n(R)) = \{v_0, v_1, \dots, v_{n-1}\}$, $V(K_n) = \{u_0, u_1, \dots, u_{n-1}\}$, $r \in R$ and $gcd(n, r) = m > 1$. Then the mapping $\Theta_{n,r,t} : V(C_n(R)) \to V(K_n)$ defined by $\Theta_{n,r,t}(v_x) = u_{x+jtm}$ and $\Theta_{n,r,t}((v_x, v_{x+s})) = (\Theta_{n,r,t}(v_x), \Theta_{n,r,t}(v_{x+s}))$ for every $x \in \mathbb{Z}_n$, $x = j + qm$, $0 \leq j \leq m-1$, $0 \leq q, t \leq \frac{n}{m} - 1$ and $s \in R$, under subscript arithmetic modulo n, for a set $R = \{r_1, r_2, \dots, r_k, n - r_k, n - r_{k-1}, \dots, n - r_1\}$ is one-to-one, preserves adjacency and $\Theta_{n,r,t}(C_n(R)) \cong C_n(R)$ for $t = 0, 1, 2, \dots, \frac{n}{m} - 1$.* □

Definition 3.6. [16] *For a given $C_n(R)$ and for a particular value of t, $0 \leq t \leq \frac{n}{m} - 1$, if $\Theta_{n,r,t}(C_n(R)) = C_n(S)$ for some $S \subseteq [1, \frac{n}{2}]$ and $S \neq xR$ for all $x \in \Phi_n$ under reflexive modulo n, then $C_n(R)$ and $C_n(S)$ are called Type-2 isomorphic circulant graphs w.r.t. r, $r \in R$. In this case, subsets R and S of $\mathbb{Z}_n$ are called Type-2 isomorphic subsets of $\mathbb{Z}_n$ w.r.t. r.*

Clearly, Type-2 isomorphic circulant graphs are circulant graphs without CI-property. We obtained the following results on Type-2 isomorphism.

Theorem 3.7. [16] *For $n \geq 2, k \geq 3, 1 \leq 2s-1 \leq 2n-1$, $n \neq 2s-1$, $R = \{2s-1, 4n-2s+1, 2p_1, 2p_2, \ldots, 2p_{k-2}\}$ and $S = \{2n-(2s-1),\ 2n+2s-1, 2p_1, 2p_2, \ldots, 2p_{k-2}\}$, circulant graphs $C_{8n}(R)$ and $C_{8n}(S)$ are Type-2 isomorphic (and without CI − property) where $gcd(p_1, p_2, \ldots, p_{k-2}) = 1$ and $n, s, p_1, p_2, \ldots, p_{k-2} \in \mathbb{N}$.* □

Theorem 3.8. [16] *For $R = \{2r-1, 2s-1, 2p_1, 2p_2, \ldots, 2p_{k-2}\}$, $n \geq 2, k \geq 3, 1 \leq t \leq [\frac{n}{2}], 1 \leq 2r-1 < 2s-1 \leq [\frac{n}{2}]$, $gcd(p_1, p_2, \ldots, p_{k-2}) = 1$ and $n, r, s, t, p_1, p_2, \ldots, p_{k-2} \in \mathbb{N}$, if $\Theta_{n,2,t}(C_n(R))$ and $C_n(R)$ are Type-2 isomorphic circulant graphs for some t, then $n \equiv 0\ (mod\ 8)$, $2r-1+2s-1 = \frac{n}{2}$, $t = \frac{n}{8}$ or $\frac{3n}{8}$, $2r-1 \neq \frac{n}{8}, 1 \leq 2r-1 \leq \frac{n}{4}$ and $n \geq 16$.* □

Definition 3.9. [1] *Let $Ad_n = \{\Phi_{n,x} : x \in \Phi_n\}$, $Ad_n(R) = \{\Phi_{n,x}(R) : x \in \Phi_n\} = \{xR : x \in \Phi_n\}$ and $Ad_n(C_n(R)) = T1_n(C_n(R)) = \{\Phi_{n,x}(C_n(R)) : x \in \Phi_n\} = \{C_n(xR) :\ x \in \Phi_n\}$ for a set $R = \{r_1, r_2, \ldots, r_k,\ n-r_k,\ n-r_{k-1}, \ldots, n-r_1\} \subseteq \mathbb{Z}_n$. Define $'\circ'$ in $Ad_n(C_n(R))$ such that $\Phi_{n,x}(C_n(R)) \circ \Phi_{n,y}(C_n(R)) = \Phi_{n,xy}(C_n(R))$ and $C_n(xR) \circ C_n(yR) = C_n((xy)R)$ for every $x, y \in \Phi_n$, under arithmetic modulo n. Clearly, $Ad_n(C_n(R))$ is the set of all circulant graphs that are Adam's isomorphic to $C_n(R)$ and $(Ad_n(C_n(R)), \circ) = (T1_n(C_n(R)),\ \circ)$ is an abelian group and we call it as the Adam's group or Type-1 group on $C_n(R)$ under $'\circ'$.*

Definition 3.10. [1] *Let $V(C_n(R)) = \{v_0, v_1, \ldots, v_{n-1}\}$, $V(K_n) = \{u_0,\ u_1,\ \ldots,\ u_{n-1}\}$, $r \in R$, $m, q, t, t', x \in \mathbb{Z}_n$ such that $gcd(n, r) = m > 1$, $x = j + qm$, $0 \leq j \leq m-1$ and $0 \leq q, t, t' \leq \frac{n}{m} - 1$. Define $\Theta_{n,r,t} : \mathbb{Z}_n \to \mathbb{Z}_n$ and $\Theta_{n,r,t} : V(C_n(R)) \to V(K_n)$ such that $\Theta_{n,r,t}(x) = x + jtm$, $\Theta_{n,r,t}(v_x) = u_{x+jtm}$ and $\Theta_{n,r,t}((v_x, v_{x+s})) = (\Theta_{n,r,t}(v_x), \Theta_{n,r,t}(v_{x+s}))$ for every $x \in \mathbb{Z}_n$ and $s \in R$, under subscript arithmetic modulo n. Let $s \in \mathbb{Z}_n$, $V_{n,r} = \{\Theta_{n,r,t} : t = 0, 1, \ldots, \frac{n}{m} - 1\}$, $V_{n,r}(s) = \{\Theta_{n,r,t}(s) : t = 0, 1, \ldots, \frac{n}{m} - 1\}$ and $V_{n,r}(C_n(R)) = \{\Theta_{n,r,t}(C_n(R)) : t = 0, 1, \ldots, \frac{n}{m} - 1\}$. Define $'\circ'$ in $V_{n,r}$ such that $\Theta_{n,r,t} \circ \Theta_{n,r,t'} = \Theta_{n,r,t+t'}$, $(\Theta_{n,r,t} \circ \Theta_{n,r,t'})(x)\ (= \Theta_{n,r,t}(\Theta_{n,r,t'}(x)) = \Theta_{n,r,t}(x + jt'm) = (x + jt'm) + jtm = x + j(t + t')m) = \Theta_{n,r,t+t'}(x)$ and $\Theta_{n,r,t}(C_n(R)) \circ \Theta_{n,r,t'}(C_n(R)) = \Theta_{n,r,t+t'}(C_n(R))$ for every $\Theta_{n,r,t}, \Theta_{n,r,t'} \in V_{n,r}$ where $t + t'$ is calculated under addition modulo $\frac{n}{m}$. Clearly, $(V_{n,r}(s), \circ)$ and $(V_{n,r}(C_n(R)), \circ)$ are abelian groups for every $s \in \mathbb{Z}_n$.*

$V_{n,r}(C_n(R))$ contains all isomorphic circulant graphs of Type-2 of $C_n(R)$ w.r.t. r, if exist. Let $T2_{n,r}(C_n(R)) = \{C_n(R)\} \cup \{C_n(S) : C_n(S)$ is Type-2 isomorphic to $C_n(R)$ w.r.t. $r\}$. Thus, $T2_{n,r}(C_n(R)) = \{C_n(R)\} \cup \{\Theta_{n,r,t}(\ C_n(R)) : \Theta_{n,r,t}(C_n(R)) = C_n(S)$ and $C_n(S)$ is Type-2 isomorphic to $C_n(R)$ w.r.t. $r,\ 0 \leq t \leq \frac{n}{m} - 1\} = \{\Theta_{n,r,0}(C_n(R))\} \cup \{\Theta_{n,r,t}(C_n(R)) : \Theta_{n,r,t}(C_n(R)) = C_n(S)$ and $C_n(S)$ is Type-2 isomorphic to $C_n(R)$, $0 \leq t \leq \frac{n}{m} - 1\} \subseteq V_{n,r}(C_n(R))$ and $(T2_{n,r}(C_n(R)), \circ)$ is a subgroup of $(V_{n,r}(C_n(R)), \circ)$. Clearly, $T1_n(C_n(R)) \cap T2_{n,r}(C_n(R)) = \{C_n(R)\}$. And $C_n(R)$ has Type-2 isomorphic circulant graph w.r.t. r if and only if $T2_{n,r}(C_n(R)) \neq \{C_n(R)\}$ if and only if $T2_{n,r}(C_n(R)) \cap \{C_n(R)\} \neq \Phi$ if and only if $|T2_{n,r}(C_n(R))| > 1$ [1].

Definition 3.11. [1] *For any circulant graph $C_n(R)$, if $T2_{n,r}(C_n(R)) \neq \{C_n(R)\}$, then $(T2_{n,r}(C_n(R)), \circ)$ is called the Type-2 group of $C_n(R)$ w.r.t. r under $'\circ'$.*

Theorem 3.12. [1] *Let p be an odd prime and $k \geq 3$. Then, for $i = 1$ to p, $d_i = (i-1)np + 1$ and $R_i = \{d_i, np^2 - d_i, np^2 + d_i, 2np^2 - d_i, 2np^2 + d_i, 3np^2 - d_i,\ 3np^2 + d_i, \ldots, (p-1)np^2 - d_i, (p-1)np^2 + d_i, np^3 - d_i, pp_1, pp_2, \ldots, pp_{k-2},\ p(np^3 - p_{k-2}),\ p(np^3 - p_{k-3}),\ \ldots,\ p(np^3 - p_1)\}$, circulant graphs $C_{np^3}(R_i)$ are Type-2 isomorphic (and without CI-property) where $gcd(p_1, p_2, \ldots, p_{k-2}) = 1$ and $n, p_1, p_2, \ldots, p_{k-2} \in \mathbb{N}$.* □

Theorem 3.13. [1] *Let p be an odd prime, $k \geq 3$, $1 \leq i \leq p$, $d_i = (i-1)np + 1$, $R_i = \{d_i,\ np^2 - d_i,\ np^2 + d_i,\ 2np^2 - d_i,\ 2np^2 + d_i,\ 3np^2 - d_i,\ 3np^2 + d_i,\ \ldots,\ (p-1)np^2 - d_i,\ (p-1)np^2 + d_i,\ np^3 - d_i,\ pp_1,\ pp_2,\ \ldots,\ pp_{k-2},\ p(np^3 - p_{k-2}),\ p(np^3 - p_{k-3}),\ \ldots,\ p(np^3 - p_1)\}$, $T2(R_i) = \{\Theta_{np^3,p,jn}(R_i) : j = 1, 2, \ldots, p\}$, $T2(C_{np^3}(R_i)) = \{\Theta_{np^3,p,jn}(C_{np^3}(R_i)) : j = 1, 2, \ldots, p\}$, $gcd(p_1, p_2, \ldots, p_{k-2}) = 1$ and $n, p_1, p_2, \ldots, p_{k-2} \in \mathbb{N}$. Then $T2_{np^3,p}(R_i) = T2(R_j)$, $T2_{np^3,p}(C_{np^3}(R_i)) = T2(C_{np^3}(R_j))$ and $(V_{np^3,p}(R_i), \circ)$, $(V_{np^3,p}(C_{np^3}(R_i)), \circ)$ and $(T2_{np^3,p}(R_i), \circ)$ are abelian groups, $1 \leq i, j \leq p$. Moreover, $(T2_{np^3,p}(C_{np^3}(R_i)), \circ)$ is the Type-2 group of order p on $C_{np^3}(R_i)$ w.r.t. $r = p$, $1 \leq i, j \leq p$.* □

Circulant graphs $C_{16}(1, 2, 7)$ and $C_{16}(2, 3, 5)$ are Type-2 isomorphic and $C_{27}(1, 3, 8, 10)$, $C_{27}(3, 4, 5, 13)$ and $C_{27}(2, 3, 7, 11)$ are also. See Figures 1–5.

4 Cartesian Product and Factorization of Circulant Graphs

Just as integers can be factored into prime numbers, there are many results on decompositions of structures throughout mathematics [6]. The standard products - Cartesian, lexicographic, tensor, and strong - all belong to a class of products introduced by Imrich and Izbicki and called $B - products$ [8]. In this section, a few important results that are obtained in our study of Cartesian product and factorization of circulant graphs, similar to the theory of product and factorization of natural numbers, are presented (For details see [15]). Graphs $C_5 \square C_6$ and $C_{30}(5, 6)$ are isomorphic and are given in Figures 12 and 13. One can see the difficulty of this study from this example.

Definition 4.1. [9] *The cross product or Cartesian product of two simple graphs $G(V, E)$ and $H(W, F)$ is the simple graph $G \square H$ with vertex set $V \times W$ in which two vertices $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are adjacent if and only if either $u_1 = v_1$ and $u_2 v_2 \in F$ or $u_2 = v_2$ and $u_1 v_1 \in E$.*

Theorem 4.2. [15] *Let G be a connected graph of order n, $n > 2$. Then, $P_2 \square G$ is circulant if and only if $G \cong H$ or $P_2 \square H$ where H is a connected circulant graph of odd order.* □

Theorem 4.3. [15] *Let G be a connected graph of order $n \geq 2$. Then $C_4 \square G$ is circulant if and only if G is circulant of odd order.* □

Theorem 4.4. [15] *If G and H are connected graphs and $G \square H$ is circulant, then G and H are circulants.* □

Graphs $P_2 \square C_3$ and $P_2 \square C_4$ are given in Figures 10 and 11 and $C_5 \square C_6$ and $C_{30}(5,6) \cong C_5 \square C_6$ in 12 and 13, respectively.

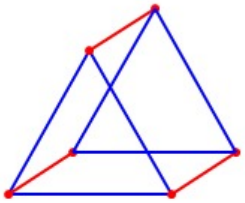


Fig.10. $P_2 \times C_3$

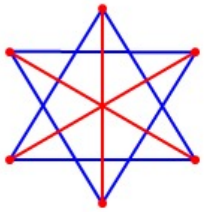


Fig.11. $C_6(2,3)$

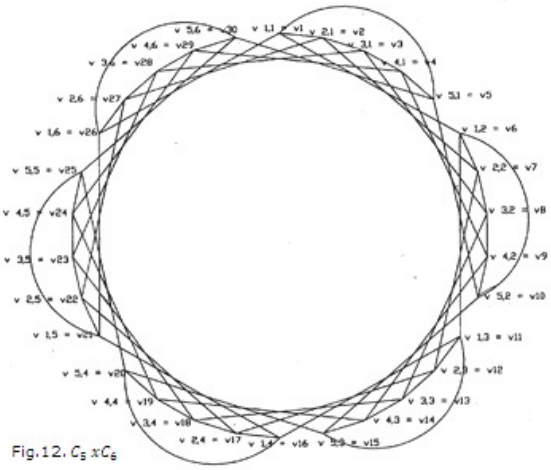


Fig.12. $C_5 \times C_6$

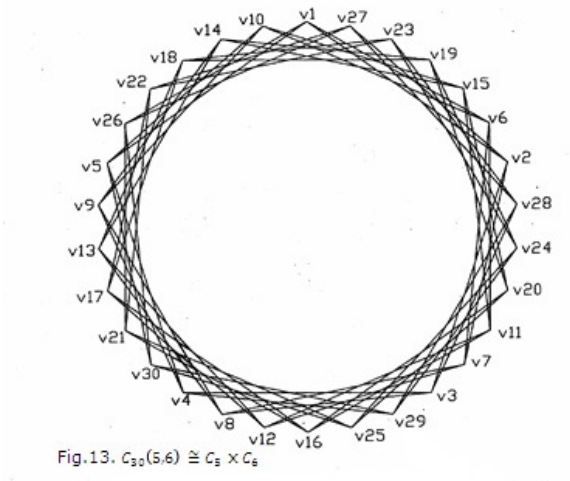


Fig.13. $C_{30}(5,6) \cong C_5 \times C_6$

Theorem 4.5. [15] *Let G and H be connected graphs, each of order > 2. Then $G \square H$ is circulant if and only if G and H are circulants and satisfy one of the following conditions:*

(i) $G \cong C_m(R)$; $H \cong C_n(S)$ *and* $gcd(m,n) = 1$.

(ii) $G \cong C_{2m+1}(R)$; $H \cong C_{2n+1}(S)$, $P_2 \square C_{2n+1}(S)$, $C_4 \square C_{2n+1}(S)$ *or* $C_{2^k(2n+1)}(S)$ *and* $gcd(2m+1, 2^k(2n+1)) = 1$, $k \in \mathbb{N}$.

(iii) $G \cong P_2 \square C_{2m+1}(R)$; $H \cong C_{2n+1}(S)$ *or* $P_2 \square C_{2n+1}(S)$ *and* $gcd(2m+1, 2n+1) = 1$.

(iv) $G \cong C_{2^k(2m+1)}(R) \neq P_2 \square C_{2^{k-1}(2m+1)}(T)$ *for any* $C_{2^{k-1}(2m+1)}(T)$; $H \cong C_{2n+1}(S)$ *and* $gcd(2^k(2m+1), 2n+1) = 1$, $k \in \mathbb{N}$.

(v) $G \cong C_4 \square C_{2m+1}(R)$; $H \cong C_{2n+1}(S)$ *and* $gcd(2m+1, 2n+1) = 1$.

(vi) $G \cong C_{2^k(2m+1)}(R) \neq C_4 \square C_{2^{k-2}(2m+1}(T)$, $P_2 \square C_{2^{k-1}(2m+1)}(U)$ *for any* $C_{2^{k-2}(2m+1}(T)$ *and* $C_{2^{k-1}(2m+1)}(U)$; $H \cong C_{2n+1}(S)$ *and* $gcd(2^k(2m+1), 2n+1) = 1$, $k \in \mathbb{N}$, $k \geq 2$. □

Definition 4.6. [15] *A non-trivial graph G is said to be prime if $G = G_1 \square G_2$ implies G_1 or G_2 is trivial; G is composite if it is not prime.*

Definition 4.7. [15] *If $C_m(R)$, $C_n(S)$ and $C_{mn}(T)$ are circulant graphs such that $C_m(R) \square C_n(S) \cong C_{mn}(T)$, then we say that $C_m(R)$ and $C_n(S)$ are divisors or factors of $C_{mn}(T)$.*

Thus for any connected circulant graph, the graph and $C_1(\,) = K_1$ are always divisors and so we call them as *improper divisors* of the circulant graph. Divisors which are integer multiple of improper divisors also be called as improper divisors of the circulant graph. This doesn't arise since we consider divisors of connected graphs only. Divisor(s) other than improper divisors is called *proper divisor(s)* of the circulant graph.

Definition 4.8. [15] *A circulant graph whose only divisors are improper is called a prime circulant graph. Other circulant graphs are called composite circulant graphs.*

Theorem 4.9. [15] ***[Factorization Theorem On Circulant Graphs]***
Let m and n be relatively prime integers. If $R \subseteq [1, \frac{m}{2}]$, $S \subseteq [1, \frac{n}{2}]$ and $T \subseteq [1, \frac{mn}{2}]$ with $T = dnR \cup dmS$ for some d such that $gcd(mn, d) = 1$, then $C_{mn}(T) \cong C_m(R) \square C_n(S)$. □

Theorem 4.10. [15] *If $n \neq 4$ and $1 \in R$, then $C_n(R)$ is a prime circulant.* □

Corollary 4.11. [15] *If $n \neq 4$ and R contains an integer relatively prime to n, then $C_n(R)$ is prime circulant.* □

Corollary 4.12. [15] *If n is a prime power other than 4 and $C_n(R)$ is connected, then $C_n(R)$ is prime circulant for all $R \neq \phi$.* □

Theorem 4.13. [15] ***[Fundamental Theorem of Circulant Graphs]***
Every connected circulant graph is the unique product of prime circulant graphs (uniqueness up to isomorphism). □

Remark 4.14. [13, 15] *If G is a connected graph such that $G \cong G_1 \square G_2 \square \ldots \square G_k$, then the diameter of G, $dia(G) = \sum_{i=1}^{k} dia(G_i)$.*

Thus, we can find the diameter of any given circulant graph, provided diameters of its prime circulant graphs are known. Also the above relation helps to generate (circulant) graphs of bigger diameters.

Remark 4.15. [15]

1. *In prime factorization of connected circulants* $C_1(\ ) = K_1$ *and* $C_2 = P_2$ *act similar to* 1 *and* 2 *among the set of all natural numbers, respectively. Thus,* $C_1(\ )$ *is a unit, like* 1 *in number theory.*
2. *There exist two types of prime circulant graphs of order n, one with periodic cycle(s) of length n and the other without periodic cycle of length n.*
3. *The theory of factorization of circulants is similar to the theory of factorization of natural numbers and one of the very few well-known mathematical structures so vividly classified (expressed) in terms of prime factors. It can be applied in cryptography.*
4. *We developed VB programs POLY215.exe and POLY315.exe to show visually how the transformations* $\Theta_{n,r,t}$ *and* $\Phi_{m,n}$ *act on* C_n *and* $C_m \square C_n$*, respectively for different values of m and n,* $m, n \in \mathbb{N}$.
5. *An interesting problem is, for a given integer n, finding the number of prime (composite) circulant graphs of order either equal to n or less than or equal to n.*
6. *One can develop theories similar to the theory of Cartesian product and factorization of circulant graphs to the other standard products of circulant graphs.*

Conclusion This study covers a few isomorphic properties of circulant graphs. One can go for a similar study on Cayley graphs.

Acknowledgment The author expresses his sincere thanks to Prof. Lowell W Beineke, Indiana-Purdue University, U.S.A., Prof. Brian Alspach, University of Newcastle, Australia, Prof. M.I. Jinnah, University of Kerala, Trivandrum, India and Prof. V. Mohan, Thiyagarayar College of Engineering, Madurai, Tamil Nadu, India for their valuable suggestions and guidance and Dr. A. Christopher, Dr. P. Wilson and Mr. R. Satheesh of S.T. Hindu College, Nagercoil, India for their assistance to develop the VB programs to show how the transformations defined on circulant graphs to obtain their isomorphic graphs is taking place. He also expresses his gratitude to Central University of Kerala (CU Kerala), Periye - 671 316, Kasaragod, Kerala, India and St. Jude's College, Thoothoor - 629 176, Kanyakumari District, Tamil Nadu, India for providing facilities and CU Kerala and Lerroy Wilson Foundation, India (www.WillFoundation.co.in) for providing financial assistance to do this research work.

References

1. V. Vilfred, "A few properties of circulant graphs: Self-complementary, isomorphism, Cartesian product and factorization" in 7^{th} International Conference on Modeling, Simulation and Applied Optimization (ICMSAO2017), American University of Sharjah, Sharjah, April 2017.
2. P. J. Davis, Circulant Matrices, Wiley, New York, 1979.
3. A. Adam, "Research problem 2-10", J. Combinatorial Theory, **3**, 393, 1967.
4. B. Alspach, J. Morris and V. Vilfred, "Self-complementary circulant graphs", Ars Com., **53**, 187–191, 1999.
5. B. Alspach and T. B. Parsons, "Isomorphism of circulant graphs and digraphs", Discrete Math., **25**, 97–108, 1979.
6. R. Arratia, A. D. Barbour and Simon Tavare, "Random Combinatorial Structures and Prime Factorizations", A.M.S.Notices, **44**, 903–910, 1997.
7. F. T. Boesch and R. Tindell, "Circulant and their connectivities", J. Graph Theory, **8**, 487–499, 1984.
8. B. Elspas and J. Turner, "Graphs with circulant adjacency matrices", J. Combinatorial Theory, **9**, 297–307, 1970.
9. F. Harary, Graph Theory, Addison-Wesley, Reading, M.A., 1969.
10. W. Imrich and H. Izbicki, "Associative products of graphs", Monatsh. Math., **80**, 277–281, 1975.
11. I. Kra and S. Simanca, "On Circulant Matrices", AMS Notices, **59** (3), 368–377, 2012.
12. C. H. Li, "On isomorphisms of finite Cayley graphs-a survey", Discrete Math. **256**, 301–334, 2002.
13. G. Sabidussi, "Graph multiplication", Math. Z., **72**, 446–457, 1960.
14. H. Sachs, "Uber selbstkomplementare Graphen", Publ. Math. Debrecen, **9**, 270–288, 1962.
15. V. Vilfred, "A Theory of Cartesian Product and Factorization of Circulant Graphs", Hindawi Pub. Corp. - J. Discrete Math., Vol. 2013, Article ID 163740, 10 pages.
16. V. Vilfred, "New Abelian Groups from Isomorphism of Circulant Graphs", Proce. of Inter. Conf. on Applied Math. and Theoretical Computer Science, St.Xaviers Catholic Engineering College, Nagercoil, Tamil Nadu, India, xiii–xvi, 2013.
17. V. Vilfred, "$\sum$-labelled Graphs and Circulant Graphs", Ph.D. Thesis, University of Kerala, Trivandrum, India, 1994.
18. V. Vilfred and P. Wilson, "New Families of Circulant graphs without Cayley Isomorphism Property with $m_i = 5$", Int. Journal of Scientific and Innovative Mathematical Research, **3** (6), 39–47, 2015.
19. V. Vilfred and P. Wilson, "New Family of Circulant Graphs without Cayley Isomorphism Property with $m_i = 7$", IOSR Journal of Mathematics, **12**, 32–37, 2016.
20. H. Zhang, "Self-complementary symmetric graphs", J. Graph Theory, **16**, 1–5, 1992.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 242-251 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Efficiency Assessment of Plasma-Aided Porous Media Surface Finishing

Peter Dineff*, Dilyana Gospodinova, Ivaylo Ivanov

Technical university of Sofia, Faculty of electrical engineering, 1756, Bulgaria

ARTICLE INFO

Article history:
Received: 22 October, 2017
Accepted: 11 November, 2017
Online: 23 December, 2017

Keywords :
Processing efficiency assessment
Plasma-aided surface finishing
Porous media

ABSTRACT

The efficiency of a plasma-aided finishing process, namely capillary impregnation, can be predicted only by setting two basic parameters of the real wood porous surface and the actual impregnating solution – surface free energy and surface tension. In general, the following processing efficiency parameter was found and the rule is true: "The plasma aided or enhanced finishing of a porous media will be more successful and this media will be more susceptible to it as the difference between its surface free energy and the surface tension or the so-called penetration-spreading parameter is positive: PSP = $(\sigma_S - \gamma_L) > 0$. If not, wetting, wicking and finishing problems will occur.

1. Introduction

This paper is an extension of a work originally presented at the International conference on electrical machines, drives and power systems "*ELMA'17*" [1].

Dielectric-barrier discharges (DBD) have been known for more than a century – the first experimental investigations focused on ozone generation were reported by *Werner von Siemens* in 1857. In 1951, *Werner Eisby* founded his company *Vetaphone* on the patent of a new plasma surface technology, which he initially named *E-treatment* and later it was known worldwide as the *Corona technology*. In fact, *Corona* was a *Dielectric-barrier discharge* in air at atmospheric pressure. This high-frequency electric discharge was controlled by a dielectric barrier and organized in the *micro-discharge mode* – it created a pattern of streamers distributed over the whole surface. This *atmospheric pressure air-plasma* (APP) *pre-treatment* was necessary to obtain sufficient wetting and adhesion on plastic films or metallic foils before high-speed roll-to-roll converting processes to ensure perfect surface finishing - printing, laminating or coating. Adding a *Corona treater* on a blown film or cast film extrusion line was essential in order to obtain a perfect *surface free energy* (SFE) prior to the following converting process. This *plasma-aided finishing* has proved to be both highly effective, cost-effective and can take place in-line [2, 3].

In the 80s of the last century, the *APP pre-activation* came to facilitate the textile finishing. In textile manufacturing finishing refers to the processes, that convert the woven, non-woven or knitted cloth into a usable material or fabric to improve the aesthetic appearance and the functional characteristics of the finish textile. This *plasma-aided finishing* has proved to be both highly effective, cost-effective and can take place not only in-line, but also out-of-the-line. The finish textiles such as carpets, rugs, mats, curtains or clothing require the implementation of plasma-aided finishing *out of the production line* [3, 4].

At the beginning of this century, the *APP-activation* (APPA) was used to enhance the wood finishing. In wood manufacturing finishing refers to the process of refining or protecting wooden surfaces. Wood finishing is the final step of the manufacturing process that gives wood surfaces desirable characteristics, including enhanced appearance and increased resistance to moisture, to fire and to attack from destructive organisms such as fungi and insects. It is well known that the inactivated by hydrophobic contamination wood surface has low *SFE*. This *wood inactivation*, [5], leads to poor wetting, adhesion, and impregnation especially for water-based varnishes and printing inks, adhesives and coating substances, as well as flame retardant water solution. The plasma-aided finishing has proved to be both highly effective, cost-effective, [6, 7], and can take place not only in-line, but mostly out-of-the-line. The three-dimensional wooden buildings and structures require removal of the plasma-aided finishing "*in situ*" out of the production line [8].

The *APPA* is essential in surface finishing processes. Without plasma-chemical surface pre-treatment, the porous and non-porous substrates will not allow further finishing processes due to the low *SFE*. The *plasma surface activation* and *functionalization* are able

* Corresponding Author: Peter Dineff, Technical University - Sofia, Faculty of electrical engineering, 1756, Bulgaria, dineff_pd@abv.bg

to ensure that perfect finishing result we all want to see when looking at finish products. The objective of this kind of activation is to increase the *SFE* and subsequently the wettability, adhesion and wicking for laminating, printing, dyeing, gluing, impregnating and other finishing processes [9, 10].

Established *APPA* processes, such as plasma or flame activation, increase the *SFE* and decrease the contact angle, and therefore improve the processing characteristics. The *Sessile Drop Test* - a contact angle measurement technique, and the *Dine Test* - an ink test technique, were two widely used methods for *SFE* assessment and therefore for checking the *need for APPA* or the efficiency of such a pre-treatment processes [11, 12, 13].

The *SFE* of a solid cannot be measured directly because solids typically show no reaction to the exertion of *SFE*. Plasma, corona or flame treated and untreated surfaces are indistinguishable to the naked eye. Consequently, practical measurements of *SFE* involve the interaction of the solid with a test liquid to determine wetting surface tension as a measure of *SFE*. Standards such as DIN EN 828: 2013-04 (E) for adhesive finishing specify contact angle measurement as a method for determining the *SFE* of surfaces for checking the efficiency of cleaning or activating plasma pre-treatment processes.

The purpose of this paper is not only to collect new evidence and extend knowledge about the plasma-aided porous media finishing, especially *capillary impregnation,* with phosphorous flame retardant (FR) containing water solution for *flame retardancy* as finishing for wood, wooden and cellulosic materials, but also to provide a process efficiency assessment for the overall plasma enhanced process, not only for the *APPA pre-treatment*.

2. Assessment of plasma-activated finishing efficiency

2.1. Model of wetting-wicking theory and efficiency assessment

Plasma-aided technique was used to facilitate the surface finishing, more precisely the *surface impregnation* of porous media such as wood, wooden and cellulosic materials [1, 14-17].

Porous materials with low *SFE* (σ_S) were subjected to plasma-chemical surface pre-activation for increasing their *SFE* and enhance the finishing. The wetting phenomenon occurs in a different way on porous media. The surface curvature of a sessile liquid drop on a porous solid depends not only on wetting but also on wicking in depth. This is a phenomenon of wetting-wicking. Wicking and wetting are essential in porous media finishing, such as capillary impregnation for flame retardancy of wood, wooden materials,woven and non-woven textile Figure 1, [1].

Both wetting and wicking together determine the variation of the sessile drop curvature and the contact angle over time. This is the wetting-wicking and contact angle phenomenon. That is why it is imperative in the case of porous media to talk about *apparent contact angle* and *apparent surface free energy.* This process of contact angle change ends with the complete absorption of the flame retardant solution at almost constant value of the contact angle - *end contact angle*, much lower than the *initial contact angle,* Figure 2.

The "*wetting-wicking theory*", expressed in terms of the thermodynamic parameters, such as the *contact angle*, the *surface tension*, and the *surface free energy*, can be the most widely used approach to porous media in wetting science at future, also referred here as "*wetting-wicking phenomena*". Wetting-wicking is defined here as the ratio between the surface energies of the solid and the surface tension of the liquid [14].

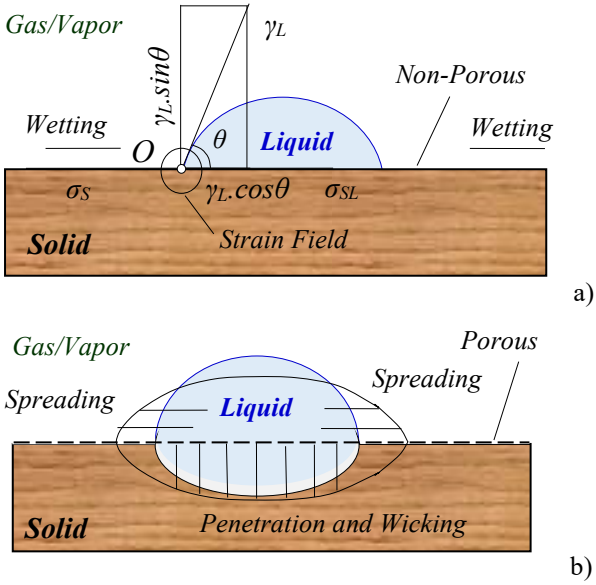


Fig. 1. Schematic illustration of two types of *wetting phenomena*: **a)** the *wetting phenomenon* on a *non-porous surface* when a liquid drop is placed on a smooth, non-porous and rigid solid, both exposed to gas/vapor, the system will not be in equilibrium and the liquid "wets out" or spreads over the solid, then the liquid exhibits a contact angle of zero against the solid (*Good,* 1993); **b)** *the wetting-wicking* or shortly *wicking phenomenon* on a *porous surface* can involve: *i* - spreading over the solid; *ii* - penetration or wicking into the porous solid (*Berg*, 1993).

According to *Young's equation*, which is proven since over 200 years (1805), there is a fundamental relationship between the *contact angle* θ, deg, the *surface free energy* (SFE) of the solid σ_S, mJ/m^2, the *surface tension* (SFT) of the liquid γ_L, mN/m, and the *interfacial tension* (IFT) σ_{SL} between liquid and non-porous or porous solid, mJ/m^2, Figure 1a, [1, 15]:

$$\sigma_S - \sigma_{SL} - \gamma_L \cos\theta = 0; \quad (1)$$

$$\Sigma = \gamma_L \cos\theta = \sigma_S - \sigma_{SL}, \quad (2)$$

where Σ is the *adhesion tension*, mJ/m^2; the adhesion tension reaches its maximum value at full wetting surface when the *IFT* reaches its minimum value: $\cos\theta = 1$ ($\theta = 0°$); $\sigma_{SL} = \sigma_S - \gamma_L$; and $\Sigma = \gamma_L$.

The liquid in porous materials is often pulled into surface pores and capillaries by *capillary action*. The capillary action or *capillarity* is defined as the movement of liquid within the spaces of a porous media due to the forces of adhesion, cohesion and surface tension. Capillarity manifests itself when the adhesion to the capillary walls is stronger than the cohesive forces between the liquid molecules. Its ability to fill-in defects, such as surface-breaking cracks, voids, pores and capillaries can be imputed to a great extent to penetration, sorption or wicking characteristics. In porous media, capillary pressure p_C is the force necessary to squeeze a droplet through a pore throat and to work against the *IFT* between solid and liquid phases. Capillary pressure p_C is proportional to the adhesion tension Σ and depends only on it at a certain characteristic pore size r [16]:

$$p_C = 2\,(\sigma_S - \sigma_{SL})/r = 2\,\Sigma/r, \quad (3)$$

where the capillary pressure p_C reaches its maximum value when the adhesion tension Σ reaches its maximum value at complete wetting: $cos\ \theta = 1$ ($\theta = 0°$); $\Sigma = \gamma_L$; and $p_C = 2\ \gamma_L/r$.

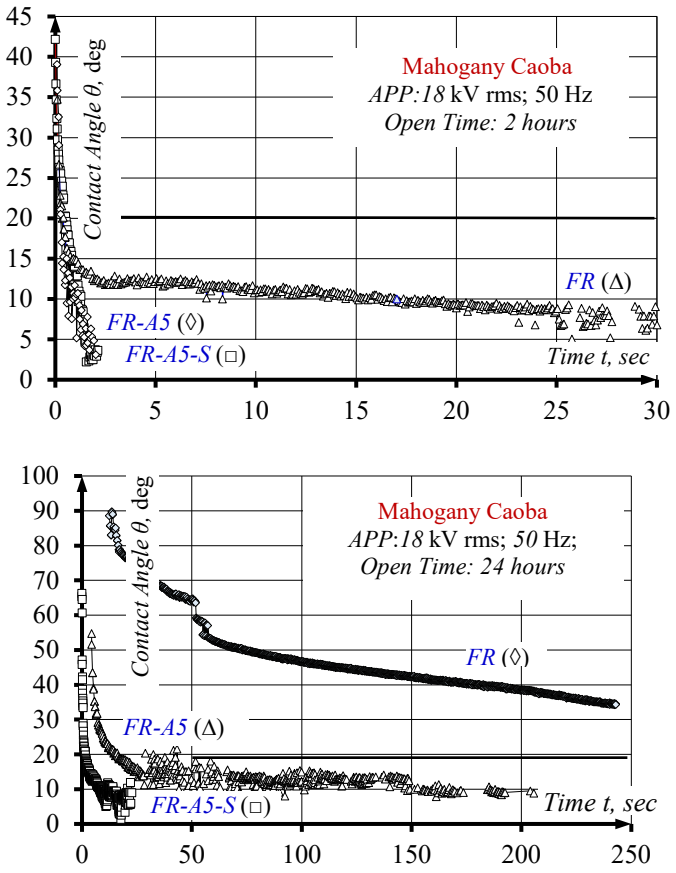


Fig. 2. Dynamics of change in *apparent contact angle* θ over time t depending on the used impregnating water solutions of phosphorous flame retardant with different surface tension: *FR* – 46.4 mN/m; *FR-A5* – 22.4 mN/m and *FR-A5-S* – 18.5 mN/m, and short and long aging time, 2 or 24 hours.

Sessile and pendant drop analyses allow *SFE* and *SFT* to be quantified and to propose two distinct relationships that illustrate in general this process in the wetting area, $\theta \in [0 \div 90]°$, – the *IFE* and adhesion tension Σ - proportional to capillary tension p_C, as a functions of the contact angle θ, Figure 3.

2.2. The best solution rule for finishing and plasma-aided finishing efficiency parameter

Both *SFE* and *SFT* are essential in porous media finishing phenomena, (1). *APPA* surface technique can be used to enhance surface finishing: materials with low *SFE* were subjected to oxidative *APPA* to increase it, Figure 3.

The difference between *SFE* of porous surface and *SFT* of impregnating solution was chosen as a finishing efficiency or *penetration-spreading parameter* (PSP) for a real porous solid such as wood: $PSP = (\sigma_S - \gamma_L)$. *Finishing* does not include evaporation of solution, dissolution or swelling of the solid by the water solution or any kind of chemical reaction between the solution and the solid substrate that changes the system composition. The higher the *SFE* of the solid, the better the finishing; the lower the *SFT*, the better the result of finishing. To sum up, large and positive difference between *SFE* and *SFT*: $\sigma_S - \gamma_L > 0$, (or $\sigma_S > \gamma_L$), will provide more effective porous surface finishing [1].

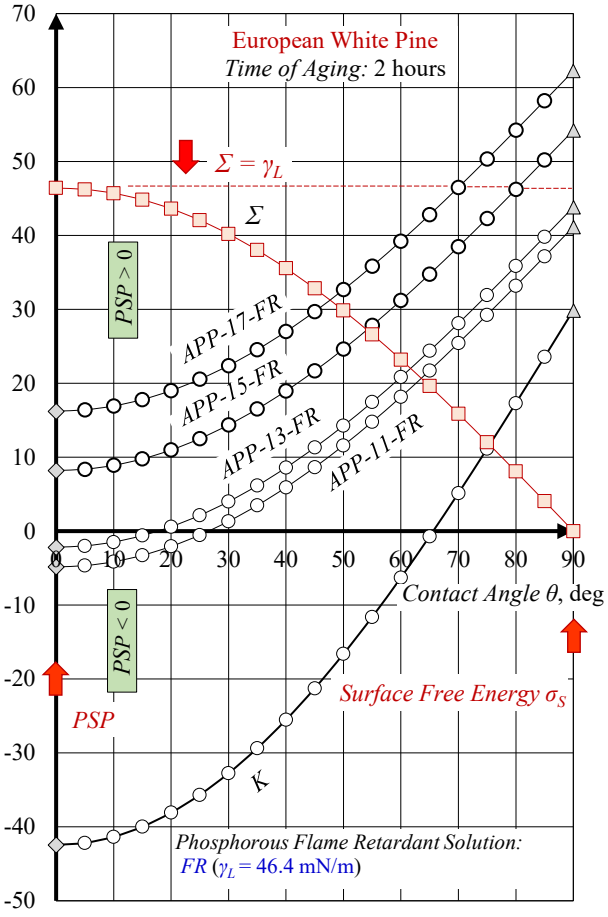


Fig. 3. Relationship between the "solid-liquid" interfacial energy σ_{SL} and the contact angle θ, measured by *Sessile drop technique*. Penetration-spreading parameter (*PSP*) is equal to the minimum *IFE* value at $cos\ \theta = 1$ ($\theta = 0°$): $PSP = \min \sigma_{SL}$. An effective finishing has a positive value of *PSP*: $PSP > 0$; a non-effective finishing has a value of less than or equal to zero: $PSP \leq 0$. Surface free energy σ_S is equal to maximum *IFE* value at $cos\ \theta = 0$ ($\theta = 90°$): $\sigma_S = \max \sigma_{SL}$.

Relationship between the adhesion tension Σ and the contact angle θ. The maximum adhesion tension Σ value is equal to the surface tension γ_L of the flame retardant water solution *FR*: $\gamma_L = 46.4$ mJ/m^2.

Atmospheric pressure plasma pre-activation at 50/60 Hz and different voltage: 11, 13, 15 and 17 kV rms: method – in situ; time of activation – 60 sec; glass barrier – 3 mm; air gap – 6 mm.

In general, the following processing efficiency parameter was found and the rule is true: "*The plasma aided or enhanced finishing of a porous media will be more successful and this media will be more susceptible to it as the difference between its SFE* σ_S *and the SFT* γ_L *or the so-called penetration-spreading parameter is positive*: $PSP = (\sigma_S - \gamma_L) > 0$. *If not, wetting, wicking and finishing problems will occur*" [1].

This new rule of processing efficiency is fully consistent with the well-known rule formulation: "*If the liquid has a dyne level* (*SFT*) *lower than the substrate dyne level* (*SFE*), $\gamma_L < \sigma_S$, *then the liquid will spread out over its entire surface in uniform wet layer. If the liquid dyne level is equal to or higher than a substrate dyne level*, $\gamma_L \geq \sigma_S$, *the liquid will become cohesive and tend to remain in droplets*" [18]:

$$\sigma_S > \gamma_L: \sigma_S - \gamma_L > 0 \text{ or } PSP > 0. \qquad (4)$$

According to *Young's equation*, there is a fundamental relationship between *IFE* and $PSP = (\sigma_S - \gamma_L)$:

$$\sigma_{SL} = (\sigma_S - \gamma_L) + \gamma_L (1 - \cos\theta) = PSP + \gamma_L - \gamma_L \cos\theta; \quad (5)$$

$$\min \sigma_{SL} = (\sigma_S - \gamma_L) = PSP;\ \theta = 0°: \cos\theta = 1, \quad (6)$$

or the minimum value of the "*solid-liquid*" *IFE* at $\theta = 0°$ is equal to the value of *PSP*: $\sigma_{SL} = PSP$, Figure 3, [1, 14].

Table 1. Contact angle, wetting and hydrophilicity change.

Contact angle θ, °	**0**	10	20	30	70	**90**
cos θ	1	$1 > \cos\theta > 0$				0
Interface free energy (*IFE*)	$\sigma_{SL} = PSP$					$\sigma_{SL} = \sigma_S$
Hydrophilicity / Hydrophobicity	Hydrophilicity				Hydro phobicity	
	Super-Hydrophilicity		Hydro philicity		Hydro phobicity	
Degree of wetting	*Spreading*	*Complete wetting*	*Wetting*	*Partial wetting*		*Incomplete wetting*
Wettability	*High wettability*					

The maximum value of the *IFE*, at $\theta = 90°$, is equal to the value of porous wood *SFE*: $\sigma_{SL} = \sigma_S$, Figures 3 and 4.

Therefore, the requirement for process efficiency finishing $PSP = (\sigma_S - \gamma_L) > 0$ can be transformed into a new requirement for effectiveness, Figure 3:

$$\sigma_S > \sigma_{SL} > PSP > 0, \quad (7)$$

which already refers directly to *IFE*: "*it is necessary for the entire range of variation of the static contact angle* ($0° \le \theta \le 90°$) in the area of the *high wetting*, *the value of the IFE to be greater than zero*: $\sigma_{SL} > 0$".

Plasma-aided capillary impregnation finishing will be more successful and the porous media will be more susceptible when the $PSP > 0$. This is the essence of the rule for obtaining effective plasma-enhanced capillary impregnation finishing.

Surface activation refers to the temporal increase of the *SFE* determined by the rise of the polar part of the *SFE*, Figure 4. This *APPA* process is based on the implantation of oxygen, leading to the formation of functional groups like *ether* (C-O), *hydroxyl* (C-OH), *carbonyl* (C=O), and *carboxyl* (O=C–OH). Such a plasma treatment enhances the affinity of the substrate for other substances and is especially needed for solids which typically have a low intrinsic *SFE*. This process is generally achieved using a standard *DBD* in open air and "in situ" or remote application Figure 5.

The *PS*-parameter depends only on two thermodynamic parameters of the finishing process – the *SFE* σ_S of the solid and the *SFT* γ_L of the liquid, which can be quantified individually using two well-known methods of contact angle measurement – the sessile drop and the pendant drop techniques [15, 16 and 17].

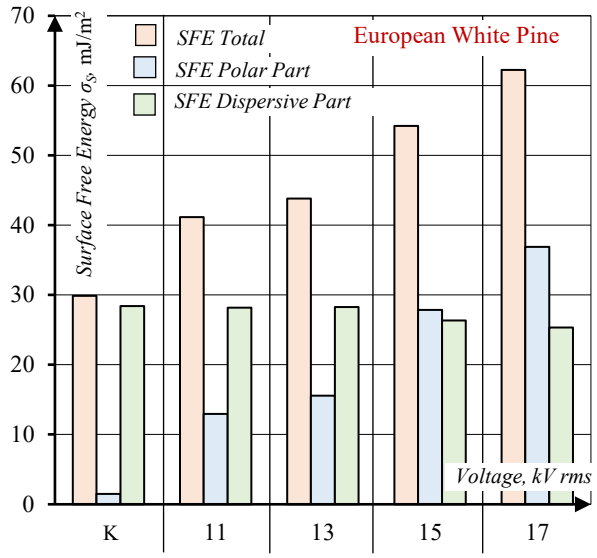


Fig. 4. Surface free energy σ_S and its dispersive and polar parts for *European white pine* after *APPA* and 2 hours of aging time and varied voltage: 11,13,15 and 17 kV rms, defined by Sessile drop technique and *theory of Wu*.

Working with the graphs and studying the relationship between the "*solid-liquid*" *IFE* σ_{SL} and *contact angle* θ in the area of high wettability, $0° \le \theta \le 90°$, it becomes clear that the graphs of the efficiency realizations of plasma-aided surface impregnation – *APP-15-FR* and *APP-17-FR* are located in the first quadrant above the abscissa, Figure 3.

One more graphic illustration has been brought about, Figure 5, depicting the efficiency parameter assessment of plasma-aided capillary impregnation with modified phosphorous *FR* solution by increasing the *SFE* and decreasing the *SFT* of the impregnating solution.

2.3. Aging, plasma activation durability and open time

It is very important to realize that the changes induced on the surface by an *APPA* are not permanent. There is aging, which is the effect of the *APPA* or *SFE* decreasing in time. Often, the activation effect will remain partially and a steady state will be reached after some hours, days or weeks. For *APPA*, the aging can be really an issue. Hence the typical *APPA* was performed "in-line", immediately before the finishing for which it is intended. Aging was usually regarded as an unwanted side-effect in plasma-aided finishing [17, 19].

The time of aging or *open time* is the period of time after an *APPA* has been applied, allowing for the finishing processes to take place – printing, bonding, painting, laminating and impregnating, during which an effective plasma-aided finishing can be achieved. Plasma pre-activation most often takes place "in-line" immediately prior to the onset of the finishing. The *time of aging* in this case is zero. The *time of aging* is different from zero when the plasma-aided capillary impregnation is applied "out-of-the-line", for example in the finishing of a wooden building construction, wood paneling, flooring or staircase. Most often, the entire surface is activated "*step by step*" for a certain time and then it is proceeded to capillary impregnation, Figure 6.

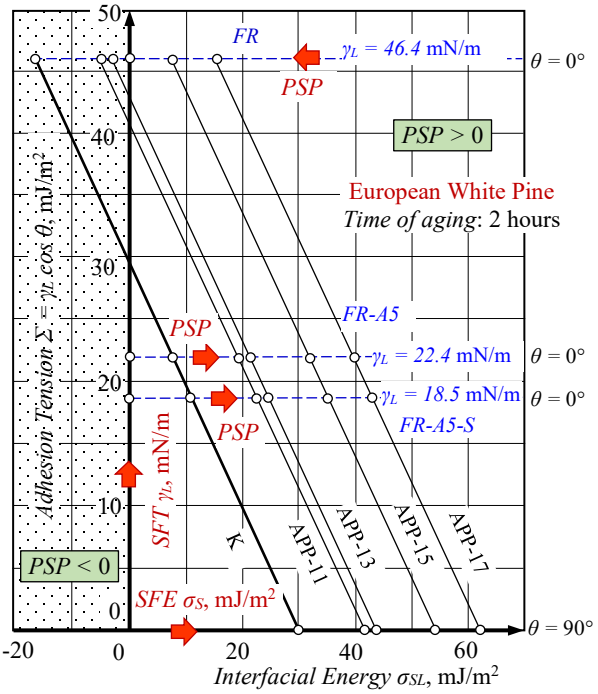


Fig. 5. Relationship between the adhesion tension Σ and the "solid-liquid" interfacial energy σ_{SL} that reveals at once the impact of: *i* – the variation of *SFE* as result of *APPA* and 2 hours of aging at varied voltage: 11, 13, 15 and 17 kV rms; *ii* – the variation of *SFT* as a result of *FR*-solution modification with surfactants: *FR-A5* – with micelle-forming anionic surfactant at 5 vol. % and *FR-A5-S* – with trisiloxane surfactant at 0.1 vol. %.

The industrialization of the plasma-aided finishing "out-of-the-line" during 2007÷09 allowed the empirical definition of two very important values for the time of aging – 2 hours or a short time of aging, allowing for successful finishing of large-scale wood wares such as paneling, flooring, wooden construction, and a day or 24 hours, a long time of aging, which is close to the *critical open time* when the effect of the plasma activation will be obliterated.

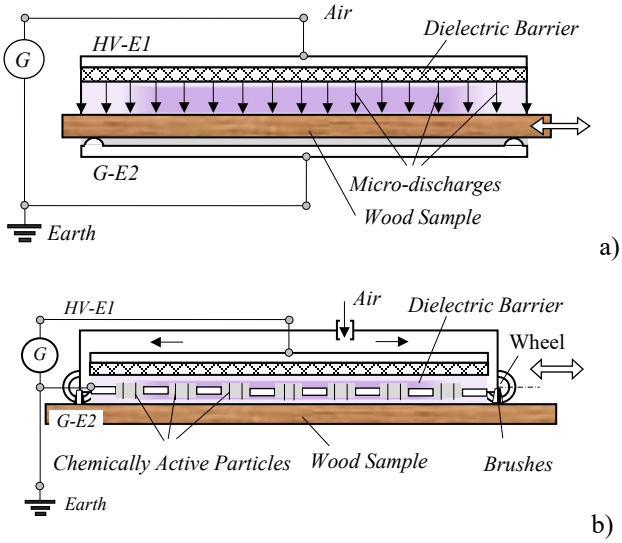


Fig. 6. Atmospheric pressure plasma pre-activation by dielectric barrier discharge in air: **a)** *APPA*, "in situ" activation where the sample is immersed in the cold plasma of the *DBD* (*V. Eisby*, 1951); **b)** *RAPPA*, remote activation where the sample interacts only with the chemically active particles produced by the *DBD* (*P. Dineff*, 2005).

HV-E1 – high-voltage electrode; *G-E2* – grounded electrode; *G* – AC electric generator (50/60 Hz).

Remember that the *SFE* level decay is extremely fast right after *APP*-activation "*in line*" – in fact an immediate loss of 10 mJ/m² is possible. This is due to the contact with process rolls, surface migration of additives and interfacial transfers between treated and untreated surfaces within the finished wound roll. Generally, the *SFE* of a film should be between 3 and 10 mJ/m² greater than the *SFT* of the ink or impregnating liquid in order to ensure acceptable performance, Figure 6, [11, 18].

The highest plasma-activation effect to the finishing, expressed by maximum *SFE* value increase, was observed with "*in situ*" *APPA*, "in line" application, Figure 6, [1, 18].

THE OBJECTIVE of this paper is to study the efficiency assessment of plasma-aided capillary impregnation as finishing for flame retardancy of wood at industrial frequency (50 Hz) and voltage of 15/18 kV rms, after short and long aging time, 2 and 24 hours, and the change of the wood impregnability monitored by the penetration-spreading parameter (PSP) and the processing efficiency rule: *PSP* > 0.

3. Experimental investigation

The development of a plasma-enhanced capillary impregnation finishing for wood flame retardancy goes through three stages:

- *The first stage* includes the selection of an effective *APPA* mode that provides the necessary high level of *SFE* after a short aging of 2 hours, Figures 3 and 9;
- *The second stage* includes the selection of an anionic micelle-forming water surfactant in a concentration to ensure the required low level of the *SFT*: 30 mN/m > γ_L > 20 mN/m, Figure 10;
- *The third stage* includes the selection of an effective *APPA* mode that will provide the necessary high level of *SFE* after a long time of aging, a day or 24 hours, and low level of *SFT,* by some adequate surfactants and concentrations, Figure 11.

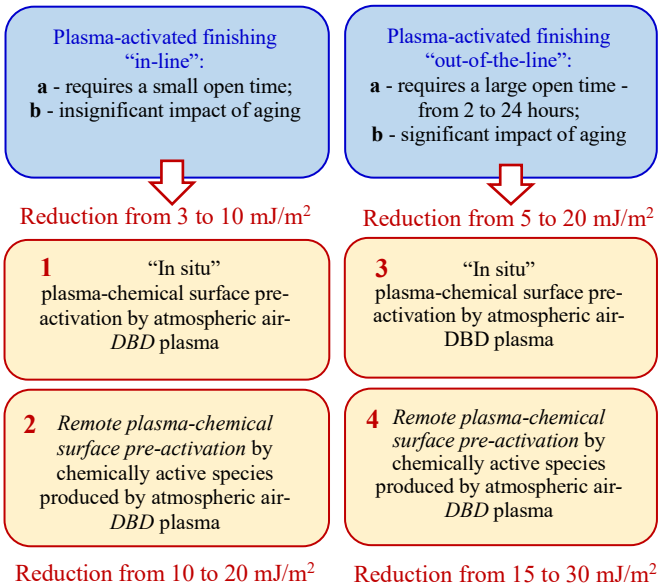


Fig. 7. Surface free energy reduction before finishing due to: *i* - contact with other non-activated surfaces, *ii* – aging after plasma activation, and *iii* – different modes of plasma surface activation provide different activation options: *APPA*, "in situ" activation immerced in cold plasma or *RAPPA*, remote activation with chemically active particles.

This experimental study was carried out on four kinds of wood samples: *European white pine* (*Pinus Sylvestris L.*) from Bulgaria and three woods from Mexican rain forest: *Tzalam* (*Lysiloma Bahamensis L.*), *Mexican white cedar* (*Cupressus lusitaca L.*), *Mahogany caoba* (*Swietenia mahogany L.*), according to the well-known method [1, 15-17]. The selected wood species form the following range of decreasing density, in kg/m^3: *Tzalam* – 780; *Mahogany caoba* – 650; *Mexican white cedar* – 470; *European white pine* – 371. There are two hard wood species with high density, $d > 540$ kg/m^3: *Tzalam* and *Mahogany caoba*, and two soft wood species with low density, $d < 540$ kg/m^3: *Mexican white cedar* and *European white pine*.

On the basis of our own former experience in plasma-aided capillary impregnation of wood and wooden materials, [9-12], an oxidative *APPA* has been applied on the test samples for 60 sec. This study employs a well-known experimental methodology, already used by us [1, 14-17].

All experimental studies pertain to *APPA* of the wood specimens in a specific burning mode of the air *DBD* at atmospheric pressure (~ 101 kPa) at the first peak of surface density of real power p_a = 26 W/m^2 and the second peak p_a = 8 W/m^2 in asymmetric rectangular coplanar electrode system with 6 mm air gap – mode *A* and *B* respectively, Figure 7.

The *APPA* was carried out with "*in situ*" application of *DBD* for 60 sec, Figure 5a. The effective operating modes of *DBD* at voltages above 13 kV rms (18.3 kV peak value) determine a surface density of real power above 4 W/m^2, Figure 7.

The aim of this study was to verify the possibility to evaluate the *PS*- efficiency parameter values for plasma and surfactant enhanced capillary impregnation trough calculation of *SFE* and *SFT* by drop shape analysis. *Drop shape analysis* is an image analysis method for determining the contact angle from the shadow image of a sessile drop and the *SFT* from the shadow image of a pendant drop. There are number of theories and models for *SFE* calculation with the help of contact angle data. The method according to *Wu* for calculating *SFE* and its two parts – polar and disperse, is suitable for high polar or plasma activated surfaces, Figure 4.

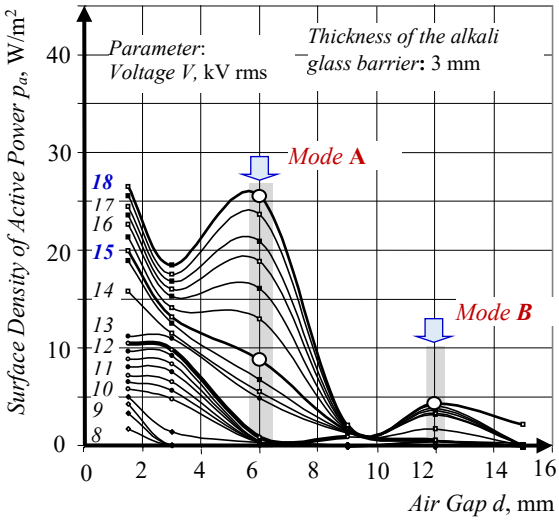


Fig. 8. Technological discharge characteristic "surface density of active power p_a – air gap size *d*" of non-equilibrium *DBD* burning at industrial frequency (50 Hz) in asymmetric rectangular coplanar electrode system with one glass barrier.

All methods described are included in the KRÜSS Drop shape analysis programs. *Drop Shape Analyzer* (DSA100 and DSA30, KRÜSS GmbH, Germany) was used for the precise measurement of contact angle *SFE* and also for measurement of the *SFT* of liquids using the pendant drop method. Measuring range (referred to image analysis) for *Sessile drop technique*: contact angle – 1°÷180°; surface free energy – 0.01÷1 000 mJ/m^2; and for *Pendant drop technique*: surface tension – 0.01÷2000 mN/m. Measurement accuracy: contact angle – 0.3°; surface tension – 0.3 mN/m [7].

Modes of *APPA* for wood samples: *mode A* – cathode-directed streamer mode at air gap of 6 mm; *mode B* – anode-directed streamer mode at air gap of 12 mm, Figure 8.

A halogen-free, phosphorous and nitrogen containing flame retardant (FR) based on ortho-phosphorous acid, urea and ammonia has been produced and studied as water solution. The impregnating *FR* water solution was based on it: dry substance of 30 wt. %; phosphorus content of about 13 wt. %, pH = 7÷8 and density of 1.15÷1.14 g/cm^3.

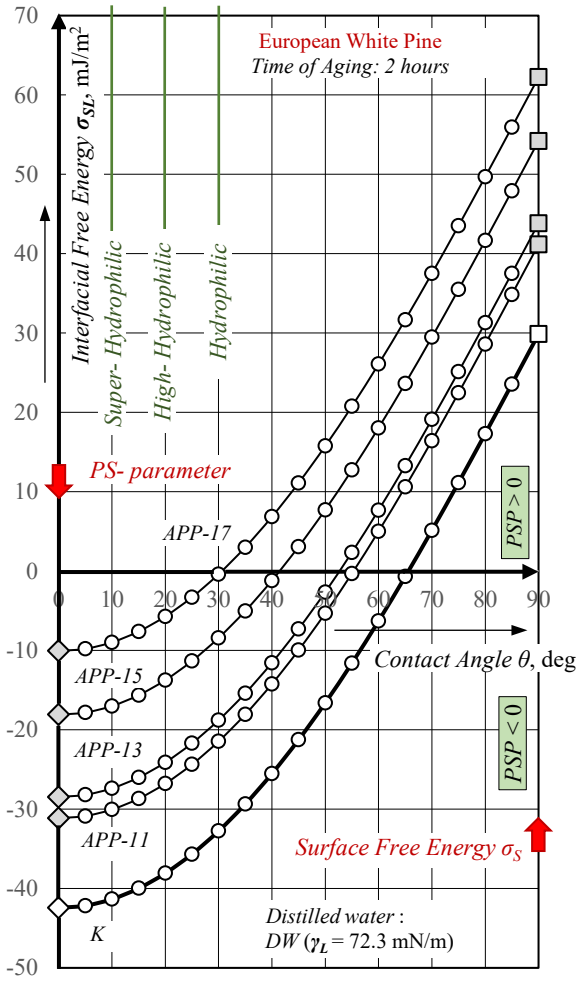


Fig. 9. The use of distilled water as a test-liquid with $SFT - \gamma_L = 72.3$ mN/m, does not carry adequate information on the efficiency of the finishing of *FR* capillary impregnation – no effective mode of *APPA* has been found at voltages up to 17 kV rms for *European white pine*: $PSP = \sigma_S - \gamma_L < 0$. All *SFE* values were determined according to the *method of Wu* for short aging time of two hours.

The non-equilibrium air (oxidative) *APPA* was combined with anionic micelle-forming (AS, "Anticrystalin A", *Chimatech, Ltd.*,

Bulgaria) and trisiloxane-ethoxylate surfactant (Y-17113, *Momentive Performance Materials* GmbH & Co. KG, Germany) enhanced capillary impregnation with well-known experimental methodology [10, 11, 12].

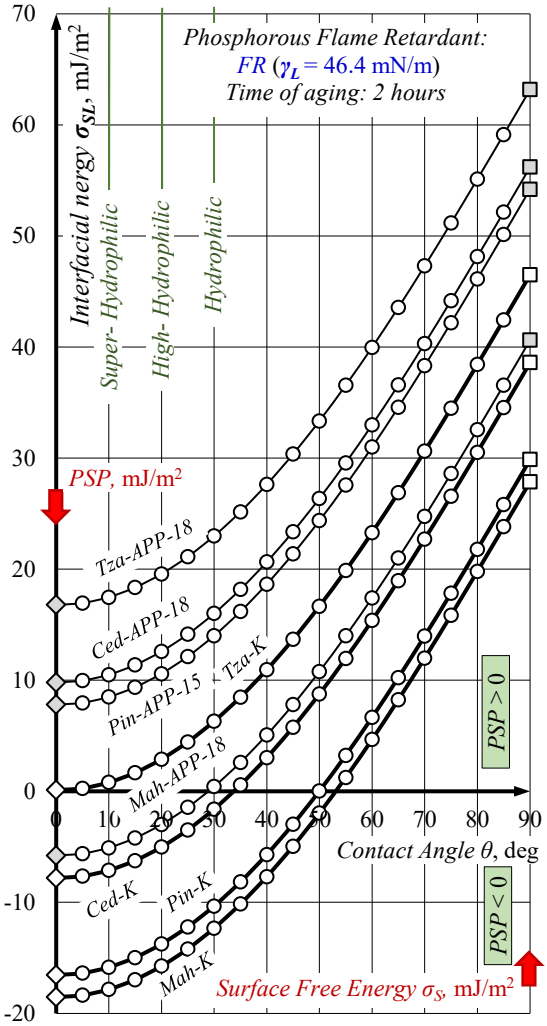


Fig. 10. Effective mode selection of *APPA* at an air gap of 6 mm and short aging time of two hours for plasma-aided capillary impregnation finishing was found only for *European white pine* (*Pin-LF-15*), *Mexican white cedar* (*Ced-LF-18*), and *Tzalam* (*Tza-LF-18*): *PSP* > 0. No steady *APPA* mode for *Mahogany caoba* (*Mah-LF-18*) was found: *PSP* = 0.1 mJ/m² ≈ 0. Plasma pre-activation modes provided trought the declared voltage – 15/18 kV rms, get the maximum *SFE* value for each species of studied woods.

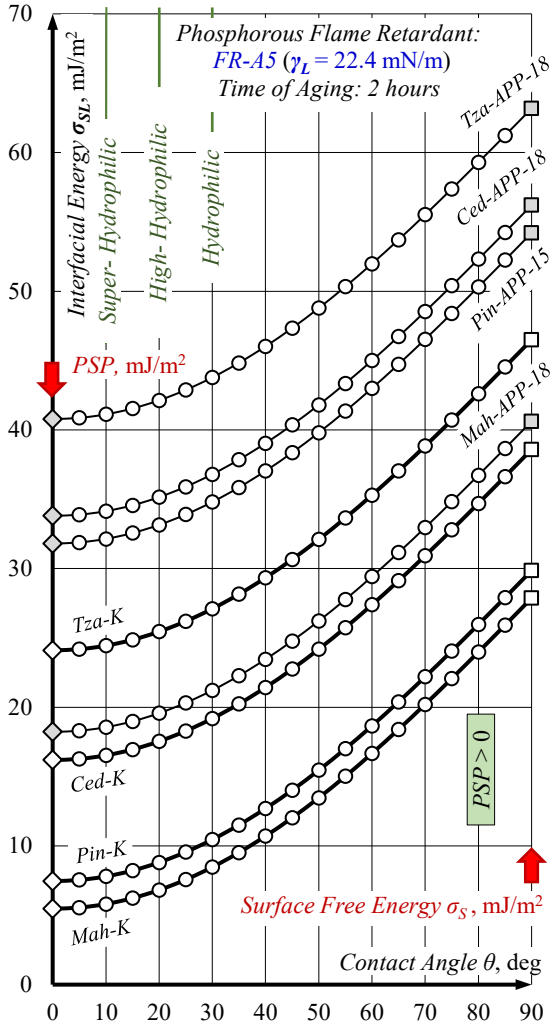


Fig. 11. Effective mode selection of *APPA* at an air gap of 6 mm and short aging time of two hours for plasma-aided capillary impregnation finishing was found for all wood species by decreasing the *SFT* of the basic phosphoric *FR* solution with *anionic micelle-forming water surfactant* at 5 vol. % (*FR-A5*): γ_L = 22.4 mN/m. The rule of processing efficiency was met for all applications of capillary impregnation – non-activated (*-K*) and *APP*-activated (*-APP-XX*) at different voltages: 11, 13, 15 and 17 kV rms: *PSP* > 0.

4. Experimental rezults and discusion

This study meets another issue which can be formulated as follows: "*is it possible a simple drop test with distilled water* (DW) *to adequately replace the entire PSP study conducted with a real impregnating FR-water solution*".

The answer can be found in the information provided in Figure 3 and 8. The use of distilled water as a test-liquid does not carry adequate efficiency assessment of the finishing – the plasma-aided capillary impregnation. The effective mode selection (*APP-XX*: *APPA* at 11, 13, 15, and 17 kV rms; 50/60 Hz) can be determined only by the use of *FR* solution – a basic *FR* water solution containing phosphor- and nitrogen flame retardant with *SFT*, γ_L = 46.4 < 72.3 mN/m, and the relationship between the "*solid-liquid*" *IFE* (σ_{SL}) and static contact angle θ measured by Sessile drop technique – *IFE* acquires a value equal to the *PS-Parameter* at *complete wetting*, i.e. contact angle equal to zero: σ_{SL} = *PSP* (θ = 0).

Distilled water has a too high *SFT*: γ_L = 72.3 > 46.4 >> 22.4 mJ/m². *APPA* cannot compensate it by increasing *SFE*, so a rule process efficiency should be enforced: $\sigma_S < \gamma_L$; *PSP* < 0, Figure 8. Normally, it fails to predict the efficiency of the plasma-aided finishing with modified γ_L = 22.4 mJ/m², or non-modified γ_L = 46.4 mJ/m², *FR*-solution, Figures 3 and 8.

All modes of *APPA* with a voltage higher than 13 kV rms (18.3 kV peak value) provide processing efficiency finishing of capillary impregnation: *PSP* > 0, Figure 3.

At the first stage, the *PSP*- testing indicates that the voltage of the *APPA* for European white pine samples should be 15 kV rms (σ_S = 54.2 kJ/m²) for a maximum processing efficiency of the *FR* (γ_L = 46.4 mJ/m²) impregnation finishing: max *PSP* = 7.8 mJ/m²

> 0, Figure 9.

The most effective voltage for the other studied species of wood was determined using the same procedure and it is 18 kV rms. They are arranged in order of decreasing *PS-parameter* according to their density, mJ/m^2: *Tzalam* – 16.78, *Mexican white cedar* – 9.8 and *European white pine* – 7.8. However, no steady *APPA* mode for *Mahogany caoba* was found: *PSP* = 0.1 mJ/m^2 ≈ 0. This result indicates that there are wood species for which a processing efficiency finishing – a capillary impregnation, cannot be applied, Figure 9.

At the second stage, the *PSP*- testing of the all studied wood species indicates that the modified *FR*- solution with less *SFT* γ_L = 22.4 mJ/m^2, performs the rule of processing efficiency after a short time of aging, *PSP* > 0, Figure 10.

At the third stage, the *PSP*- testing of all studied wood samples indicates that the *APPA* and the finishing of capillary impregnation with the modified impregnating *FR* solution (FR-A5), γ_L = 22.4 mJ/m^2, satisfy the rule of finishing processing efficiency: *PSP* > 0, Figure 11.

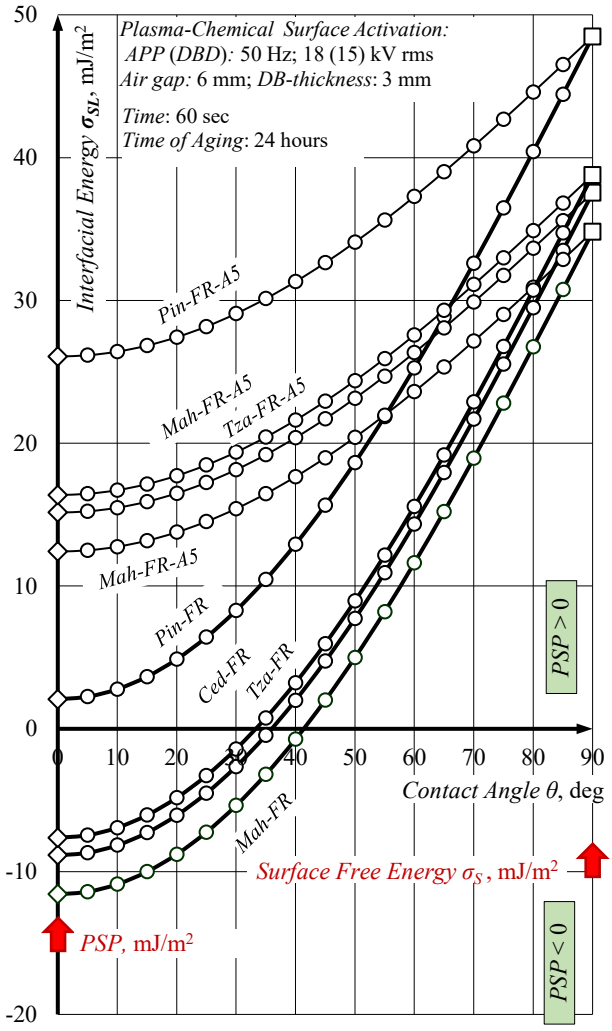


Fig. 12. Effective mode of *APPA* at air gap of 6 mm and short and long aging time - 2 and 24 hours, for plasma-aided capillary impregnation finishing was found for all studed wood samples by decreasing the *SFT* of the *FR* impregnating solution with anionic micelle-forming surfactant at 5 vol. % (*-FR-A5*): γ_L = 22.4 mN/m. The rule of processing efficiency was met for all applications of plasma-aided finishing with this modified solution: *PSP* > 0.

This study proves our claim that the plasma-aided capillary impregnation can be applied "out-of-the-process-line" after a long time of aging 24 hours, between *APPA* and the capillary impregnation finishing.

The plasma-aided capillary impregnation of porous woods has resulted in the formation of a *FR*-functional coating with a certain thickness, continuity and defect-free surface. In porous media, the *capillary pressure* p_C is proportional to the *adhesion tension* Σ and depends only on it at certain characteristic pore size, (3).

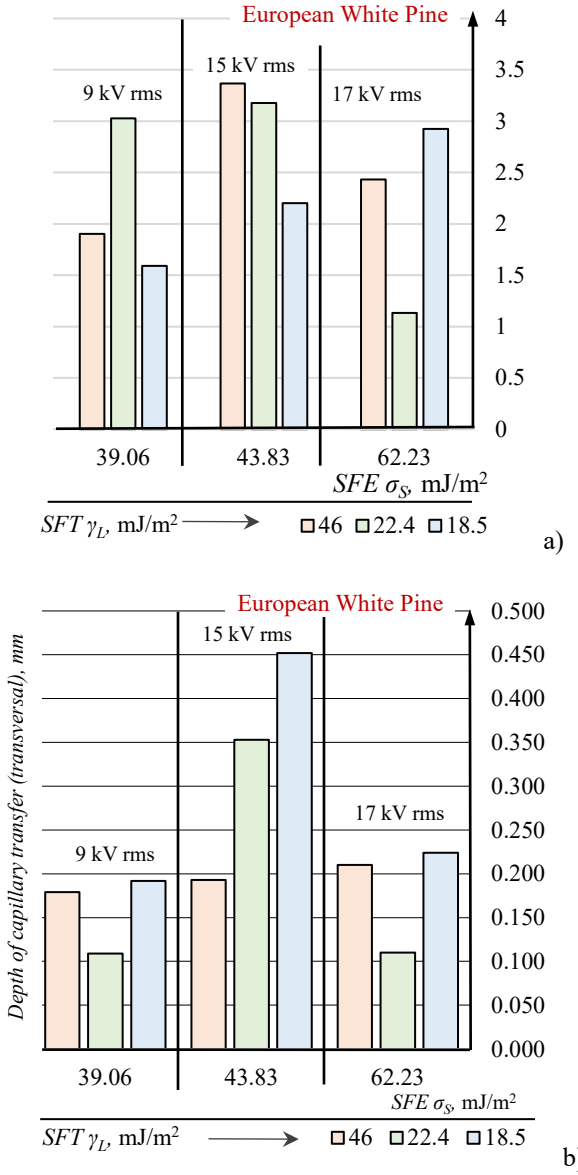


Fig. 13. Effect of *APPA mode* (voltage) – 9, 15 and 17 kV rms (50/60 Hz; air gap: 6 mm), the short time of aging – 2 hours, the used impregnating phosphorous *FR* solutions and consumption rate – constant rate: 0.139 l/m^2 (dm^3/m^2) on the depth of "capillary" transfer to the wood capillary direction: *longitudinal* – capillary displacement (**a**) and *transversal* – wicking through the capillary walls (**b**). The best quality of the plasma-aided capillary impregnation for flame retardancy of *European white pine* is at 15 kV rms.

It was expected that the largest *DBD* voltage of 17 kV rms should ensure greater depth of wicking transfer as the *SFE* and its greatest polar part, which determines greater value of *PS*-efficiency parameter. Although the *SFE* and its polar part are greater, Figure 4, just like the *PS*-parameter, Figure 3, the depth of

capillary transfer – longitudinal to the capillary structure and transversal to it, or through the capillary walls, is less than 15 kV rms, Figure 12. What is typical for this mode (15 kV rms) of plasma and surfactant enhanced capillary impregnation, is that the polar part of the *SFE* becomes greater than the dispersive part of *SFE*, Figure 4. The transition to more intensive mode of *APPA* (17 kV rms) does not give a positive result in spite of increasing *SFE* and its polar part, and *PS*-efficiency parameter.

Both the capillary displacement and the wicking through the capillary walls are two visible forms of maximum solution transfer, Figure 12. The maximum *capillary displacement* was determined by a maximum value of capillary pressure p_C or adhesion tension Σ respectively. This is determined by the maximum basic *FR* solution, Figure 5 and 13. On the contrary, the maximum *wicking through the capillary walls* was obtained at the minimum value of adhesion tension Σ; this is provided by the modified *FR-A5-S* solution. Wicking, however, is essential for the impregnation process. More extensive studies on different species of wood are needed to confirm this experimental fact, which has been observed only for European white pines so far.

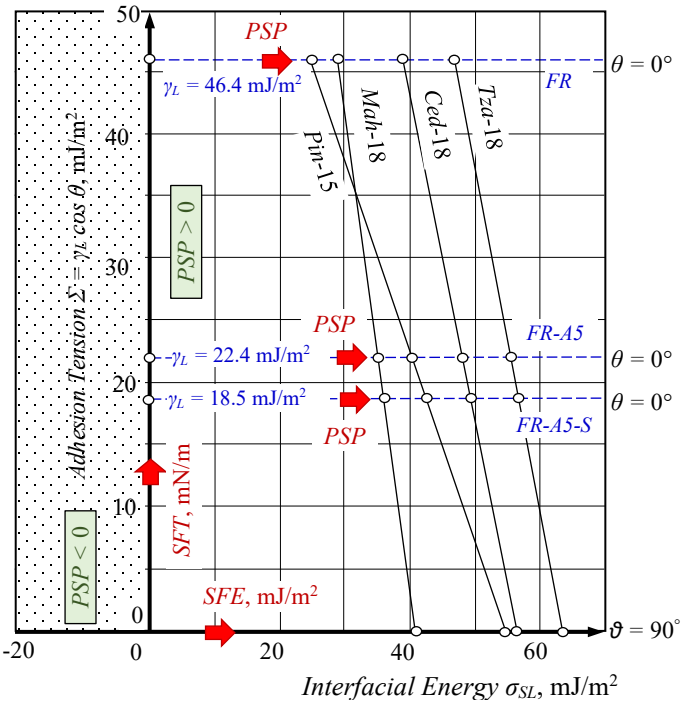


Fig. 14. Relationship between the *adhesion tension* Σ and the "*solid-liquid*" *interfacial energy* σ_{SL} that reveals at once the impact of: *i* – the variation of *SFE* as result of *APPA* at short time of aging (2 hours), air gap of 6 mm and voltage 15 kV rms; *ii* – the variation of *SFT* as a result of *FR*-solution modification with surfactants: *FR-A5* – anionic micelle-forming surfactant at 5 vol. %; *FR-A5-S* – trisiloxane neutral surfactant at 0.1 vol. %.

In conclusion, an attention has to be paid to the fact that the transfer of the *FR*-solution in depth is mainly carried out in the more complicated way through the capillary walls. The modification of the basic *FR*-solution through the combination of anionic micelle-forming and trisiloxane surfactant thus acquires a new meaning.

The graphical representation of the relationship between *adhesion tension* Σ and "*solid-liquid*" *IFE* σ_{SL} fully illustrates the two pathways for achieving the finishing efficiency *PSP* of plasma-aided capillary impregnation by increasing the *SFE* of porous media and reducing the *SFT* of the impregnating *FR* solution, Figures 5 and 13.

By selecting the voltage, the quality of the *APPA* expressed by the process efficiency parameter *PSP* was controlled: the *SFE* of the European white pine samples was increased; the *SFT* of the impregnating *FR* solution was decreased. The process efficiency parameter *PSP* meets the rule of plasma-aided finishing efficiency, *PSP* > 0, at and above 15 kV rms, Figure 5. For all wood species, the maximum process efficiency parameter *PSP* is at a voltage of 18 kV rms, Figure 13.

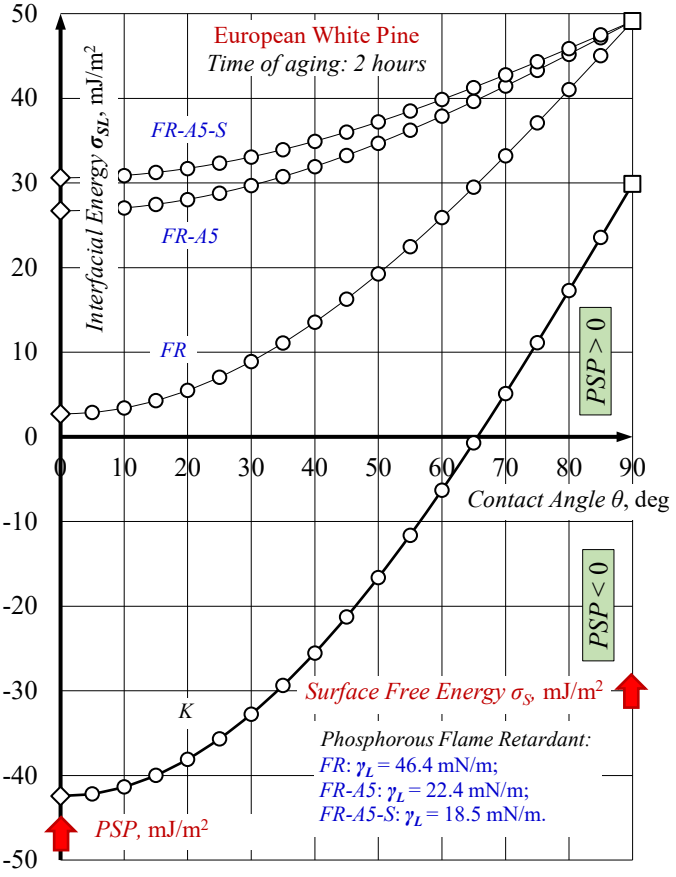


Fig. 15. Effective mode selection of plasma-aided capillary finishing of *European white pine* after two hours aging time from air *RAPPA* at air gap of 6 mm by a modification of the phosphoric *FR*-solution with an *anionic micelle-forming water surfactant* at 5 vol. % (*FR-A5*) and neutral trisiloxane surfactant at 0.1 vol. %. The rule of processing efficiency was met for all plasma-aided applications of capillary impregnation.

The new created plasma-aided finishing can find its wide application only if it can be envisaged and managed so as to observe the process efficiency rule: *PSP* > 0, Figure 15.

The prerequisite for this is the possibility for *SFE* measurement "*in situ*" out of the lab. We believe that this approach can be successful, as the market now has a technical device for measuring the *SFE* "in situ" – pocket goniometers, mobile or hand-held surface and contact angle analyzers.

There is also an innovative "mobile" analyzer with two test liquids using a dispenser with two parallel drops with "one click" for direct analysis of the static contact angles and the results obtained from the *SFE*.

Conclusion

Plasma, surfactant or plasma and surfactant aided finishing were used to enhance the finishing of capillary impregnation and thus the flame retardancy of wood: wood materials and constructions – timber, lumber, plywood, cardboard, paper, flooring, paneling, staircase and wooden structure with low *SFE* were subjected to *APPA* and *RAPPA* for increasing their *SFE*.

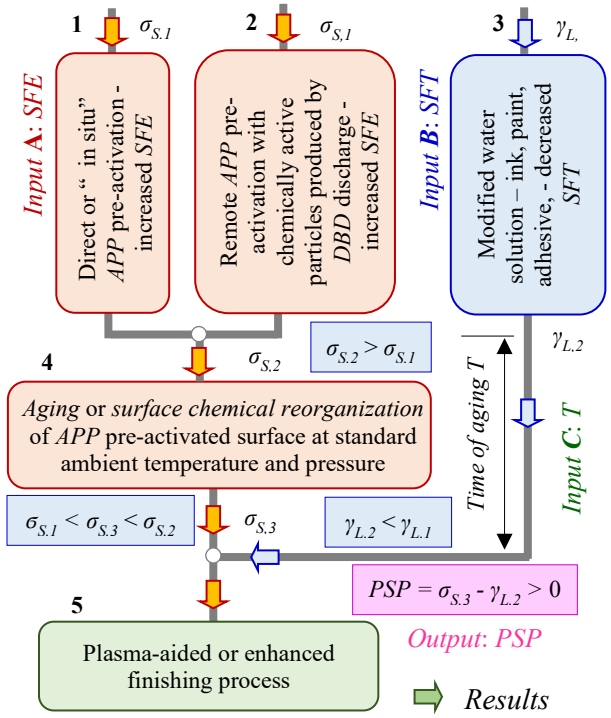


Fig. 16. Schematic representation of the plasma-aided impregnation finishing model – flow of information and processes involved in the processing efficiency and results to be obtained. Management of the plasma-aided finishing with two input factors – *SFE* of the porous solid and *SFT* of the liquid. The time of aging is set depending on the *APPA* and impregnation used technology.

In general, a processing efficiency parameter named *PSP* was found: "*Plasma, surfactant or plasma and surfactant enhanced finishing of capillary impregnation will be more successful and the material will be more susceptible when the PSP is positive:* $PS > 0$".

This process efficiency parameter was defined by the difference between the *SFE* of the porous wood and the *SFT* of the impregnation *FR*- solution: $PS = \sigma_S - \gamma_L$. This efficiency parameter is equal to the value that gets the *IFE* when the water solution wets the surface completely: $PS = \sigma_{SL}$, $\theta = 0°$ ($\cos\theta = 1$).

The efficiency of a plasma-aided impregnation finishing can be predicted only by setting both basic parameters of the real porous wood surface and the actual *FR*-impregnating solution – the *SFE* and the *SFT*. Successful plasma aided or enhanced finishing, including the capillary impregnation, can be expected if the process efficiency parameter takes a positive value, $PSP > 0$, after a prescribed aging time – 2 or 24 hours.

The evolution of plasma-enhanced wicking phenomena in time after *APPA* – the short and long-time aging of plasma activated surface, after 2 and 24 hours, was investigated for four species of wood by the *PS*-efficiency parameter change. All species of wood were ranked according to the plasma aided capillary impregnation process efficiency as per its *PSP*.

The existing methods and instruments for measuring the *SFE* of solids and *SFT* of liquids, including portable measurement systems, [20], allow this approach to determine the process efficiency of the plasma-enhanced finishing such as impregnation, printing, painting, laminating to be recommended for application "in-line" and "out-of-the-line".

References

[1] Dineff, P.; Gospodinova, D.; Ivanov, I. Processing efficiency of plasma-aided porous media finishing. Prossidings of the 15-th international conference on electrical machines, drives and power systems (ELMA), IEEE 2017, IEEE Part Number CFP17L07-ART, pp 334-338.

[2] Fridman, A. *Plasma chemistry*. Cambridge University Press, Cambridge, England, New York; 2008.

[3] *Plasma technologies for textiles*. Edited by Shishoo, R. Published by Woodhead Publishing Limited, Abington, Cambridge, England, 2007.

[4] *Plasma Technology for Hyperfunctional Surfaces. Food, Biomedical and Textile Application*, Edited by Rauscher, H.; Perucca, M.; Buyle, G. WILEY-VCH Verlag GmbH & Co., Weinheim Germany, 2010.

[5] Sernek, M. Comparative analysis of inactivated wood surfaces. Doctoral thesis, 2002. Glasser, W.; Kamke, F. (Co-Chairs). Virginia Polytechnic Institute and State University. Blacksburg, Virginia, USA.

[6] Viöl, W. Method for modifying wooden surfaces by electrical discharges at atmospheric pressure. WO039944A1 (EP1233854B1; US6818102). Priority Date: 01.12.1999; Publication Date: 16.11. 2004.

[7] Cernak, M.; Rahel, J. Apparatus and method for treatment of wood, wood fibres and wood-based materials. WO Patent WO2008085139A1 (EP2095694A1), Priority date: 05.12.2006; Publication date: 17.07.2008.

[8] Dineff, P.; Kostova, L. Method for plasma chemical surface modification. WO Patent WO2006133524A2 (BG 66022B1); Priority Date: 14.06.2005; Publication date: 12.04.2007.

[9] Podgorski, L.; Chevet, B.; Onic, L.; Merlin, A. Modification of wood wettability by plasma and corona treatments. *Internat. J. Adhes. Adhes* 2000, 20(2), pp 103÷111. Doi: 10.1016/S0143-7496(99)00043-3.

[10] Avramidis, G.; Wascher, R.; Militz, H.; Viöl, W. Impact of air-plasma treatment at atmospheric pressure on wood and wood extractives. Int. Wood P. 2016. 7(2) pp 76-79. 8-th European Conference on Wood Modification. Doi: 10.1080/20426445.2016.1162429.

[11] Nolan, M. Dyne testing –why test in the first place? Converting Supply, Inc., Ontario, Canada, 2011.

[12] Why test inks cannot tell the full truth about surface free energy. Application report No 272, Krüss GmbH, Hamburg Germany, 2014. www.kruss.de.

[13] Wolkenhauer, A.; Avramidis, G.; Militz, H.; Viöl, W. Plasma treatment of heat treated beech wood - investigation on surface free energy. Holzforschung 2008. 62(4), pp 472-474. Doi: 10.1515/HF.2008.074.

[14] Dineff, P.; Gospodinova, D.; Ivanov, I. Study on processing efficiency of plasma-aided capillary impregnation of porous media through a penetration-spreading parameter. Porceedings of the X-th international conference advanced manufacturing technologies & XI-th international conference advanced materials and operations (AMT&AMO 2016). 2016, pp 54-65.

[15] Dineff, P.; Ivanov, I.; Gospodinova, D. Efficiency processing parameter for plasma-aided surface impregnation: I. Penetration-spreading parameter. 19-th international symposium on electrical apparatus (SIELA) 2016. IEEE DOI: 10.1109/SIELA.2016. 7542991.

[16] Dineff, P.; Ivanov, I.; D. Gospodinova, "Efficiency Processing Parameter for Plasma-Aided Surface Impregnation: II. Interfacial Energy, Adhesion and Capillary Pressure. 19-th international symposium on electrical apparatus (SIELA) 2016; IEEE DOI: 10.1109/SIELA.2016.7542992.

[17] Dineff, P.; Ivanov, I.; Gospodinova, D.; Stanimirova, R. Efficiency Processing Parameter for Plasma-Aided Surface Impregnation: III. Plasma Surface Activation. 19-th international symposium on electrical apparatus (SIELA) 2016; IEEE DOI: 10.1109/SIELA.2016.7542993.

[18] Corona and plasma for extrusion. Vetaphone A/S, Denmark, 2017. www.vetaphone.com.

[19] Wolkenhauer, A., Avramidis, G.; Hauswald, E.; Militz, H., Viöl, W. Sanding vs. plasma treatment of aged wood: a comparison with respect to surface energy. Intern. J. Adhes. Adhes. 2009. 29(1), pp 18-22. Doi: 10.1016/j.ijadhadh.2007.11.001.

[20] Mobile surface anlyzer – MSA. Krüss GmbH, 2017. www.kruss.de.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 252-260 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Heuristic Based Approach for Voltage Stability Improvement using FACTS Devices

Hajer Jmii*, Asma Meddeb, Souad Chebbi

LATICE Laboratory, University of Tunis, Tunis 1008, Tunisia

ARTICLE INFO

Article history:
Received: 30 October, 2017
Accepted: 17 November, 2017
Online: 23 December, 2017

Keywords:
Heuristic method
SVC
TCSC
UPFC
Voltage Stability

ABSTRACT

The aim of this paper is to develop a hybrid device for voltage stability enhancement using two kinds of FACTS (Flexible AC Transmission System) namely SVC (Static Var Compensator) and TCSC (Thyristor Controlled Series Capacitor). The idea behind the proposed method is to maintain safe and satisfactory power system operation in a lesser costing manner by taking advantage of the performances of SVC and TCSC at the same time. We propose to evaluate the efficacy of the combined device to UPFC, as it is a hybrid FACTS and it is the most versatile compensator. For purpose of identifying the placement of the devices, we opt for a heuristic based approach. The methodology is tested with the IEEE 14-Bus system using the software EUROSTAG, and the simulation results reveal the efficiency of the proposed method for enhancing voltage stability.

1. Introduction

Today's power system is more and more operating close to its stability limits due to the unceasing growing of the power demand [1, 2]. Therefore, transmission lines are prone to the overloading, especially when the system is unable to cope up with power transition [3]. In addition, in some cases, the transmission of the energy is a hard task because of the long distance between the generation station and the loads. This may increase power losses and threats to the voltage stability. Voltage stability is defined as the system's ability to retain voltages at all the buses in the whole network within the specified boundaries after a disturbed condition [4]. Several system collapses have been reported in recent years [5] making the need for rapid and accurate control systems more and more insistent [6]. As power electronics components continue to develop, a variety of control devices as FACTS (Flexible AC Transmission System) have been prospered [7, 8]. The introduction of such a technology in the power system provides the control of the transmission line impedance, the voltage magnitude, and the phase angle. UPFC, SVC, and TCSC are three of the main FACTS devices which have acquired a well-recognized term for higher and smoother controllability in power systems.

The use of FACTS technology for system stability support has been studied in a large number of works [9-11]. A. Motiebirjandi *et al.* [12] studied the impact of UPFC on damping oscillations of the generator rotor. For this purpose, authors proposed system critical modes and residue factor methods for the optimal placement. In addition, they applied particle swarm optimization (PSO) to optimize the parameters of the UPFC. Reference [13] proposed the use of SVC to improve power system transient stability and damp oscillations in case of three-phase short-circuit. In reference [14], the authors demonstrated the performance of UPFC compared to SVC, in terms of improving power system stability. In order to enhance transient stability, the authors in [15] compared the performances of UPFC to different FACTS devices, namely TCSC, STATCOM (Static Synchronous Compensator) and SVC. B. Bhattacharyya et al. [16] were interested in the cost implication and power loss of installing UPFC alongside with SVC and TCSC. After determining the optimal emplacement and parameters of the FACTS using specific algorithms, authors have reported the gain obtained when adding the hybrid compensator to the other FACTS. Likewise, authors in [17] discussed the use of SVC, TCSC, and UPFC in the improvement of dynamic and transient system stability. They compared the three FACTS based on their mathematical models and operation modes. It was found that UPFC provided the most rapid control and the highest performances in stabilizing the system. P. Pandey *et al.* [18] studied the contribution of the SVC and UPFC in enhancing the voltage profile of a grid connected distributed generation system. They demonstrated through simulations, the satisfactory operation of the two FACTS especially the hybrid device. Reference [19] presented the application of a heuristic based procedure to

*Hajer Jmii, University of Tunis, Tunis 1008, Tunisia, +21694014865, hajerjmii@yahoo.com

parameters, type and location problems of UPFC, SVC and TCSC. Simulations showed satisfactory results of the proposed method.

This paper is an extension of work initially published in the Proceedings of the International Conference on Sciences of Electronics, Technologies of Information and Telecommunications [20]. The aim of this work is to improve power system voltage stability by using a hybrid device composed of SVC and TCSC. The benefits of this idea are to achieve a safe and secure power system operation in a less costly way. We choose to compare the performance of the proposed device to UPFC, as it belongs to the family of hybrid FACTS and it is the most powerful tool in the current control systems. The placement of the devices is identified by applying a heuristic based approach and the proposed method is tested on the IEEE 14-Bus system via the software EUROSTAG.

Firstly, in section 2 we present statistics about FACTS devices. Second, the voltage instability phenomenon is explained mathematically. Then, the mathematical models and structures of the proposed FACTS are detailed. Section 5 outlines the proposed approach. In section 6, we focus on the case study and the discussion of simulation results. Finally, we summarize the main points of this paper in the conclusion.

2. Statistics about FACTS Devices

Since the development of power electronics, a great interest has been granted to FACTS devices. The research considers this technology as a substantial and a timely topic. For this reason, an important number of studies discussing the application of FACTS to power systems have been published. In power system stability field, the publications related to FACTS are also numerous as described in Figure 1. It was found that the interest to the flexible controllers is more and more increased which reflects the efficiency of such a solution for instability problems. According to reference [21], when looking at the statistics for SVC, TCSC and UPFC publications, we noted that they are dominant compared to the other FACTS. Nonetheless, SVC is reaping the major concern in researcher's studies with 114 publications until 2004.

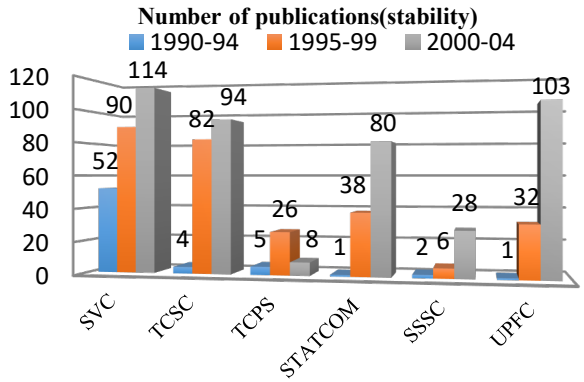


Figure 1. Statistics of the publications related to the FACTS

Actually, these statistics are reflecting the installed FACTS systems in the world. Reference [22] presented approximate results of a survey on worldwide integrated FACTS. It should be mentioned that SVC is the most exploited FACTS with a total power of 90.000MVA. While 2.000MVA are distributed on 10 incorporated TCSC into power systems, there are only 3 real networks equipped with UPFC with 250MVA of generated power.

To understand the reasons for the widespread use of SVC in comparison to the others FACTS particularly UPFC, we must look at system planner's choices. Recognizing that the investment cost is always a considerable constraint, economical solutions for power system instability are generally preferred. The cost of a FACTS device is highly dependent on the complexity of its model which is determined by the number of semiconductors used. Thereby, consisting of two voltage source converters, UPFC has the highest cost among the various FACTS [22]. According to reference [23], it is estimated at 0.33 million$ for 1MVAR generated power while the cost of SVC is approximately 0.19 million$ and the instrument and investment costs of TCSC are estimated at 0.22 million$ for 1MVAR generated power. Thus, when the network is already equipped with SVC, we can get well-improved stability by adding TCSC and at the same time economize 0.11million$ for each MVAR generated.

Otherwise, the cost-efficiency is not always to install the least costly FACTS device, but rather to choose the appropriate one based on the type of instability problem that we want to solve. As these problems are, on the one hand, unpredictable and on the other hand, closely linked, it is required to invest in equipment which can assure the control of more than one network parameter. Currently, UPFC has the potential to act on three parameters, namely: phase angle, line impedance, and bus voltage either simultaneously or separately. So, confronted with these constraints, we thought to find a solution for network instability, combining the economical side with the wide application area and efficiency side. Hence, our idea is to create an equivalent of UPFC by joining the action of TCSC, as a series FACTS, to that of SVC which assure the shunt compensation. With this approach, power system stability is enhanced at lower cost.

3. Voltage Instability

In power system stability field, a particular interest is accorded to the control of voltage and reactive power. The main issue is to avoid voltage instability which can result in a widespread problem known as voltage collapse [7]. For this purpose and in order to guarantee the reliability and the efficiency of power system operation, the following points must be respected in voltage and reactive power control:

- Voltages of all network buses must be within the permissible interval. According to reference [24], it is specified as [0.94pu; 1.06pu] for MV network.
- The reactive power flow in transmission lines must be minimized as much as possible. This leads to reducing line losses: RI^2 and XI^2.

We can illustrate the voltage instability by considering the system of Figure 2. This 2-bus network consists of a voltage source supplying a load through a transmission line.

The apparent impedance at bus 2 is expressed by:

$$S_2 = P_2 + jQ_2 = V_2 I^* \quad (1)$$

Where:

$$\bar{I} = \frac{\overline{V_1} - \overline{V_2}}{jX}$$

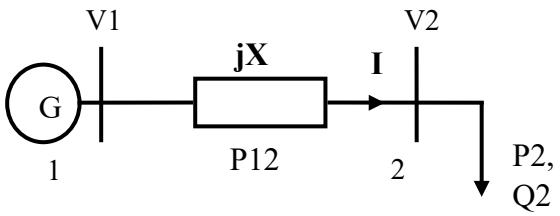


Figure 2. Equivalent model of a 2-bus system

Therefore equation (1) may be expressed as :

$$S_2 = V_2 \times \frac{V_1 \cos(\delta) + jV_1 \sin(\delta) - V_2}{jX} \quad (2)$$

$$S_2 = \frac{V_1 V_2 \sin(\delta)}{X} + j\frac{V_1 V_2 \cos(\delta) - V_2^2}{X} \quad (3)$$

The active and reactive powers supplied to the load are :

$$P_2 = \frac{V_1 V_2 \sin(\delta)}{X} = P_{\max} \sin(\delta) \quad (4)$$

$$Q_2 = \frac{V_1 V_2 \cos(\delta) - V_2^2}{X} \quad (5)$$

From equations (4) and (5) and using: $\cos^2(\theta) + \sin^2(\theta)$, we obtain:

$$P_2^2 + (Q_2 + \frac{V_2^2}{X})^2 = \frac{V_1^2 V_2^2}{X^2} \quad (6)$$

Therefore the voltage at bus 2 can be extracted as:

$$V_2^2 = \frac{1}{2}\left[V_1^2 - 2Q_2 X \pm V_1 (V_1^2 - \frac{4P_2^2 X^2}{V_1^2} - 4Q_2 X)^{1/2}\right] \quad (7)$$

Where :

V_1 : Busbar 1 voltage

V_2 : Busbar 2 voltage

I : Line current

X: Line reactance

δ : Generator internal angle

P_2 and Q_2 : Active and reactive powers supplied to the load.

Equation (7) describes the evolution of the voltage at bus 2 in function of the load increase at the same bus. We can deduce that in addition to the active and reactive powers, bus 2 voltage depends on the line reactance.

4. FACTS Models

In order to compare the performances of the FACTS devices, it is required to describe their functionalities supported with mathematical models and equivalent schemes. That is what we are going to present in this subsection.

4.1. TCSC

TCSC is one of the most important FACTS equipment, which has been used to modify the series impedance of the transmission line in order to improve system stability. Typically, it consists of a thyristor-controlled reactor in parallel with a fixed capacitor, which is equivalent to an adjustable reactance [25, 26]. The reactance of a transmission line equipped with a TCSC is expressed as follows:

$$X_{jk} = X_{jk}^{line} + X_{jk}^{TCSC} \quad (8)$$

Figure 3 shows a TCSC integrated between bus j and bus k with its equivalent model while Figure 4 describes the configuration of the series FACTS.

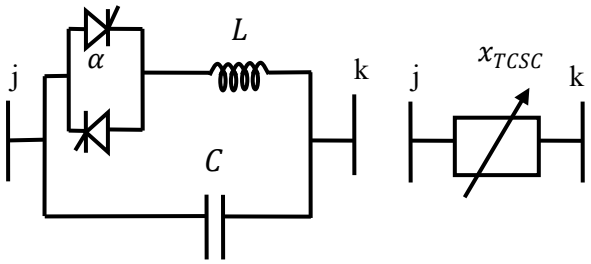


Figure 3. TCSC configuration

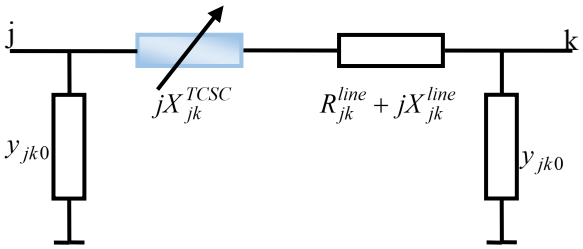


Figure 4. Equivalent model of TCSC

Before the integration of the TCSC in the transmission line jk, the admittance matrix Y is given by equation (9) [27]:

$$Y = \begin{pmatrix} \underline{y}_{jk} + \frac{\underline{y}_{jk0}}{2} & -\underline{y}_{jk} \\ -\underline{y}_{jk} & \underline{y}_{jk} + \frac{\underline{y}_{jk0}}{2} \end{pmatrix} \quad (9)$$

Where:

$$y_{jk} = G_{jk} + jB_{jk} \quad (10)$$

$$G_{jk} = \frac{R_{jk}^{line}}{(R_{jk}^{line})^2 + (X_{jk}^{line})^2} \quad (11)$$

$$B_{jk} = -\frac{X_{jk}^{line}}{(R_{jk}^{line})^2 + (X_{jk}^{line})^2} \quad (12)$$

The addition of the TCSC modifies the admittance matrix as follows:

$$Y' = \begin{pmatrix} \underline{y}'_{jk} + \frac{\underline{y}_{jk0}}{2} & -\underline{y}'_{jk} \\ -\underline{y}'_{jk} & \underline{y}'_{jk} + \frac{\underline{y}_{jk0}}{2} \end{pmatrix} \quad (13)$$

Where:

$$y'_{jk} = G'_{jk} + jB'_{jk} \quad (14)$$

$$G'_{jk} = \frac{R_{jk}^{line}}{(R_{jk}^{line})^2 + (X_{jk})^2} \quad (15)$$

$$B'_{jk} = -\frac{X_{jk}}{(R_{jk}^{line})^2 + (X_{jk})^2} \quad (16)$$

On EUROSTAG, to model a series and variable admittance, we must create two fictive buses in a transmission line and insert a current injector to each bus, as described in Figure 5. The two fictive buses must be connected to the network by lines with high reactance value to avoid injector shutdown in case of an opening line.

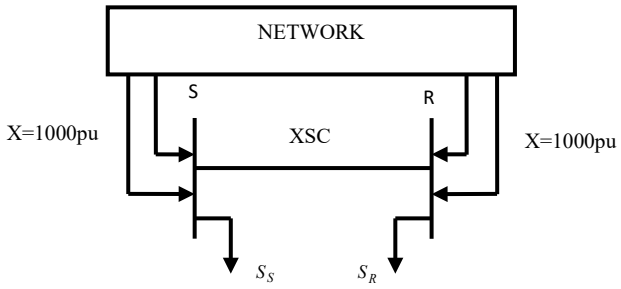


Figure 5. TCSC model on EUROSTAG

4.2. SVC

SVC is a static reactive power compensator whose output is adjusted for exchanging a capacitive or inductive current with the network to typically control bus voltage [28, 29]. In the steady state as well as in transient regime, this device is able to maintain voltage within the desired limits. Figure 6 shows the dynamic model of SVC. It can be modeled as variable shunt admittance with a thyristor controller. However, by neglecting the losses of SVC, we can consider it as ideal, so the admittance is purely imaginary and is described by the equations (17) and (18):

$$G_{SVC} = 0 \quad (17)$$

$$y_{SVC} = jB_{SVC} \quad (18)$$

The susceptance can be capacitive or inductive. Indeed, in the case of reactive power excess, SVC absorbs the increased amount through the inductor and in the opposite case; the capacitor covers the reactive demand.

The capacitive power injection model of SVC at the rated voltage is given by equation (19) [30]:

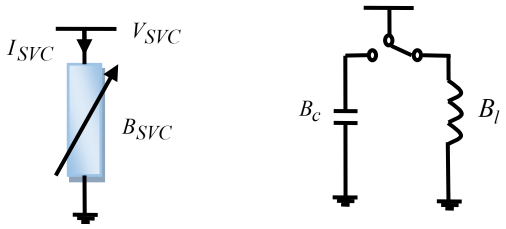


Figure 6. SVC model

$$Q_{SVC} = -V_N{}^2 B_{SVC} \quad (19)$$

Where B_{SVC} must be controlled according to equation (20):

$$B_{SVC}{}^{\min} \le B_{SVC} \le B_{SVC}{}^{\max} \quad (20)$$

$B_{SVC}^{\max}$ designates the capacitive limit state while $B_{SVC}^{\min}$ designates the inductive one. If SVC susceptance reaches its limits without maintaining the voltage of the bus where it is connected, it loses the capability of voltage control and it becomes similar to a fixed susceptance.

The connection of SVC to a bus j, as illustrated in Figure 7, changes the admittance matrix only at the element Y'_{ii}, in fact, the admittance of the compensator is added according to the following equation:

$$Y'_{ii} = Y_{ii} + y_{SVC} \quad (21)$$

The admittance matrix is so expressed by equation (22).

$$Y = \begin{pmatrix} y_{ij} + \frac{y_{ij0}}{2} + y_{SVC} & -y_{ij} \\ -y_{ij} & \frac{y_{ij0}}{2} + y_{ij} \end{pmatrix} \quad (22)$$

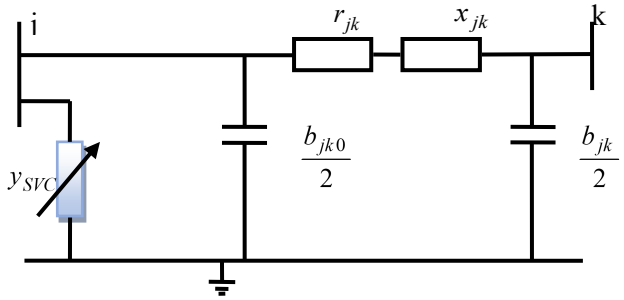


Figure 7. SVC connected to a bus

EUROSTAG adopts the model of Figure 8 and represents SVC as an impedance injector connected to a bus of the electrical network.

4.3. UPFC

UPFC is the most recognized hybrid compensator, it is the third generation of FACTS devices, it is the first and only device that has the ability to control simultaneously or separately voltage, line impedance and the phase shift of the bus voltage. UPFC consists

of two transformers, one connected in series and the other in shunt with the transmission line as shown in Figure 9 [31].

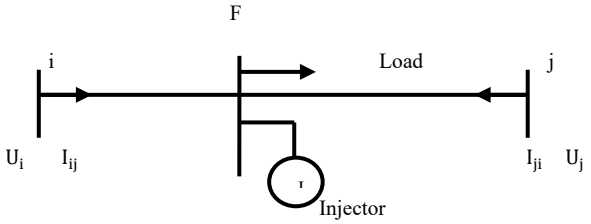


Figure 8. SVC model on EUROSTAG

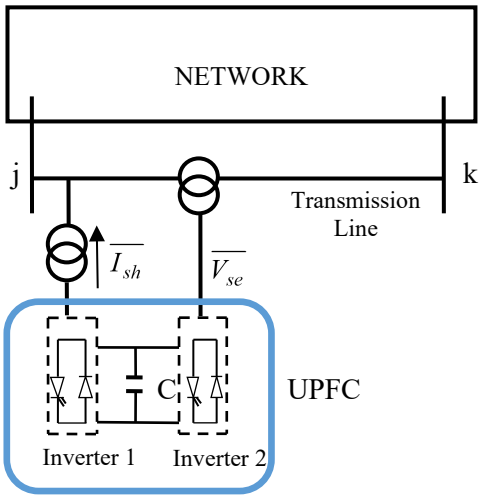


Figure 9. UPFC inserted in a power system

The inverter 2 injects the voltage $\overline{V}_{se}$ which is controllable in amplitude and in phase so that it can perform the serial compensation function of the active power. On the other hand, the inverter 1 is used, through the continuous connection, to provide the active power required for the inverter 2, it serves to compensate the reactive power since it can absorb or inject the active power into the network. Indeed, UPFC allows both the active power control and the line voltage control. It can switch from one to the other function instantaneously by changing the control of the inverters separated by a capacitor.

The equivalent circuit of UPFC is represented in Figure 10 [32]. From this scheme, we can extract the active and reactive power flows of the shunt and series converters which are expressed by the equations (23) - (25).

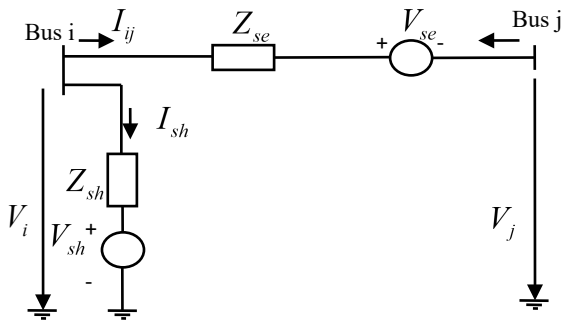


Figure 10. Equivalent circuit of UPFC

$$P_{sh} = V_i^2 g_{sh} - V_i V_{sh}(g_{sh}\cos(\theta_i - \theta_{sh}) + b_{sh}\sin(\theta_i - \theta_{sh})) \quad (22)$$

$$Q_{sh} = -V_i^2 b_{sh} - V_i V_{sh}(g_{sh}\sin(\theta_i - \theta_{sh}) - b_{sh}\cos(\theta_i - \theta_{sh})) \quad (23)$$

$$P_{ij} = V_i^2 g_{ij} - V_i V_j(g_{ij}\cos\theta_{ij} + b_{ij}\sin\theta_{ij}) - V_i V_{se}(g_{ij}\cos(\theta_i - \theta_{se}) + b_{ij}\sin(\theta_i - \theta_{se})) \quad (24)$$

$$Q_{ij} = -V_i^2 b_{ij} - V_i V_j(g_{ij}\sin\theta_{ij} - b_{ij}\cos\theta_{ij}) - V_i V_{se}(g_{ij}\sin(\theta_i - \theta_{se}) - b_{ij}\cos(\theta_i - \theta_{se})) \quad (25)$$

Where:

$$g_{sh} + jb_{sh} = \frac{1}{Z_{sh}}$$

$$g_{ij} + jb_{ij} = \frac{1}{Z_{se}}$$

$$\theta_{ij} = \theta_i - \theta_j$$

V_i and V_j: voltages at buses i and j

P_{sh} and Q_{sh}: active and reactive power flows of the shunt inverter

P_{ij} and Q_{ij}: active and reactive power flows of the series inverter

V_{sh} and V_{se}: shunt and series voltage sources

Z_{sh} and Z_{se}: shunt and series coupling transformer impedances.

The modeling of the UPFC shunt part on EUROSTAG is simple; it is represented by a current injector. As for the modeling of the series part, we must open the line where we want to insert UPFC and place at its extremities two current injectors (Figure 11). The opening of the line is assured by a high reactance value.

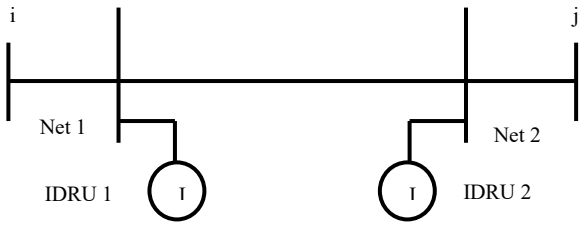


Figure 11. UPFC model on EUROSTAG

5. Proposed Approach

The main purpose of our paper is to find an economical and efficient way for sustaining voltage stability by comparing the performances of SVC-TCSC to those of UPFC. In order to achieve

this, it is important to take into consideration the identification of the proper placement of the FACTS in the system.

There is a variety of methods for finding optimal solutions [19, 33]. However, in practice, the time factor is more important than optimally solving the problems so that engineers and system planners often resort to heuristic methods. Heuristic methods known also as approximate methods are used for the resolution of optimization problems. These methods guarantee to find in a polynomial time at least a good and a feasible solution, but not necessarily the optimal one [34].

There are different types of heuristic algorithms, as shown in Figure 12, among which we are interested in reduction heuristics. The principle of such a method is to simplify the problem and to reduce the domain of the solutions. This implies to define the boundaries of the problem by identifying the properties that must satisfy a good solution.

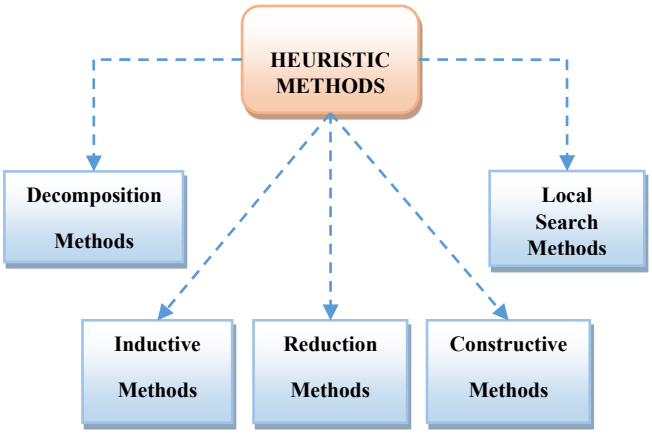


Figure 12. Types of heuristic algorithms

Figure 13 shows the flowchart of the proposed method. The normal state of the system must be obtained first of all so that we can analyze the load flow results and determine the weakest zone of the system that needs to be compensated. In this way, the location problem is simplified and the domain of solutions is reduced to the weakest zone. Based on the load flow analyses, a preliminary selection of the FACTS placement is made. Then, the system with the FACTS devices is executed. The evaluation of the selected location is performed by comparing the voltage magnitude to the specified lower and upper boundaries. A proper FACTS placement must maintain the voltage of all the buses within the permissible limits. If this criterion is not respected, another location has to be considered, and the evaluation based on the defined boundaries is repeated. Once the best location of the devices is identified, the next step is to perform the comparison between the proposed FACTS. The comparison takes into consideration the voltage magnitude, the oscillations, and the power flows.

With this method, we can get satisfactory results and avoid complicated and slow routines of the optimization methods.

6. Simulation

6.1. Test System

The proposed case study is the classical system IEEE 14-bus network, it consists of:

- Two generator buses and eleven load buses.
- Fifteen transmission lines.
- A three-winding transformer and two step-up transformers with two windings.
- Three synchronous compensators connected to buses 3, 6 and 8.

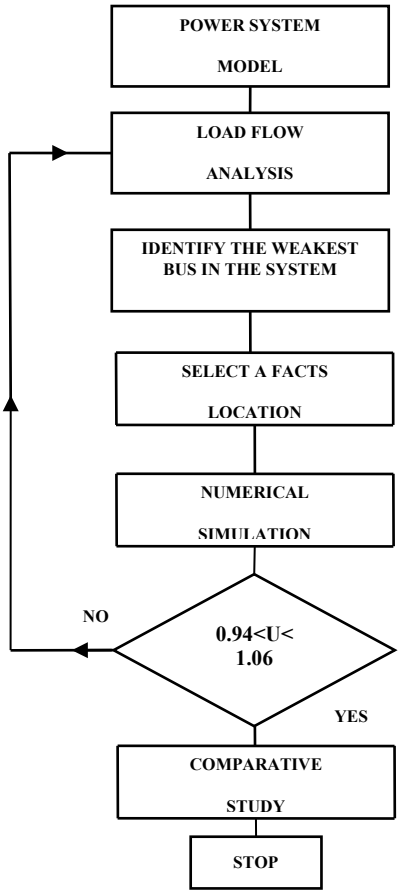


Figure 13. Flowchart of the proposed method.

All data relating to this test network are extracted from reference [32] and we used EUROSTAG software package [35] for the simulations. This software is a powerful tool dedicated to dynamic simulations. Its major advantage is the high rapidity of its algorithm. Irrespective of the nature of the disturbance and the system size, it can rapidly visualize the behavior of the network until it returns to an equilibrium state. Furthermore, this simulation is not accompanied by any deterioration in the accuracy of the calculations. In addition, it gives the possibility to define the modeling level adapted to the type of the performed study.

6.2. Simulation Results

6.2.1. Identification of FACTS Location

To compare the performances of UPFC and the proposed SVC-TCSC, first of all, it is mandatory to identify the appropriate placement of the FACTS. This step has to provide a firm basis for comparison. Therefore, we plan a step change in load at time t= 200s from 5% of the initial load by a step of 5%. The procedure

continues until reaching the maximum loading capability of the system.

The results of the base state signaled that bus 14 has got the lowest voltage amplitude as shown in Figure 14; hence we summarized the voltage levels of bus 14 during various step changes in Table 1.

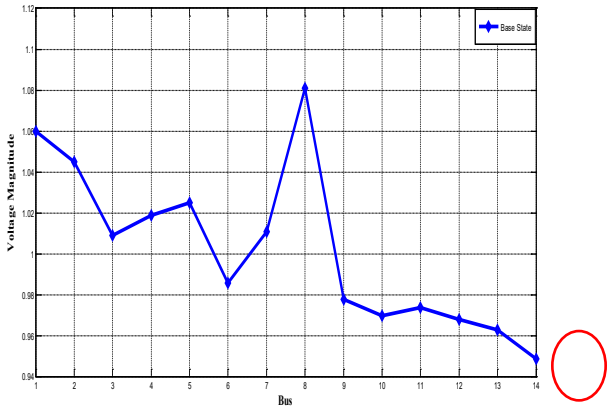


Figure 14. Voltage profile of IEEE 14-bus system in normal state

Table 1. Network behavior under incremental load increase.

Network Settings	Load Increase in %					
	5%	*10%*	*20%*	*25%*	*30%*	*33%*
U14 (pu)	0.966	0.962	0.952	0.948	0.946	Voltage collapse
State System	Stable					

We note the occurrence of a voltage dip in bus 14, which is increased as more the loads are augmented. The voltage level is still acceptable and the network is maintaining its stability until 32% of load increase. When reaching the case of 33%, EUROSTAG fails to execute the simulation which means that the system is collapsing.

After identifying the weakest bus of the system, we make a preliminary selection of the FACTS location. Hybrid or series controller both must be integrated through a transmission line. According to the network architecture, as schematized in Figure 15, bus 14 is related to bus 13 and bus 9, so we have two possibilities to place the device.

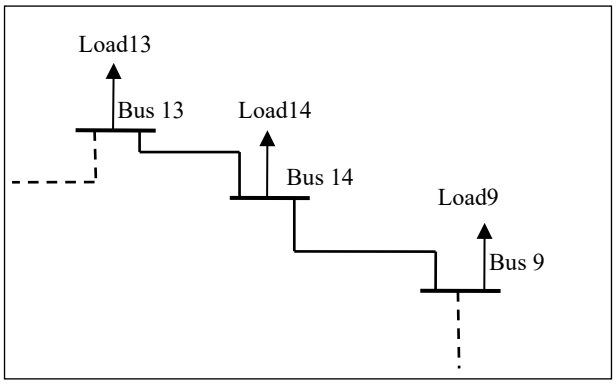


Figure 15. Scheme of the architecture of the IEEE 14-bus system

First of all, we integrate UPFC in the middle of line 13-14. Then, we apply an increase of 20% of the total load of the system. As shown in Figure 16, there is no improvement in the magnitude of bus 10 voltage. In addition, lower amplitudes equal to 0.925pu and 0.977pu are obtained respectively in bus 14 and bus 9, compared to the uncompensated system. Thus, it is concluded that line 13-14 is not a good location for the FACTS to perform the comparative study.

Now, we integrate UPFC through line 9-14 and we rerun the system. A meaningful improvement of the voltage magnitude is observed. According to Figure 16, UPFC has increased the voltage amplitude of bus 14 from 0.952pu to 0. 1.007pu with less severe oscillations, bus 10 from 0.976pu to 0.985pu and bus 9 from 0.982pu to 0.991pu. As we have got satisfactory results, the line 9-14 is retained as a good placement to insert the proposed FACTS and then to compare their potentialities.

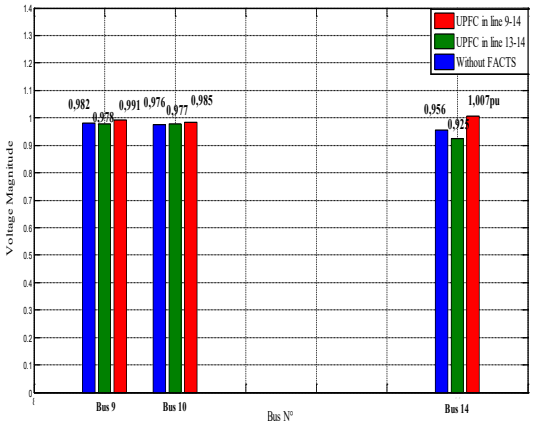


Figure 16. Voltage profile with different UPFC locations.

6.2.2. *Performance comparison between the FACTS*

We connect UPFC and SVC-TCSC to the network, separately. Firstly, UPFC was integrated into the middle of line 9-14 with a capacity of 60MVAR for each one of its inverters, and then, with keeping the same dimensions, we inserted TCSC through the line 9-14 and SVC to bus9.

Based on the results of the prior section, we apply a total load increase of 30% of the initial load. In this case, the test network is under heavy loaded conditions.

We are interested in the comparison, to the evolution of bus voltage as well as the active and reactive power flows in transmission lines.

Taking bus 14 as an example as shown in Figure 17, it can be seen that both UPFC and SVC-TCSC have a bearing on the voltage magnitude by increasing it significantly, while it was under the minimum acceptable value before compensation. However, UPFC showed a higher capability of enhancing voltage level, from 0.90pu to 0.96pu. In addition, we obtained a rapid damping of oscillations after only 7s either by UPFC or SVC-TCSC.

The impact of the FACTS on the oscillatory regime is clearer in the voltage curve of bus1, plotted in Figure 18. Nonetheless, we note that the damping action of SVC-TCSC is more

considerable compared to that of UPFC, in fact, it reduced the highest oscillation from 1.08pu to 1.071pu and rapidly established the steady state with well-amortized oscillations.

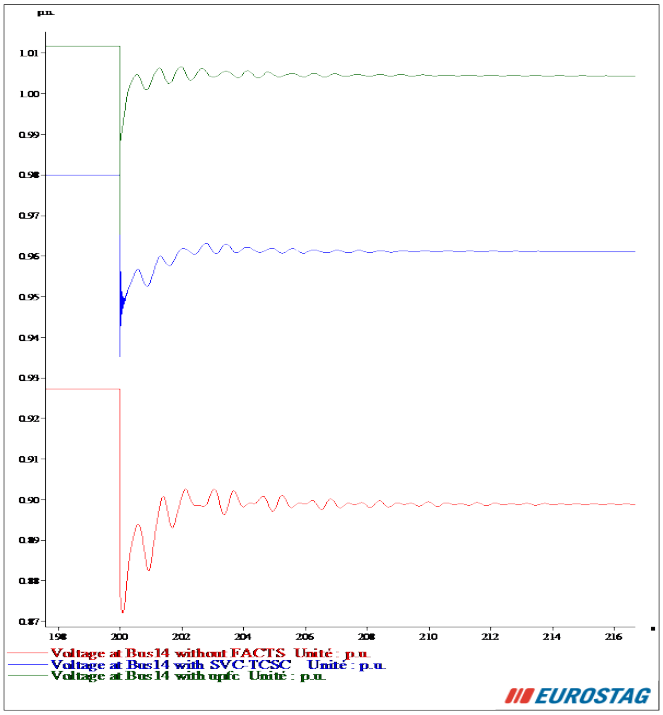


Figure 17. Voltage of bus14 with and without FACTS.

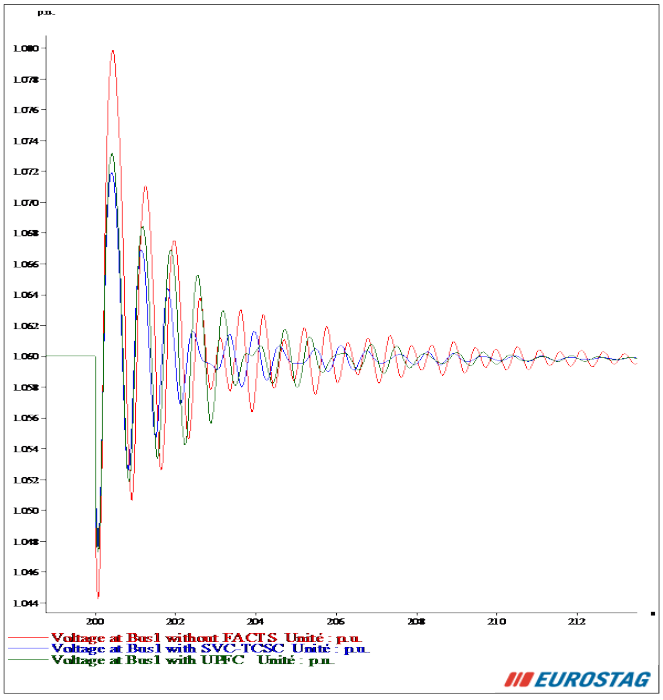


Figure 18. Voltage of bus1 with and without FACTS.

According to the results of Table 2, an excellent enhancement in the active power has been observed with the action of UPFC. However, we didn't note any important change in its level when used SVC-TCSC. As for the reactive power flow, it encountered a considerable decrease by using UPFC; it was decremented from 5.7 MVAR to 2.5MVAR through the line 9-10 as an example. This was observable also when introducing the combination SVC-TCSC. The amount of reactive power in line 6-12 is reduced by about 12% and in line 1-2, from 20MVAR to 19.2MVAR. Nevertheless, we recorded an increase of 10.5% in the transmitted MVARs across line 9-10.

Figure 19 shows the active power transmitted in line 1-2 with and without FACTS. When focusing on the transient regime, it must be mentioned that even in the temporal evolution of powers, the damping action of SVC-TCSC is more efficient; it gives more attenuated oscillations and joins the stable state in a shorter time.

Table 2. Active and reactive power flows with and without FACTS.

Simulation case	Line 9-10		Line 1-2		Line 6-12	
	Active Power MW	*React. Power MW*	*Active Power MW*	*React. Power MW*	*Active Power MW*	*React. Power MW*
Without FACTS	8.3	5.7	209	20	9.8	3.4
With SVC-TCSC	8.3	6.3	209	19.2	10	3
With UPFC	9.2	2.5	216	18.1	10.8	1.5

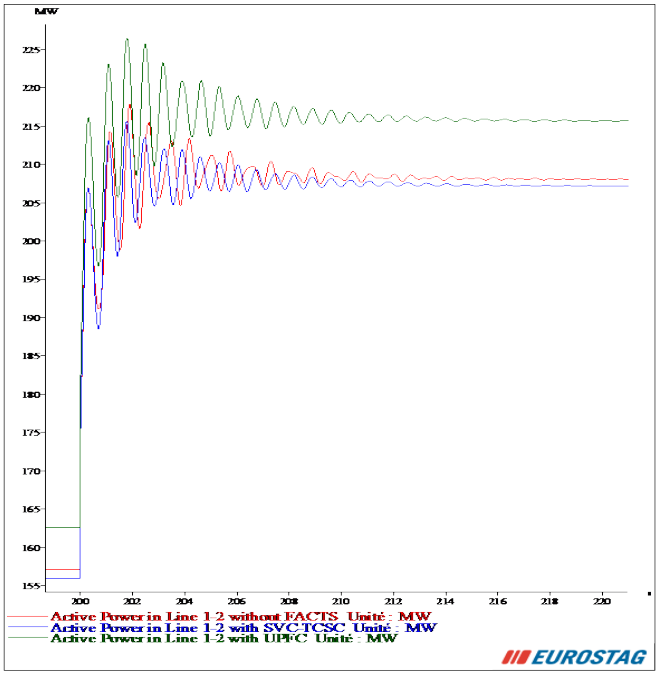


Figure 19. Active power flow in line 1-2 with and without FACTS.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

7. Conclusion

This paper presented a hybrid method for voltage stability enhancement using SVC and TCSC. The advantage of the proposed approach is the lesser cost of achieving a safe power

system operation. The IEEE 14-bus system was utilized to test the performance of the combination SVC-TCSC and compare it to the UPFC, the most pricey and multi-functional FACTS. The identification of the best placement of the FACTS devices was performed using a simple heuristic method. The simulation results demonstrated that the hybrid FACTS based on SVC and TCSC is able to significantly improve the voltage profile of the whole network in addition to the satisfactory action in amortizing the oscillations. The findings are encouraging to rather improve the performance of the SVC-TCSC in controlling power flow in lines and introduce other techniques for optimal location.

References

[1] B. Singh, "Applications of FACTS Controllers in Power Systems for Enhance the Power System Stability: A State-of-The-Art," International Journal of Reviews in Computing, Vol. 6, 15th July 2011.

[2] P. Kundur, "Power system stability and control," McGraw-Hill, Inc., 1993.

[3] K. Damor, . D. Patel, V. Agrawal, H. Patel, "Improving Power System Transient Stability by using FACTS Devices," International Journal of Engineering Research & Technology (IJERT), Vol. 3, Issue 7, July 2014.

[4] S. Rath, B. Sahu, P. Dash, "Power System Operation and Control using FACTS Devices," International Journal of Engineering Research & Technology (IJERT), Vol. 1, Issue 5, July 2012.

[5] S. Dheebika, R. Kalaivani, "Optimal Location of SVC, TCSC and UPFC Devices for Voltage Stability Improvement and Reduction of Power Loss using Genetic Algorithm," International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE), 2014.

[6] Sh. Ahmad, F.M. Albatsh, S. Mekhilef, H. Mokhlis, "An Approach to Improve Active Power Flow Capability by Using Dynamic Unified Power Flow Controller," *IEEE Trans*., 2014.

[7] A. Gupta, P.R. Sharma, "Static and Transient Stability Enhancement of Power System by Optimally Placing UPFC (Unified Power Flow Controller)," *IEEE trans.*, 2013.

[8] H. Jmii, A. Meddeb, S. Chebbi, "Efficiency Limits of SVC in Improving Voltage Stability," 3rd Inter. conf. (GEEE), April, 2016, Hammamet, Tunisia.

[9] P. Sivachandran, T. Hariharan, R. Pushpavathy, "Improvement of Power System Stability using D-FACTS Controllers: A Review," ARPN Journal of Engineering and Applied Sciences, Vol. 10, No. 2, February 2015.

[10] B. Rao, P. Chanti, N. Lavanya, S. Sekhar, Y. kumar, "Power System Stability Enhancement Using Fact Devices," International Journal of Engineering Research and Applications, pp.339-344, Vol. 4, Issue 4(Version 1), April 2014.

[11] M. Kamarposhti, M. Alinezhad, H. Lesani, Nemat Talebi, "Comparison of SVC, STATCOM, TCSC, and UPFC Controllers for Static Voltage Stability Evaluated by Continuation Power Flow Method," IEEE Electrical Power & Energy Conference, 2008.

[12] A. Motiebirjandi, D. Fateh, "Optimal placement method of multi UPFCs to damp power system oscillations," International Transactions on Electrical Energy System, 2017.

[13] M. Benaissa, S. Hadjeri, S.A. Zidi, "Impact of PSS and SVC on the Power System Transient Stability," Advances in Science, Technology and Engineering Systems Journal, pp. 562-568, Vol. 2, No. 3, 2017.

[14] A. Meddeb, H. Jmii, S. Chebbi, "UPFC and SVC Devices for Transient Stability Enhancement," 4th International Conference on Automation, Control Engineering and Computer Science (ACECS) Proceedings of Engineering and Technology – PET, pp.82-87, Vol.19, 2017.

[15] S. Titus, B.J. Vinothbabu, I. Nishanth, "Power System Stability Enhancement under Three Phase Fault with FACTS Devices TCSC, STATCOM and UPFC," International Journal of Scientific and Research Publications, Vol. 3, Issue 3, March 2013.

[16] B. Bhattacharyya, V.K. Gupta, S. Kumar, "UPFC with series and shunt FACTS controllers for the economic operation of a power system," Ain Shams Engineering Journal, 2014.

[17] S. Varma, "FACTS Devices for Stability Enhancements," International Conference on Green Computing and Internet of Things, 2015.

[18] P. Pandey, B. Bag, "A Comparative study on UPFC and SVC towards Voltage Profile improvement of a Grid Connected Distributed Generation System," *IEEE tran.s*, 2015.

[19] P. Stefanov, A. Savić, G. Dobrić, "Power System Optimization Using Parallel Scenario Algorithm," IEEE International Energy Conference, 2014.

[20] H. Jmii, A. Meddeb, S. Chebbi, "An approach for improving voltage stability by combination of SVC and TCSC," 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), pp.134–141, 2016.

[21] M.A. Abido, "Power System Stability Enhancement using FACTS Controllers: A Review," The Arabian Journal for Science and Engineering, vol.34, no1B, Apr 2009.

[22] X.P. Zhang, C. Rehtanz, B. Pal, "Flexible AC Transmission Systems: Modeling and Control," 2nd Edition, Springer, Feb 2012.

[23] M. Sadi, S. Ali, "A Comprehensive Analysis Of Transient Stability Enhancement Methods Of Electric Power System," *IEEE trans*, 2015

[24] Characteristics of voltage in public distribution networks, Standard EN 50160, 2007.

[25] M. Zellagui, H.A. Hassan, A. Chaghi, A. Ghorbani, "A Comparative Study of Ground Fault Analysis for a Practical Case of a Transmission Line Equipped with Different Series FACTS Devices," Automatika Journal, pp 262–274, 2015.

[26] D. Murali, M. Rajaram, N. Reka, "Comparison of FACTS Devices for Power System Stability Enhancement," International Journal of Computer Applications, Vol. 8, No.4, October 2010.

[27] Y. Amrane, M. Boudour, M. Belazzoug, "A new hybrid technique for power systems multi-facts optimization design," International Transactions on Electrical Energy Systems, 2014.

[28] C.V. The., K.L. Minh., T.T. Quoc, N.B. Khue., "FACTS devices applications on power system to improve the angle stability," IEEE, 2006.

[29] S. Varma, "FACTS Devices for Stability Enhancements," *IEEE trans.*, 2015.

[30] P. Kumkratug, "Application of Lyapunov Theory and Fuzzy Logic to Control Shunt FACTS Devices for Enhancing Transient Stability in Multimachine System," Journal of Electrical Engineering & Technology, vol. 7, no. 5, pp. 672-680, 2012.

[31] M. Haque, "Evaluation of First Swing Stability of a Large Power System with Various FACTS Devices," IEEE Transactions on Power Systems, Vol. 23, No. 3, August 2008.

[32] F. Milano, "Power System Modelling and Scripting," Springer, ETSII, University of Castilla - La Mancha 13071, Ciudad Real Spain 2010.

[33] R. Dubey, S. Dixit, G. Agnihotri, "Optimal Placement of Shunt FACTS Devices Using Heuristic Optimization Techniques: An Overview," International Conference on Communication Systems and Network Technologies, 2014.

[34] R. Marti, G. Reinelt, "Exact and Heuristic Methods in Combinqtorial Optimization,"Available:https://books.google.tn/books?id=4Lk8pPYYI3cC&printsec=frontcover&hl=fr#v=onepage&q&f=false.

[35] Eurostag, Eurostag Software Release Notes, Tractebel-EDF, Release 5.1, Dec 2010.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 261-266 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Enhancement of Voltage Stability through Optimal Location of UPFC

Hajer Jmii*, Asma Meddeb, Souad Chebbi

LATICE Laboratory, University of Tunis, Tunis 1008, Tunisia

ARTICLE INFO

Article history:
Received: 31 October, 2017
Accepted: 02 December, 2017
Online: 23 December, 2017

Keywords:
Numerical method
Optimal location
UPFC
Voltage Stability

ABSTRACT

This paper proposed the identification of the proper placement of UPFC (Unified Power Flow Controller) using a series of methodological numeric simulations to improve voltage stability. For this purpose, the critical zone of the system is determined, then, comparative analyses depending on different emplacements of UPFC are performed. The dynamic model of UPFC and the proposed method are tested on the IEEE 14-bus system. The finding shows that the proper location of UPFC helps in improving voltage profiles and increasing the maximum loading capacity.

1. Introduction

Power systems have been forced to operate close to their stability limits and around their full capacities. This is due to the economic and environmental pressures imposed on the power generation and the construction of new transmission lines [1-4]. On the other hand, the amount of power transmitted across transmission lines is limited for safety and stability considerations. The system is prone to be collapsed when the power flowing exceeds the limits in case of an arbitrary incident [5]. System planners are so more and more focusing on how to manage the obstacles facing the power transmission.

Over the last decades, FACTS (Flexible AC Transmission Systems) devices have opened new perspectives for managing the transmission of the power [6]. Different types of FACTS controllers are available for this aim, but the most powerful and prominent one is UPFC. UPFC is able to control the three parameters of the transmission system: phase angle, impedance, and voltage. In addition, it provides voltage support, improves transient stability and amortizes oscillations [7, 8].

However, according to references [9, 10], the major disadvantage of UPFC is the expensive cost of installation and reactive power generation which is due to the great efforts of its voltage source converters. Therefore, an optimal location of the device helps in achieving good performances and enhancing power system stability with a reduced investment cost [11].

Previous studies have been focused on the optimal placement of UPFC and the evaluation of its impact on system stability using various methods and techniques [12-14]. The authors in reference [15] presented an evolutionary algorithm-based approach to identify the optimal placement and settings of the hybrid device. Reference [16] has investigated the use of UPFC to increase the load ability margin and improve the system stability. The emplacement of the FACTS is decided by the stability indices which determine the critical line, the appropriate line for introducing UPFC. In reference [17], H. Wang investigated the impact of UPFC on the stability of a multi-machine power system. A. Ray *et al.* [18] studied the application of UPFC for controlling power flow in the transmission line and improving the voltage profile. They computed the VSI (Voltage Stability Index) for each line to integrate UPFC into the line which has got the highest VSI value. In reference [19], authors opted for the CPF (Continuation Power Flow) and LSI (Line Stability Index) for the identification of the suitable location of UPFC. Through simulation results, it was demonstrated that optimal placement of the hybrid FACTS improves the static and transient stability of power system when subjected to large disturbances. S. Lee *et al.* [20] addressed the determination of UPFC location and rating for voltage stability improvement using the analysis of the Jacobian matrix. Intelligent techniques were applied in reference [21] for the optimal position of UPFC along with series FACTS to minimize the operation cost of the power system. In [22], authors proposed two evolutionary optimization techniques to optimally place and size the hybrid controller in view of reducing the active power losses.

This paper is an extension of work initially published in the Proceedings of the International Conference on Control Engineering &Information Technology [23]. In this work, we proceed with an efficient method for the identification of the proper placement of UPFC, for the purpose of improving voltage

*Hajer Jmii, University of Tunis, Tunis 1008, Tunisia, +21694014865, hajerjmii@yahoo.com

stability. We plan a series of methodological dynamic simulations to find the critical zone of IEEE 14-bus case study. The zone having the lowest voltage magnitudes is taken as the best location of UPFC. Then, we perform comparative studies to justify the placement selected. The dynamic model of UPFC and the numerical simulation method are implemented within the software EUROSTAG.

2. Necessity of FACTS devices

Consider a single machine infinite bus system interconnected through a transmission line having reactance X, as shown in Figure 1. The power transferred across the transmission line is given in equation (1) [24, 25].

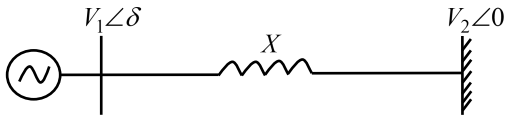


Figure 1. Single machine infinite bus system

$$P = \frac{V_1 V_2}{X}\sin\delta \quad (1)$$

V_1 and V_2 are the voltages of the synchronous generator and the infinite bus respectively where: $V_1\angle\delta$ and $V_2\angle 0°$.

The power variation is a sinusoidal function of the difference between the two voltage angles as shown in Figure 2. The maximum power that can be transmitted across the line corresponds to $\delta = 90°$ and by considering $V_1 = V_2 = V$, it is given by:

$$P_{\max} = \frac{V_2}{X} \quad (2)$$

$P_{\max}$ depends on the value of the reactance X which imposes the theoretical limit of the steady state transmitted power. However, the line resistance R creates power losses quantified by RI^2 and thus sets the practical limit of the line. The physical characteristics of the conductor could also limit the power transmission, which is known as the thermal limit.

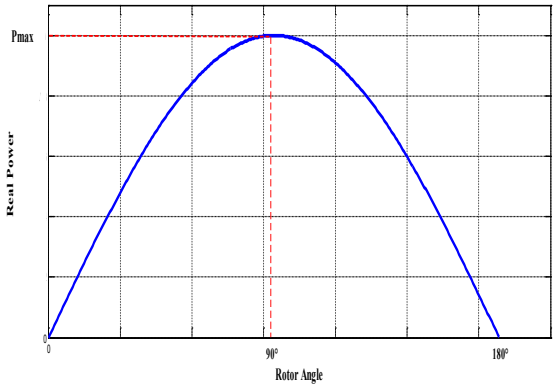


Figure 2. Power-angle curve

From equation (1), we note that the power can be controlled by two parameters; it can be enhanced by increasing the voltage magnitude of the synchronous machine and the infinite bus or by increasing the angular difference of the voltages. However, there are limits imposed on either the voltage magnitude or the angular difference, which have to be considered. The voltage variation must not exceed $\pm 5\,\%$ of the nominal voltage while the difference between the angles is limited to less than $30° - 35°$ to maintain the transient stability [26, 27].

Therefore, it is inferred that the power transmission is limited by:

- Thermal limitations.
- Physical limitations.
- Stability.

It is possible to deal with the thermal limitations by renovating the transmission lines and opting for a higher current rating conductor. Nonetheless, this solution does not necessarily guarantee voltage within the acceptable boundaries or the flow and the controllability of the power, but it is feasible by line compensation. Electro-mechanical devices used for line compensation cannot achieve rapid compensation and are prone to the wear quicker than the static equipments. The solid-state based technology, FACTS, provides the control of one or more parameters to magnify the loading capability and to develop controllability. As the current in a transmission line has the property to be controlled, it is possible to use a FACTS device for flowing power across the line during normal or disturbed conditions. This returns to the ability of FACTS devices to control all power flow parameters, namely phase angle, bus voltage and line impedance. In other words, the FACTS technology gives the possibility to maintain acceptable voltage magnitude and power flow.

3. UPFC

The UPFC is a hybrid device consisting of two voltage source converters interconnected through a common DC link capacitor, as shown in Figure 3 [28-30]. The inverter 1 is connected in shunt with the transmission line while the inverter 2 is series connected. It injects a series connected. It injects a series voltage controllable in phase and magnitude. The decomposition of this voltage leads to a quadrature component responsible for controlling the real power and an in-phase component which controls the reactive power in the line. Similarly, the injected shunt voltage is decomposed into two components. The quadrature component regulates the voltage of the DC link as for the in-phase component; it controls the exchange of the reactive power with the system.

The model of UPFC is represented in Figure 4 [30, 31]. From this scheme, we can extract the active and reactive power flows of the series and shunt converters which are expressed by the equations (3) - (6).

$$P = \frac{VV_k}{(1+K)X}\sin(\theta_{jk} + \theta) \quad (3)$$

$$Q = \frac{1}{(1+K)^2 X}[V^2 - VV_k\cos(\theta_{jk} + \theta)] \quad (4)$$

$$-\frac{1}{(1+K)^2 X}[{V_k}^2 - VV_k\cos(\theta_{jk} + \theta)]$$

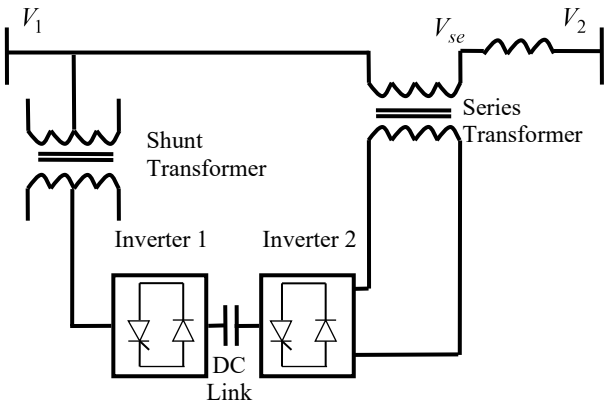


Figure 3. UPFC incorporated in a power system

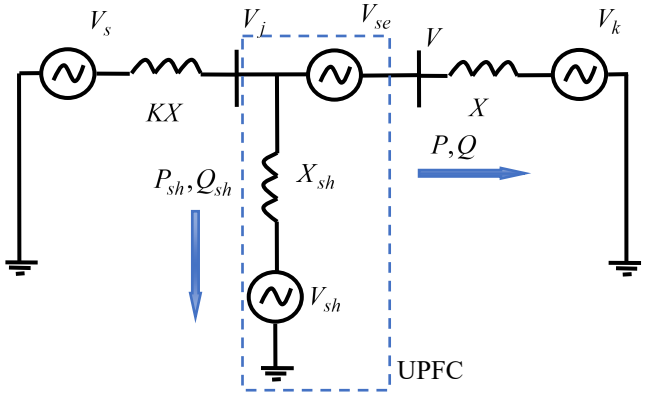


Figure 4. Model of UPFC

Where:

$$V^2 = V_j^2 + V_{se}^2 + 2V_jV_{se}\cos\phi_{se}\ \ \phi_{se} = \tan^{-1}\frac{\gamma}{\beta},\ \ \theta = \tan^{-1}\frac{\gamma}{1+\beta}$$

Similarly, the active and reactive powers of the shunt inverter are expressed by:

$$P_{sh} = \frac{V_jV_{sh}}{X_{sh}}\sin\varepsilon \quad (5)$$

$$Q_{sh} = \frac{V_j^2}{X_{sh}} - \frac{V_jV_{sh}}{X_{sh}}\cos\varepsilon \quad (6)$$

Where: $\varepsilon = \tan^{-1}\frac{\xi}{1+\eta}$ and $V_{sh}^2 = (1+\eta)^2V_j^2 + \xi^2V_j^2$

KX: Short circuit level;

$\phi(x)$: Gaussian function

γ : Relative quadrature series voltage;

β : Relative in-phase series voltage;

ξ : Relative quadrature shunt voltage;

η : Relative in-phase shunt voltage;

V : Right YPFC voltage;

V_{se} : Series voltage;

V_{sh} : Shunt voltage;

P : Real power;

Q : Reactive power;

P_{sh} : Shunt real power;

Q_{sh} : Shunt reactive power.

The modeling of the UPFC shunt part on EUROSTAG is simple; it is represented by a current injector. As for the modeling of the series part, we must open the line where we want to insert UPFC and place at its extremities two current injectors (Figure 5). The opening of the line is assured by a high reactance value.

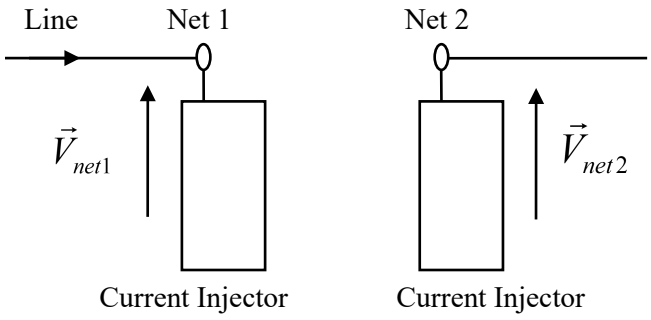


Figure 5. Model of UPFC in EUROSTAG

4. Simulation

4.1. Case Study

To test the proposed method, we used the IEEE 14-bus test system. This system contains two generators each one is equipped with voltage and speed regulators and three synchronous compensators to produce reactive power. It also has two transformers with two windings, a three-winding transformer, fifteen transmission lines and eleven loads. The data of the above mentioned system are taken from reference [31].

EUROSTAG [32] adopts the model considering the voltage behind transient reactance for representing all the generators and constant impedance to model loads and transformers.

4.2. Simulations Results

4.2.1. Identification of the Critical Zone:

The IEEE 14-bus network can be decomposed into a transmission system (69 KV) referred by « zone A » and a distribution one (13.8 KV) referred by « zone B » interconnected through three step-down transformers. To identify the critical zone, we proceeded by an incremental load increase of each zone separately. The part of the network experiencing voltage collapse rapidly is considered as the critical zone.

According to the steady state, we noted that bus 14 is the most insecure bus; it has got the lowest magnitude compared to the other buses. For this reason, we choose to adopt this bus as an index for voltage stability assessment.

Figure 6 shows the main results of increasing load. In zone A, the voltage of bus 14 underwent a progressive decrease of 0.002pu from one case to another until collapsing in the case of 49%. As for zone B, it could not withstand over than 39% of load increase where bus 14 has got the limit voltage amplitude. While the increase of the total system load leads to a severe voltage drop which ends with a collapse when reaching 25%.

Noteworthy that distribution system is the critical zone. All thanks to voltage and speed regulators of synchronous machines which reinforce transmission system stability.

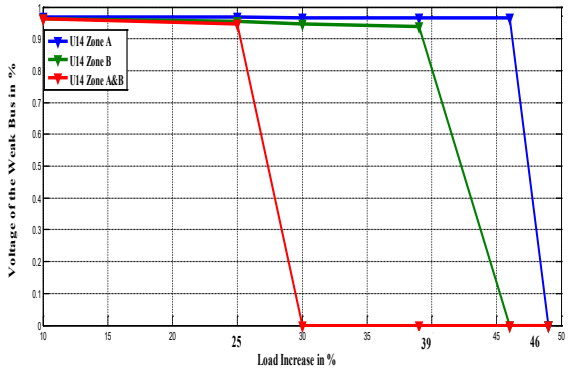


Figure 6. Identification of the critical zone of IEEE 14-Bus system.

4.2.2. UPFC Application:

Having identified the critical zone, we placed UPFC between buses 9 and 14, sized at 60MVAR for each one of its converters. Line 9-14 is close to the weakest bus so that we expect satisfying results when introducing UPFC into it. To evaluate the contribution of the controller, we rerun the system under 20 % of load increase at time t=250s.

As we can see in Figure 7, UPFC was able to considerably increase bus14 voltage amplitude just after perturbation and with well-damped oscillations compared to the uncompensated system. This result indicates the voltage support supplied by the UPFC. Note that we obtained similar results for the other buses of the network.

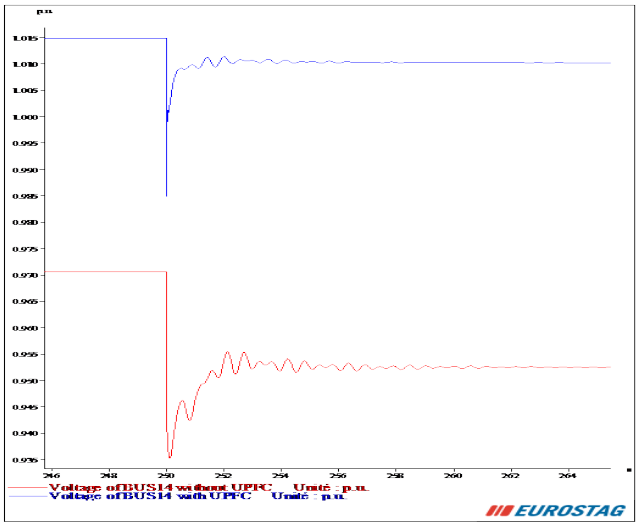


Figure 7. Temporal evolution of voltage at bus14 for 20 % of load increase with and without UPFC

The test network along with UPFC at different locations is tested now for comparison purpose. In addition to line 9-14, we choose two placements in the transmission system: lines 1-2 and 2-3 which are intended for transmitting the highest amount of reactive and active powers. Thereby, enhancing power flowing in these lines can be effective in improving voltage stability of the whole network. Figure 8 and Figure 9 illustrate the temporal evolution of voltage at bus1 and bu2 respectively. It should be noted that the implementation of UPFC in line 1-2 has considerably damped voltage oscillations compared to the other cases. While, this location had no effect on the voltage magnitude as shown in Figure 10. UPFC in line 9-14 gave best voltage levels at all network buses than when it is incorporated in line 2-3, where we obtained the lowest voltage magnitudes.

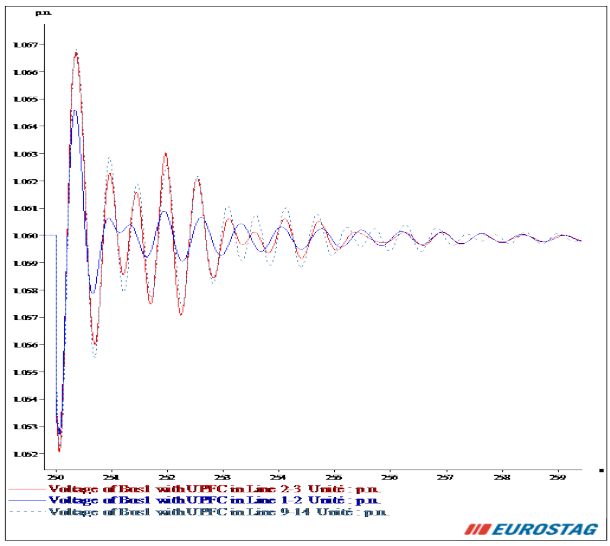


Figure 8. Temporal evolution of voltage at bus 1 for 20% of load increase with different locations of UPFC.

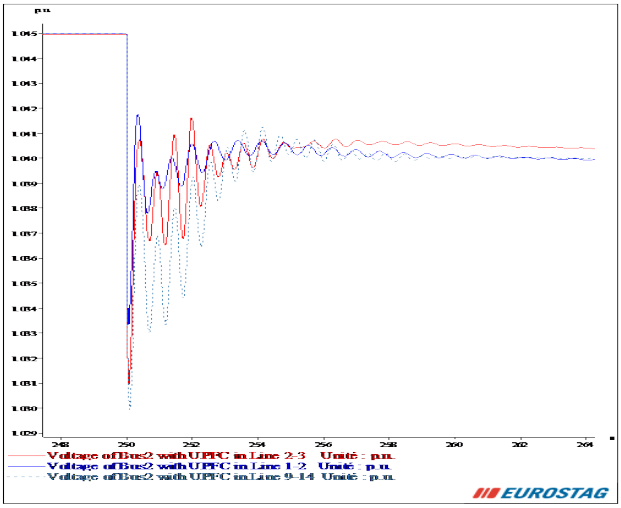


Figure 9.Temporal evolution of voltage at bus 2 for 20% of load increase with different locations of UPFC.

Figure 11 illustrates the active power losses of the system. It is clearly observable that the active power losses had the lowest value in the case of UPFC inserted through line 9-14, it is about 13.5MW. However, a large amount of active losses is detected with line 2-3 placement and about 14 MW in line 1-2 case.

Now, we evaluate the impact of UPFC location on the loadability margin of the power system. From Table 1, we can note that the contribution of the controller in improving the maximum loading capability was very meaningful when installing in the distribution system. In fact, it was able to increase the loadability margin in zone A from 48% to 59% and from 39% to 70% in zone B. As for the whole network, it became able to

maintain a stable state up to 33.5% of load increase while its maximum loadability was about only 25% before compensation. Nevertheless, when introduced in the transmission system, UPFC was less effective in increasing loadability margin. That was observable notably in line 2-3, where only the loading capacity of zone A; reached a higher percentage compared to the case without UPFC.

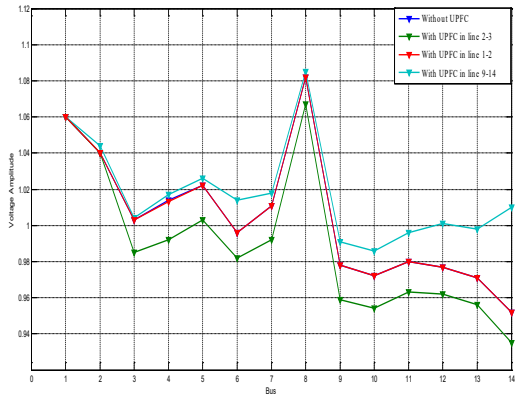


Figure 10.Voltage profiles of IEEE 14-Bus system for 20% load increase with different locations of UPFC

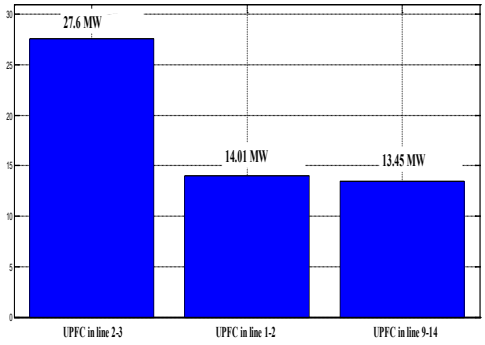


Figure 11. Active power losses for 20% load increase with Different UPFC Locations.

Table 1. Load ability Margins with Different Locations of UPFC.

Network Zone	Loadability Margin in %			
	Without UPFC	With UPFC in line 2-3	With UPFC in line 1-2	With UPFC in line 9-14
Zone A	48	52	55	59
Zone B	39	39	39	70
Zones A&B	25	25	32	33.5

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

5. Conclusion

In this paper, the impact of the proper installation of UPFC on improving voltage stability has been investigated. The placement of UPFC is done by using a numerical simulation based method. The identification of the critical zone of the system can determine the best location of the hybrid device. Then, the performance of the proper placement is compared to different selected locations in the system for justification purpose. It was found that UPFC incorporated through line 9-14 in the critical zone, helps in improving voltage magnitude, reducing active power losses as well as increasing maximum loading capacity of the system. However, the integration of the controller in the transmission system gave rapid and well- damped oscillations. Thereby, the optimal placement of the UPFC is highly dependent on the objective of the integration.

References

[1] H. Jmii, A. Meddeb, S. Chebbi, "Efficiency Limits of SVC in Improving Voltage Stability,"3rd Inter, conf, GEEE, April, 2016, Hammamet, Tunisia.

[2] B. Singh, "Applications of FACTS Controllers in Power Systems for Enhance the Power System Stability: A State-of-The-Art," International Journal of Reviews in Computing, Vol. 6, 15th July 2011.

[3] A. Gupta, P.R. Sharma," Optimal location and setting of SVC and TCSC devices using non-dominated sorting particle swarm optimization," Electric Power Systems Research, 2009.

[4] P. Kundur., "Power system stability and control,"McGraw-Hill, Inc., 1993.

[5] S. Amara, H. Hsan, "Power system stability improvement by FACTS devices: a comparison between STATCOM, SSSC and UPFC," 1st International Conference on Renawble Energies and Vehicular Technology, 2012.

[6] M.A. Abido, "Power System Stability Enhancement using FACTS Controllers: A Review," The Arabian Journal for Science and Engineering, vol.34, no1B, Apr 2009.

[7] Gyugyi L., " Unified power flow control concept for flexible AC transmission systems," Proceedings of IEE Part-C IEE Proceedings of Generation, Transmission and Distribution, 139:323–331, 1992.

[8] J. Singh, H. Qazi, M. Ghandhari, " Load curtailment minimization by optimal placement of Unified Power Flow Controller," International Transactions on Electrical Energy systems, 2016.

[9] V.G. Mathad, B.F. Ronad, S.H. Jangamshetti, "Review on Comparison of FACTS Controllers for Power System Stability Enhancement," International Journal of Scientific and Research Publications, vol. 3, Issue 3, March 2013

[10] X.P. Zhang, Ch. Rehtanz, B. Pal, " Flexible AC Transmission Systems: Modeling and Control," 2nd Edition, Springer, Feb 2012.

[11] Q.H. Wu, Z. Lu, M.S. Li, T.Y. Ji, " Optimal Placement of FACTS Devices by A Group Search Optimizer with Multiple Producer," *IEEE trans*, 2008.

[12] S. Gupta, R. Tripathi, R. Shukla, " Voltage stability improvement in power systems using FACTS controllers: State-of-the-art review," https://www.researchgate.net/publication/251987213, 2010.

[13] M. Benaissa, S. Hadjeri, S.A. Zidi, "Impact of PSS and SVC on the Power System Transient Stability," Advances in Science, Technology and Engineering Systems Journal, pp. 562-568 , Vol. 2, No. 3, 2017.

[14] M. Sanjari, O. Mousavi, G. Gharehpetian, "Assessing the risk of blackout in the power system including HVDC and FACTS devices," European Transactions on Electrical Power, 2012.

[15] H. Shaheen, G. Rashed, S. Cheng, "Application of differential evolution algorithm for optimal location and parameters setting of UPFC, 19:911–932, 2009.

[16] A. Jupta, P.R. Sharma,"Optimal Placement of FACTS Devices for Voltage Stability Using Line Indicators," *IEEE trans*, 2012.

[17] H. Wang,"Applications of modelling UPFC into multi-machine power systems," IEE Proceedings online no. 19990170, DOI: 10.1049/ip-gtd:19990170, 1999.

[18] A. Ray, J.O. Chandle , "Voltage Stability Enhancement during Excess Load Increments through Optimal Location of UPFC Devices," IEEE

International Conference on Technological Advancements in Power & Energy, 2015.

[19] A. Jupta, P.R. Sharma, " Static and transient stability enhancement of power system by optimally placing UPFC (Unified Power Flow Controller)," *IEEE trans*, 2013.

[20] S. Lee, J. Liu, C. Chu, " Modelling and locating unified power-flow controllers for static voltage stability enhancements," International Transactions on Electrical Energy systems, 2012.

[21] B. Bhattacharyya, V.K. Gupta, S. Kumar, "UPFC with series and shunt FACTS controllers for the economic operation of a power system," Ain Shams Engineering Journal, 2014.

[22] H.I. Shaheen, G.I. Rashed, S.J.Cheng, "Optimal Location and Parameters Setting of Unified Power Flow Controller Based on Evolutionary Optimization Techniques," *IEEE trans*, 2007.

[23] H. Jmii, A. Meddeb, S. Chebbi, "Proper Placement of UPFC for the Improvement of Voltage Stability of IEEE-14 Bus System," 4th International Conference on Control Engineering & Information Technology (CEIT) Tunisia, Hammamet, December 2016.

[24] R. Mathur, R. Varma, "Thyristor-Based FACTS Controllers for Electrical Transmission Systems," J. Wiley & sons.

[25] S. Rath, B. Sahu, P. Dash, "Power System Operation and Control using FACTS Devices," International Journal of Engineering Research & Technology (IJERT), Vol. 1, Issue 5, July 2012.

[26] I. Khan, M. Mallick, M. Rafi, M. Mirza, "Optimal placement of FACTS controller scheme for enhancement of power system security in Indian scenario," Journal of Electrical Systems and Information Technology 2, pp.161–171, 2015.

[27] Aboreshaid, S.,Billinton, R., "Probabilistic evaluation of voltage stability," IEEE Trans.PowerSyst.14(February(1)), pp. 342–348, 1999.

[28] S. Morsli, "Robust Control of UPFC in MultiMachines Power Network," M.S. thesis, Dept. Electron. Eng., Oran Univ., Algeria, 2012.

[29] A. Meddeb, H. Jmii, S. Chebbi, " Comparison of UPFC, TCSC and SVC for Improving Voltage Stability," 3rd Inter, conf, ACECS , March, 2016, Hammamet, Tunisie.

[30] M. Farrag, G. Putrus, "An on-line training radial basis function neural network for optimum operation of the UPFC," EUROPEAN TRANSACTIONS ON ELECTRICAL POWER, 21, pp.27–39, 2010.

[31] F. Milano, "Power System Modelling and Scripting," Springer, ETSII, University of Castilla - La Mancha 13071, Ciudad Real Spain 2010.

[32] Eurostag, Eurostag Software Release Notes, Tractebel-EDF, Release 5.1, Dec 2010.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 267-271 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

2x1 Microstrip Patch Array Antenna with Harmonic Suppression Capability for Rectenna

Nur Aisyah Amir*, Shipun Anuar Hamzah, Khairun Nidzam Ramli

Faculty of Electrical and Electronic, Universiti Tun Hussein Onn Malaysia, 86400, Malaysia

A R T I C L E I N F O

Article history:
Received: 30 October, 2017
Accepted: 02 December, 2017
Online: 23 December, 2017

Keywords :
Key 1 Rectangular inset feed antenna
Key 2 PBG structure
Key 3 Band pass

A B S T R A C T

This paper is an extension of work originally presented in 2016 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE). A 2×1 microstrip patch array antenna integrated with photonic bandgap (PBG) and stubs is designed and analyzed. The performance of the PBG and stubs structure are explained and analyzed in terms of the elimination of the resonance at the harmonic frequencies of the antenna. The proposed antenna is designed on FR-4 substrate with thickness of 1.6 mm and operated at 2.45 GHz frequency suitable for rectenna design application. From the simulated result, the first harmonic frequency (5.4 GHz), the second harmonic frequency (6.6 GHz) and the third harmonic frequency (7.8 GHz) are successfully suppressed. For instance, the radiation to the forward of the stubs-PBG antenna is suppressed at more than 15 dB at the second and third harmonic frequencies.

1. Introduction

The term rectenna is commonly denoted as rectifying antenna at microwave power transmission system. It is a combination of the antenna and high efficient rectifier circuit. The ability of the wireless power transmission system is very important and it depends on the ability of each process; the conversion of signal DC-to-RF at the transmitter, at the transmission and at the conversion of signal RF-to-DC at the receiver. Hence, the receiving antenna is one of the main elements in the rectenna system development.

One of the most popular types of receiving antenna that is used for rectenna application is the microstrip patch antenna [1 - 4]. Some of the advantages of employing this type of antenna is low profile, low cost, light-weight and suitable if integrated with RF devices. However, it also exhibits disadvantages such as the excitation of surface waves that exist in the substrate layer. Surface waves are exceptionable because when a patch antenna radiates, a part of the total available radiated power becomes trapped along the surface of the substrate. It will diminish the total available power for radiation to space wave, and the harmonic frequency is produced. In addition, for arrays antenna, surface waves have a significant impact on the mutual coupling between array elements [5]. Moreover, in rectifying circuit, diode that used to convert the RF signal to DC power supply also generates the unwanted radiation of harmonic frequencies. Thus, the rectenna cannot operate well and will disturb the overall performance of the system.

Several design methods have been proposed in the past in order to overcome these unwanted generated harmonic problems. One of the most popular methods is called the photonic band gap (PBG) or electromagnetic band gap (EBG) [6 - 8] and defected ground structure (DGS) [9 - 11]. Several other techniques for instance antenna with slit and stub structure [12], circular sector patch antenna [13], slot antenna and notch antenna [14], Low Pass Filter (LPF) [15 - 17] also have been discussed earlier as a one of the method to suppress the harmonic frequencies.

According to that, a harmonic suppression microstrip array patch antenna is proposed in order to improve the system performance. In this paper, the 2x1 microstrip patch array antenna integrated with stubs and two-dimensional (2-D) PBG pattern in the ground plane beneath the square patch, is proved experimentally and the capability of the stubs and PBG structure for the suppression of the harmonics are discussed [18].

2. Antenna design procedure

2.1. Numerical method

The structure of the microstrip patch array antenna with PBG substrate (PBG antenna) is presented in Figure 2. A 2x1 array antenna is fabricated on the FR-4 substrate with dimensions of 143 x 52 x 1.6 mm^3 and the relative permittivity of $\varepsilon r = 4.4$. The source signal is fed directly to the antenna by the 50Ω microstrip line with

*Nur Aisyah Amir: Universiti Tun Hussein Onn Malaysia, 01127865561, aisyahamir91@gmail.com

width length (Wd2) = 4.13 mm. The rectangular inset feed antenna dimensions are calculated prior to designing the array antenna by using [19]:

$$w=\frac{c}{2f_0\sqrt{\frac{(\varepsilon_r+1)}{2}}} \quad (1)$$

$$\varepsilon_{eff}=\frac{\varepsilon_r+1}{2}+\frac{\varepsilon_r-1}{2}\left(1+12\frac{h}{w}\right)^{-\frac{1}{2}} \quad (2)$$

$$L_{eff}=\frac{c}{2f_o\sqrt{\varepsilon_{eff}}} \quad (3)$$

$$\Delta L=0.412h\frac{(\varepsilon_{eff}+0.3)\left(\frac{w}{h}+0.264\right)}{(\varepsilon_{eff}-0.258)\left(\frac{w}{h}+0.8\right)} \quad (4)$$

$$L=L_{eff}-2\Delta L \quad (5)$$

Where:

W = Width of the patch
L = Length of the patch
Ɛr = Dielectric constant of substrate
Ɛreff = Effective dielectric constant of substrate
c = Speed of light
f_0 = Frequency reference
h = Height of dielectric substrate

There are advantages of PBG structure which are increasing the antenna's performance in terms of the gain and radiation pattern and also react as a filter to suppress the harmonic frequency. The PBG structure is designed based on the formula [20]:

$$\lambda=\frac{c}{f_o\sqrt{\varepsilon_r}} \quad (6)$$

$$r=\lambda/8 \quad (7)$$

$$a=\lambda/2 \quad (8)$$

After calculation, the 9 x 5 squares with the dimensions 7 x 7 mm^2 are etched in the ground plane at the period (g_b) of 15 mm as shown in Figure 1.

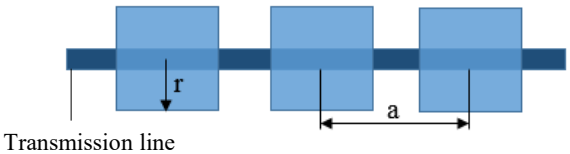
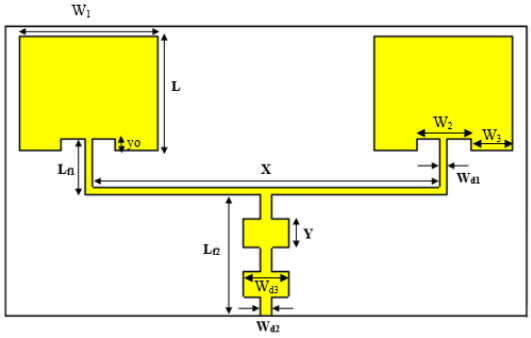


Figure 1: PBG structure at ground plane

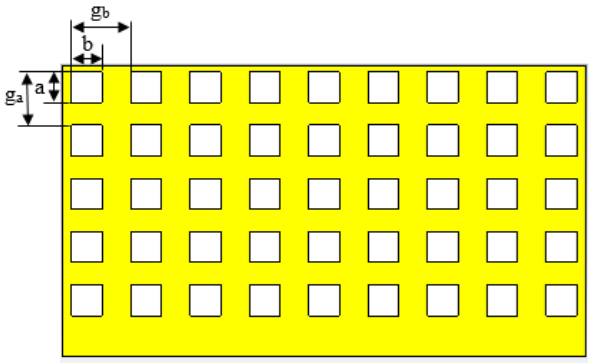
2.2. Design simulation

By using CST microwave software, the rectangular inset feed antenna is designed and simulated to get the operating frequency at 2.45 GHz. Some enhancement are made to obtain the best performance of the antenna. After completing the single element geometry, a 2x1 array antenna is designed. The parameters of the feed line network can be chosen by setting the feed line impedance to 50 Ω (Z1 = 50 Ω), which splits into two 100 Ω (Z2 = 100 Ω) [21]. Similarly, some optimization is done in order to get the greatest achievement of the array antenna.

After done with the conventional microstrip patch array antenna design, the 9 x 5 PBG square are etched at the ground plane. However, some size optimization of PBG square size is done in order to match with the 2x1 array patch antenna design.



(a)



(b)

Figure 2: 2x1 microstrip patch array antenna with stubs and PBG structure (a) view from the front (b) view from the back.

Table 1: Dimension of antenna (unit: mm)

Parameter	Dimension	Parameter	Dimension
L	29.16	**W_{d3}**	11.3
W_1	37.61	**X**	88.32
W_2	10	**Y**	5
W_3	13.805	**a**	7
L_{f1}	5.5	**b**	7
L_{f2}	15.7	**g_a**	15
W_{d1}	0.71	**g**	15
W_{d2}	4.13		

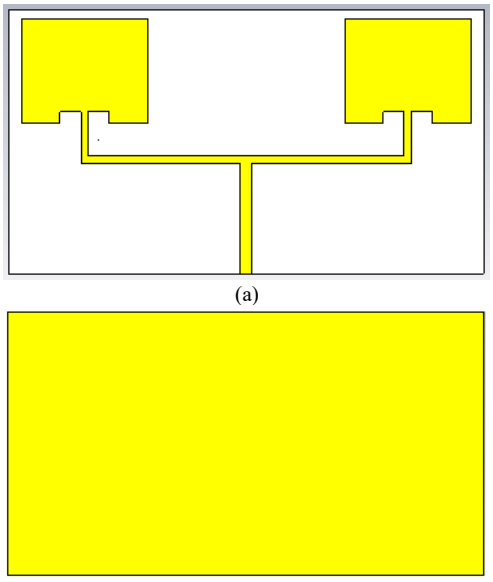
3. Result and discussion

This proposed antenna is divided into three parts which are normal array patch antenna, PBG array antenna and array patch antenna with PBG and stub structure. Each part is discussed and analyzed according to their characteristics.

3.1. Normal array patch antenna

The normal patch antenna without PBG structure shows in Figure 3 is simulated and measured for the comparison of the basic

characteristics. The resonant frequency of the fundamental is adjusted to be 2.45 GHz for the antenna. The simulated and measured S-parameter (return loss) result of the array antenna without stubs and PBG structures are shown in Figure 4. As can see, the return loss is at -13 dB and has a few harmonic frequencies which are not good for an antenna.



(a)

(b)

Figure 3: 2x1 rectangular inset feed array antenna (a) front view (b) back view

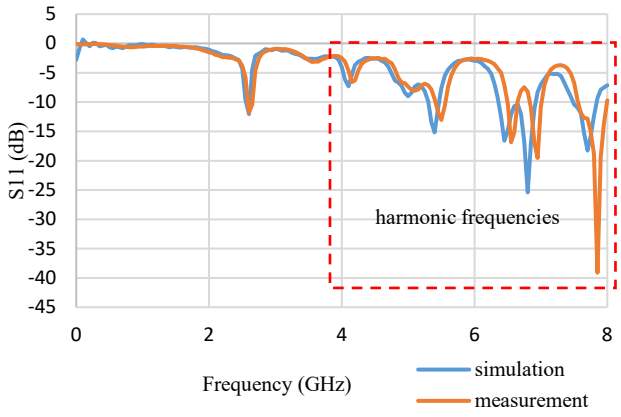


Figure 4: Simulated and measured result for S-parameter (return loss) of 2x1 rectangular inset feed array antenna

3.2. PBG structure

As can see in Figure 4, there are a lots of harmonic frequencies produces above 4 GHz. Hence, PBG structure is beneath at the ground in order to suppress the harmonic frequencies. At the same time, the gain of the antenna also can be improved. Figure 5 shows the PBG structure at ground plane. This arrangement of the PBG lattice produces the band pass characteristics of the transmission parameter at more than -20 dB from 4.6 GHz to 8 GHz shown in Figure 6 (a) when the uniform microstrip line is fabricated on this substrate instead of the patch antenna. The effect of the PBG can be seen by observing the S-parameter (return loss) in Figure 6 (b). It shows that the harmonic frequencies above 4.6 GHz is successfully suppressed when the S-parameter of PBG structure is compare with return loss of 2x1 rectangular inset feed array antenna with and without PBG structure. While Figure 7 shows the comparison of return loss of 2x1 rectangular inset feed array antenna with PBG structure only between simulation and measurement.

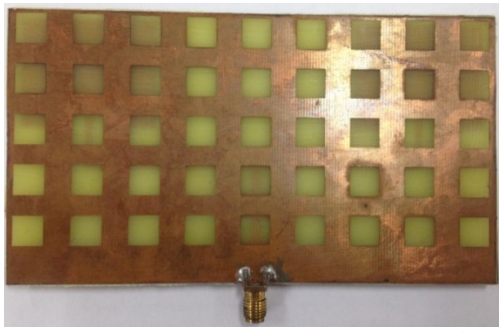
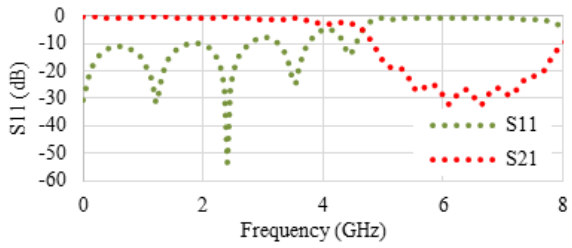
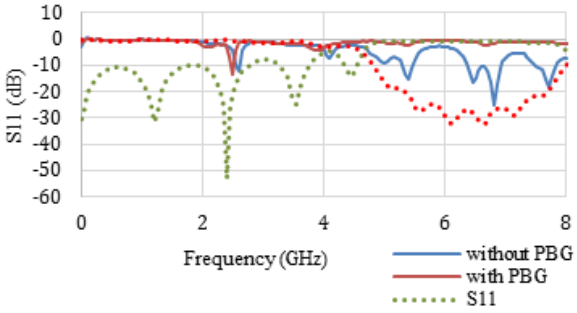


Figure 5: Prototype of PBG structure at ground plane



(a)



(b)

Figure 6: The effect of PBG structure (a) S-parameter of PBG square structure (b) comparison S-parameter of PBG square with return loss of 2x1 rectangular inset feed array antenna with and without PBG structure

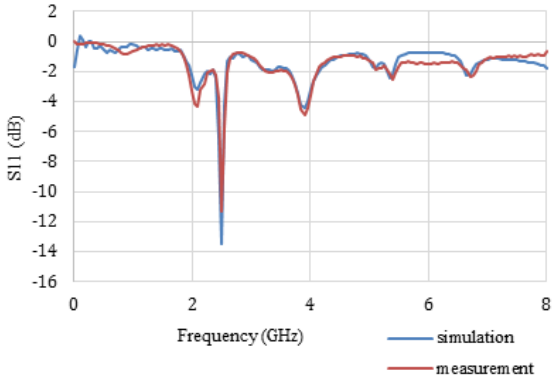


Figure 7: Simulated and measured result for return loss of 2x1 rectangular inset feed array antenna with PBG structure only

3.3. Array patch antenna with PBG and stub structure

The final design is combination of normal array patch antenna with PBG and stubs structure as shown is Figure 8. The S-parameter (return loss) of the array antenna with PBG and stubs structure, and the array antenna without PBG and stubs structure are shown in Figure 9.

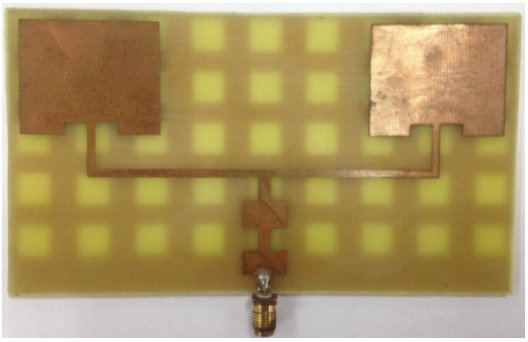


Figure 8: Prototype of 2x1 rectangular inset feed array antenna with PBG structure and stub

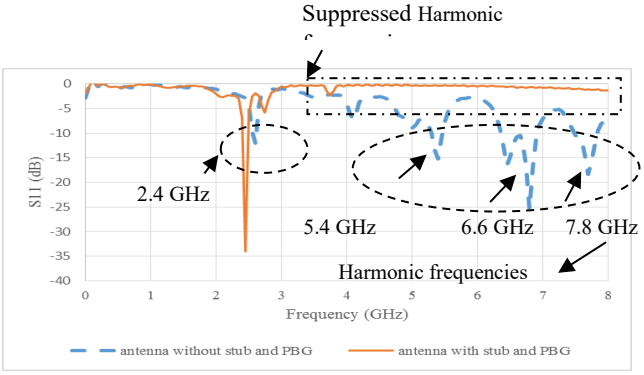


Figure 9: Measured S-Parameter (return loss) of the proposed array antenna. Graph of the comparison for antenna with and without stubs and PBG

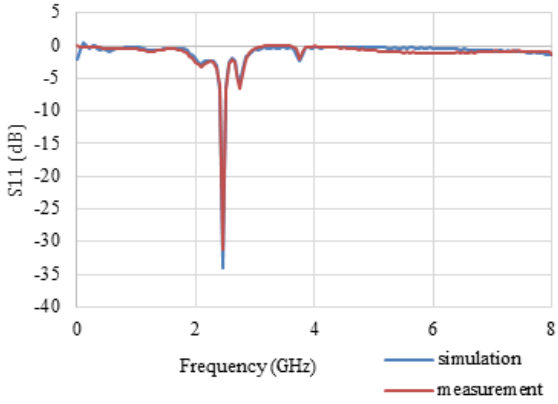


Figure 10: Comparison of return loss between simulation and measurement for 2x1 array patch antenna with PBG and stub structure

Figure 9 shows, a comparison between conventional array antenna with array antenna with PBG and stubs structure indicates that the harmonic frequencies are indeed suppressed. According to the theoretical design, the return loss is observed at 2.45 GHz. However from measurement result, the S-Parameter (return loss) is observed at 2.6 GHz for conventional antenna while for stubs-PBG antenna is at 2.45 GHz. Through observation, the antenna performances are increased. The return loss is revised from -12.05 dB to -34.99 dB after the stub and PBG is added. Besides, based on to the characteristic of PBG, it also can suppress the surface wave in the band-gap. A periodic structure surrounding the patches in addition to one underneath the patches is essential for effective suppression of the harmonic and surface wave. Figure 10 depicts the comparison between simulated and measured return loss of the proposed array antenna.

3.4. Gain

As mentioned before, besides suppressing the harmonic frequency, PBG structure also can increase the performance of the antenna in terms of gain. As can be seen in Figure 11 (a) and (b), there are comparison antenna's gain between simulation and measurement for normal array patch antenna and PBG array patch antenna.

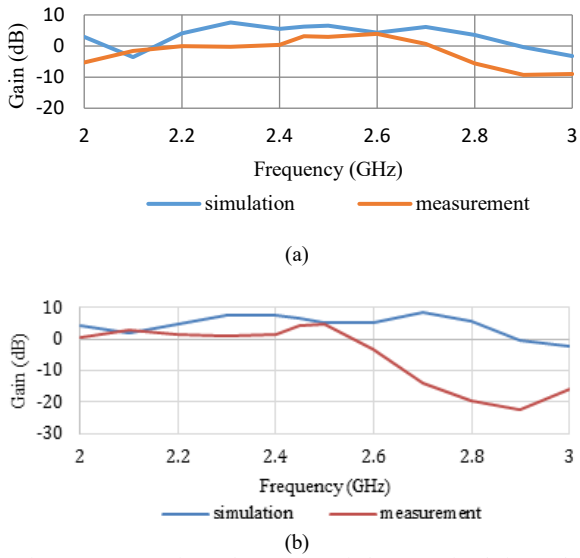


Figure 11: Comparison of antenna's gain between simulation and measurement (a) normal array patch antenna (b) PBG array patch antenna

Figure 11 (a) shows the gain comparison between simulation and measurement for normal array patch antenna. From the graph, the gain for 2.45 GHz is 6.263 dB for simulation while 3.15 dB for measurement. The gain increases to 6.436 dB when the PBG structure is added as shows in Figure 11 (b). While for measurement, the gain is 4.31 dB.

In Table 2, due to the effective suppression of the harmonic and the surface wave, and changes in the current distribution, the antenna array with stub and PBG has higher gain compared to the normal array antenna. The gain increase about 82.41% which is from 3.156 dB to 5.757 dB. While for normal array antenna and array antenna with PBG structure, the result shows a minimal gain increase by 36.66%. Thus, it shows that with the addition of stub structure at the feed line, the gain's performance will improve more. Moreover, the mixed current distribution is useful for diminishing the mutual coupling between the patches and decreasing the maximum side-lobe level.

Table 2: Measured antenna gain base on structure added

Antenna structure	Gain (dB)
Array antenna without PBG and stub	3.156
Array antenna with PBG	4.313
Array antenna with PBG and stub	5.757

4. Conclusion

2x1 microstrip patch array antenna has been presented in this effort. The proposed structure is integrated with two-dimensional PBG pattern in the ground plane under the square patch and stubs. The effectiveness of the PBG and stubs structure for harmonic suppression are analyzed. The simulation and measurement works verified that the combination of the array antenna with PBG and stubs is efficient in eliminating unwanted frequencies. Moreover, the performances of the antenna in term of gain also increase.

Acknowledgment

The authors would like to thank the Centre Graduate Study UTHM for supporting this study. Much appreciation also goes to the Electromagnetic Compability Lab (EMC Centre), Universiti Tun Hussein Onn Malaysia for providing the measurement facilities.

References

[1] G. A. Vera, A. Georgiadis, A. Collado, and S. Via, "Design of a 2.45 GHz rectenna for electromagnetic (EM) energy scavenging," Radio and Wireless Symposium (RWS), 2010 IEEE (pp. 61-64). IEEE, 2010.

[2] U. Olgun, C. C. Chen and J. L. Volakis, "Wireless power harvesting with planar rectennas for 2.45 GHz RFIDs,"Electromagnetic Theory (EMTS), 2010 URSI International Symposium on (pp. 329-331). IEEE, 2010

[3] H. Takhedmit, B. Merabet, L. Cirio, B. Allard, F. Costa, C. Vollaire and O. Picon, "A 2.45-GHz low cost and efficient rectenna," Antennas and Propagation (EuCAP), 2010 Proceedings of the Fourth European Conference on (pp. 1-5). IEEE, 2010

[4] J. Zhang, Y. Huang and P. Cao, "Harvesting RF energy with rectenna arrays," Antennas and Propagation (EUCAP), 2012 6th European Conference on (pp. 365-367). IEEE, 2012

[5] D. M. Nashaat Elsheakh, M. F. Iskander, E. A.-F. Abdallah, H. A. Elsadek, and H. Elhenawy, "Microstrip Array Antenna with New 2D-Electromagnetic Band Gap Structure Shapes to Reduce Harmonics and Mutual Coupling," Prog. Electromagn. Res. C, vol. 12, pp. 203–213, 2010. http://onlinewww.jpier.org/PIERC/pierc12/17.09112008.pdf

[6] M. N. Aktar, M. S. Uddin, M. R. Amin, and M. A. Ali, "Enhanced gain and bandwidth of patch antenna using ebg substrates," Int. J. Wirel. Mob. Networks, vol. 3, 2011. https://www.researchgate.net/profile/Mst_Nargis_Aktar/publication/49965985_Enhanced_Gain_and_Bandwidth_of_Patch_Antenna_Using_EBG_Substrates/links/0912f50973187d409b000000.pdf

[7] Y. Horii and M. Tsutsumi, "Harmonic control by photonic bandgap on microstrip patch antenna," Microwave and Guided Wave Letters, IEEE, 9(1), 13-15, 1999

[8] W. Zhang, J. Mao, X. Sun, R. Qian, and D. Zhang, "Microstrip patch antenna array on ground with circular PBG," Microwave and optical technology letters, 41(2), 127-130, 2004.

[9] V. S. Kushwah and G. S. Tomar, "Size reduction of microstrip patch antenna using Defected Microstrip Structures," Proc. - 2011 Int. Conf. Commun. Syst. Netw. Technol. CSNT 2011, pp. 203–207, 2011.

[10] R. Dua, H. Singh, and N. Gambhir, "2.45 GHz Microstrip Patch Antenna with Defected Ground Structure for Bluetooth," … J. Soft Comput. Eng. …, no. 6, pp. 262–265, 2012. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.2510&rep=rep1&type=pdf

[11] M. S. Ghaffarian, G. Moradi, and R. Zaker, "Harmonic suppression slot antenna using rectangular defected ground structure," Electrical Engineering (ICEE), 2011 19th Iranian Conference on (pp. 1-4). IEEE, 2011.

[12] R. A. Rahim, F. Malek, S. F. W. Anwar, S. L. S. Hassan, M. N. Junita, and H. F. Hassan, "A harmonic suppression circularly polarized patch antenna for an RF ambient energy harvesting system," CEAT 2013 - 2013 IEEE Conf. Clean Energy Technol., pp. 33–37, 2013.

[13] F. J. Huang, T. C. Yo, C. M. Lee, and C. H. Luo, "Design of circular polarization antenna with harmonic suppression for rectenna application," IEEE Antennas Wirel. Propag. Lett., vol. 11, pp. 592–595, 2012.

[14]R. A. Rahim, M. N. Junita, S. I. S. Hassan, and H. F. Hassan, "Harmonics suppression circular polarization elliptical shape microstrip patch antenna," Proc. 2014 2nd Int. Conf. Technol. Informatics, Manag. Eng. Environ. TIMEE 2014, pp. 147–150, 2015.

[15] M. Han, S. Jung and H. Sohn, "High efficient rectenna using a harmonic rejection low pass filter for RF based wireless power transmission," Wireless Communications Systems (ISWCS), 2014 11th International Symposium on (pp. 423-426). IEEE, 2014.

[16] D. Gangwar and R. L. Yadava, "Design and analysis of a pentagonal rectenna," Signal Processing and Integrated Networks (SPIN), 2014 International Conference on (pp. 654-658). IEEE, 2014.

[17] J. Zhang, Y. Huang and P. Cao, "A wideband cross dipole rectenna for RF wireless harvesting," Antennas and Propagation (EuCAP), 2013 7th European Conference on (pp. 3063-3067). IEEE, 2013.

[18] N. A. Amir, S. A. Hamzah, K. N. Ramli, and M. Esa, "2× 1 microstrip patch array antenna with harmonic suppression capability" Applied Electromagnetics (APACE), 2016 IEEE Asia-Pacific Conference on(pp. 272- 276). IEEE, 2016.

[19] Joshua, M.S, Mayank, M, Prafull, and S, "Design and Optimization of Microstrip Patch Antenna," Int. J. Emerg. Trends Technol. Comput. Sci., 2013.

[20]I. Garde, M. J. Yábar, and C. Del Río, "Simple modeling of DGS to design 1D-PBG low-pass filters," Microw. Opt. Technol. Lett., vol. 37, no. 3, pp. 228–232, 2003. http://onlinelibrary.wiley.com/doi/10.1002/mop.10878/full

[21] Y. S. H. Khraisat, "Design of 4 elements rectangular microstrip patch antenna with high gain for 2.4 GHz applications," Mod. Appl. Sci., vol. 6, no. 1, pp. 68–74,2012.http://ccsenet.org/journal/index.php/mas/article/view/12540/9541

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 272-274 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Cyclical Wave Bolt for Sound Waves in a Gas Stream

Vladimir Arabadzhi*

Institute of Applied Physics, Division of Geophysical Research, 603950 Russian Federation

ARTICLE INFO

Article history:
Received: 31 October, 2017
Accepted: 02 December, 2017
Online: 23 December, 2017

Keywords :
Cyclical Wave Bolt, Sound in Gas Stream, Rotary Switch, Impeller, Bundle of Tubes, Echelon of Waveguides Bundle of Tubes, Automotive Silencer

ABSTRACT

This article is devoted to the problem of blocking the propagation of sound in a gas stream. This can be useful in the problems of pipeline acoustics and the design of automotive silencers, when it is necessary to ensure the opacity of the boundary (cross section of the pipe) for sound simultaneously with the free (ideally) flow through this boundary. The tool for solving this problem is the rapid periodic overlapping of waveguide (gas-conducting) channels in a system of parallel waveguides. The design of the proposed device assumes the following technological requirements: high mechanical rigidity of the elements, high accuracy of their manufacture and high rotor speed. Fulfillment of these conditions allows creating an effective device for blocking sound in the gas flow. Such a device may have small wave dimensions (with respect to the wavelength of sound) and a small expenditure of mechanical power to push gas through it very wide frequency range of silencing.

1. Introduction

In practice, sometimes it is necessary to place a filter in the path of the gas flow (with velocity v) in the pipe (and simultaneously sound with speed c in the tube as an acoustic waveguide). This filter must pass the flow of gas and block sound waves on its output. Such situations are encountered in problems of the acoustics of pipelines and also in the problem of designing exhaust silencers for cars with internal combustion engines. Next we will call such a device a silencer. The area of the inlet and outlet cross-section of the silencer is assumed equal S_0 , the average velocity of the flow at the inlet and outlet of the silencer is V_0 (see Figure 1). The time-average gas pressures at Q the outlet and inlet of the silencer are, respectively, p_0 and $p_0 + \Delta p$, where $\Delta p > 0$. The quality of the silencers is characterized by the following parameters:

(a) the degree

$$Q_1 = P_{out} / P_{in} \tag{1}$$

of sound pressure suppression (of course preferably $Q_1 << 1$);

(b) the frequency range

$$\omega_{\min} \le \omega \le \omega_{\max} \tag{2}$$

of sound suppression, preferably when

$$Q_2 = \omega_{\min} /(\omega_{\max} - \omega_{\min}) << 1 \tag{3}$$

(c) the wave dimensions

$$Q_3 = \ell \omega_{\min} / c \tag{4}$$

of the device (where ℓ is the geometric size of the silencer, c - is a sound speed, preferably when $Q_3 << 1$);

$$Q_4 \approx 2\Delta p / \rho V_0^2 << 1 \tag{5}$$

(d) the relative loss of power to propel the gas through the silencer (ρ - gas mass density, preferably when $Q_4 << 1$). For a silencer, this condition is necessary

$$Q_1 = P_{out} / P_{in} \tag{6}$$

Always this combination is desirable (in addition to (1))

$$Q_2 << 1,\ Q_3 << 1,\ Q_4 << 1. \tag{7}$$

Violation of the conditions (7) (or one of these conditions) is usually (for systems with parameters that are constant in time) a payment for the fulfillment of the condition (6). Widely known are the following three approaches [1] in the design of such devices: when narrow-band ($Q_2 > 1$) absorption is permissible

*Vladimir Arabadzhi: Institute of Applied Physics (RAS), 603950, Nizhny Novgorod, Ulianov st. 46, Russia, v.v.arabadzhi@appl.sci-nnov.ru

and one can be satisfied with resonant absorption of active suppression of sound; when large silencer dimensions are permissible ($Q_3 \geq 1$) and one can be satisfied by a Helmholtz resonator with a resonant frequency below the minimum sound frequency $\omega_{\min}$ in the flow at low energy costs to propel the gas; when the high energy expenditure for pushing the gas through a classical rigid plate, perforated with small holes ($Q_4 \sim 1$).

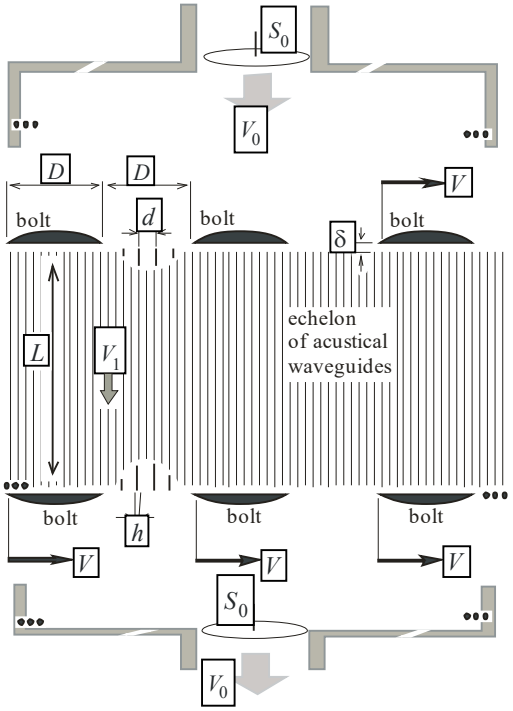


Figure 1. Two-dimensional version of cyclical wave bolt

2. Cyclical Wave Bolt

The above solutions [1] for blocking sound in a stream (with or without sound absorption) were based on systems with time-constant parameters. We now consider a system with parameters modulated in time [2]. We consider the infinite echelon (see Figure 1) of waveguides with rigid vertical plane walls, located at a characteristic distance

$$d >> h \tag{8}$$

from each other, where h is characteristic thickness of walls. Two horizontal systems of rigid thin bands (wave bolts) of the width D, separated from each other by the distance

$$D >> d \tag{9}$$

move horizontally (to the right) with the velocity V near the upper and lower edges of the echelon of above mentioned waveguides of length L. A sound wave is incident on top of this system and an excessive pressure Δp is applied to push the gas flow through this system. The gap between bolts and ends of waveguides is limited by condition $\delta < d$. We call such a parametric structure a cyclical wave bolt (CWB). From above, a sound wave falls on the echelon of waveguides, periodically blocked by wave bolts. If the conditions

$$L / c = D / V \tag{10}$$

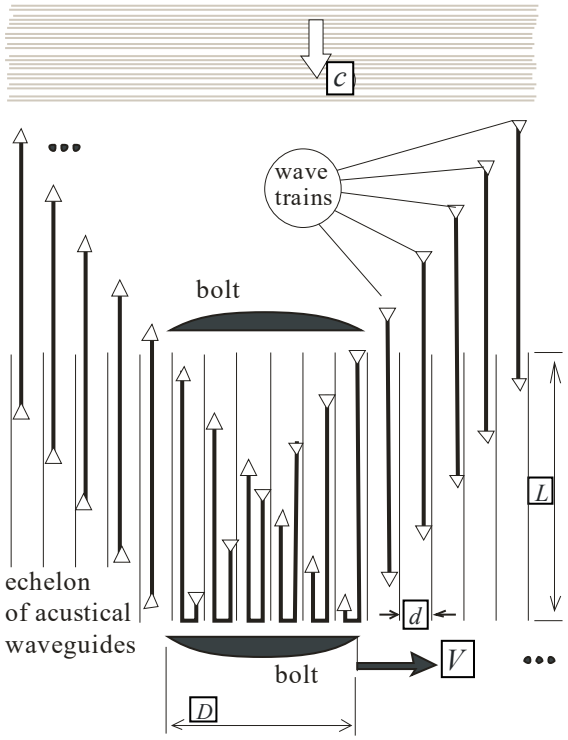


Figure 2. Space diagram of seven incoming wave trains, moving inside the echelon of waveguide channels

of synchronism are satisfied, the incident wave either reflects upward from the upper running bolts, or from the lower ones (Figure 2). Due to condition (10), each waveguide is open at both ends for a certain period of time D / V and is closed by wave bolts at both ends during a subsequent time interval of the same duration D / V. During the first interval (with open bolts), the sound wave fills this waveguide, during the subsequent interval (with closed bolts), the wave turns completely to reverse motion (due to full reflection from closed wave bolts). Thus, any perturbation propagating in a gas at a speed c cannot overcome CWB. On the other hand, for a gas flow this system is transparent. The only limitation of its transparency for the flow is the friction of the gas against the walls of the waveguides, which leads to additional energy loss. Waves, reflected not from the bolts, but from the open ends of the waveguides, do not leave the CWB. In principle, the operation of the system will be the same if the bolts are made stationary, and the train vertical walls is made to move at speed V.

3. Three-Dimensional Version of Cyclical Wave Bolt

Figure 3 shows a three dimensional cylindrical analogue (axially symmetric device) of the infinite two dimensional system described above.

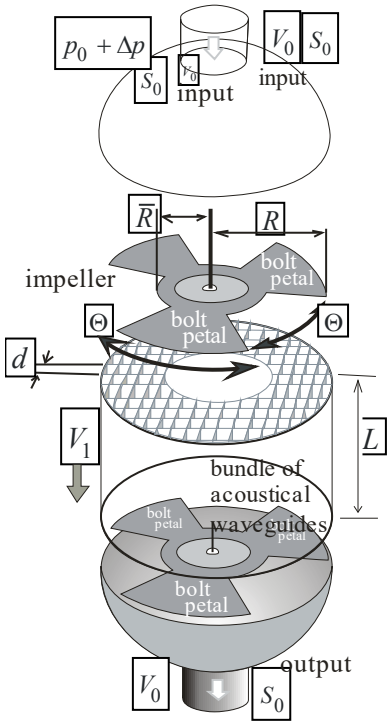


Figure 3. Three-dimensional version of silencer.

An infinite spatially periodic two dimensional bolt system is replaced by rotating (with angular frequency Ω) acoustically rigid (for sound in gas) impeller or rotary switch with N petals (bolts) with angular sectors $\Theta = \pi / N$ of the petals and the same angular sector between adjacent petals. Both ends of each waveguide of length L with a characteristic transverse dimension d (with rigid walls of thickness $h << d$) are open during time intervals $\tilde{t} + 2Tn \le t \le \tilde{t} + T + 2Tn$ and are closed at intervals $\tilde{t} + T + 2Tn \le t \le \tilde{t} + 2T + 2Tn$ of time by bolts (or petals of rotating impeller), where $\tilde{t}$ is a certain initial moment, determined by the point of placement of this waveguide and the phase of rotation of the impeller, where

$$T = \Theta / \Omega = L / c . \qquad (11)$$

The relation (11) is analog of synchronism condition (10). A bundle of tubes (waveguides) is located between the radii $\bar{R}$ and R. Any wave train that gets into the waveguide of length L in time $T = L / c$ has time to fully rotate to exit this waveguide. The maximum rigidity of structural elements and the maximum rotor speed are determined by modern state of technology level. We note that the gas velocity $V_1 = \chi V_0$ inside the waveguide tubes must ensure that the Mach number is small, i.e. $V_1 / c << 1$. By decreasing the value $\chi = S_0 / \pi(\bar{R}^2 - R^2) < 1$ we simultaneously reduce the Mach number inside the waveguides and the relative power loss Q_4. The rotor of the impeller is driven by an electric motor with a stabilized rotational frequency Ω. The frequency of incident waves in the flow is bounded from above only by the condition of one dimensionality of sound propagation in the rigid tube, i.e. its characteristic transverse scale d, i.e. $\omega_{\max} << c / 2\pi d$. Minimum cross-sectional dimension d of the tubes is limited below by the condition of a fairly low energy expenses to push gas trough the waveguides. The maximum cross-sectional dimension of the tubes is bounded above by the condition $d / V << L / c$. That the switching time of one tube is small compared with the propagation time of sound along the tube. Minimal length $L = c\Theta / \Omega$ of the tubes-waveguides is limited from below by the maximum achievable frequency Ω of the rotation of impeller. The minimum gap δ between the impeller (wave bolts, petals) and the ends of the tubes is limited from below by the manufacturing accuracy and stiffness.

Conclusions

Silencer considered can provide a low level of sound at the output, less engine power loss, a wider bands of sound suppression at smaller wave dimensions or, in other words, joint fulfillment of conditions (6), (7). The loss of power to push gas through acoustic waveguides can be reduced by shortening L and by increasing the rotational speed Ω of the impeller (until too much power is required to rotate it). The input and output of device can be interchanged, it acts symmetrically: it does not pass through itself the perturbation propagating in gas at the speed of sound.

References

[1] H. Fuchs, Applied Acoustics: Concepts, Absorbers, and Silencers for Acoustical Comfort and Noise Control: Alternative Solutions – Innovative Tools – Practical Examples, Springer Science & Business Media, 2013.

[2] V. Arabadzhi, Solutions to Problems of Controlling Long Waves with the Help of Micro-Structure Tools, Bentham Science Publishers Ltd., 2011.

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 275-284 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

A Multilingual System for Cyberbullying Detection: Arabic Content Detection using Machine Learning

Batoul Haidar*,[1], Maroun Chamoun[1], Ahmed Serhrouchni[2]

[1]Saint Joseph University, Lebanon

[2]Telecom ParisTech, France

ARTICLEINFO

Article history:
Received: 12 November, 2017
Accepted: 03 December, 2017
Online: 23 December, 2017

Keywords:
Cyberbullying
Machine Learning
Natural Language Processing
Arabic Natural Language Processing

ABSTRACT

With the abundance of Internet and electronic devices bullying has moved its place from schools and backyards into cyberspace; to be now known as Cyberbullying. Cyberbullying is affecting a lot of children around the world, especially Arab countries. Thus, concerns from cyberbullying are rising. A lot of research is ongoing with the purpose of diminishing cyberbullying. The current research efforts are focused around detection and mitigation of cyberbullying. Previously, researches dealt with the psychological effects of cyberbullying on the victim and the predator. A lot of research work proposed solutions for detecting cyberbullying in English language and a few more languages, but none till now covered cyberbullying in Arabic language. Several techniques contribute in cyberbullying detection, mainly Machine Learning (ML) and Natural Language Processing (NLP). This journal extends on a previous paper to elaborate on a solution for detecting and stopping cyberbullying. It first presents a thorough survey for the previous work done in cyberbullying detection. Then a solution that focuses on detecting cyberbullying in Arabic content is displayed and assessed.

1. Introduction

As stated in the Abstract, this work is an extension to the author's previous work [1] which was published in the EMS2016 conference.

Children and teens were exposed to physical bullying before the abundance of internet, computers and handheld devices. Nowadays, bullying is performed using cyber technology. Around 50% of the youth of America are suffering from cyberbullying [2]. As for the Arab world: 20.9% of middle-school adolescents report bullying in UAE, 31.9% in Morocco, 33.6% in Lebanon, 39.1% in Oman and 44.2% in Jordan [3].

Awareness of cyberbullying is rising in the Arab world. Arab News [4] declares that they heard of numerous cyberbullying incidents in Saudi Arabia. One of the rare reports [5] on cyberbullying states that 60% of Gulf Countries' youth openly admit the presence of cyberbullying amongst their peers. This study also states that only quarter the predators online do bully their victims offline. Which means that internet have encouraged three quarters of the predators to bully others, while they wouldn't have considered bullying face to face.

Most of the previous research dealing with cyberbullying focused on the effects of cyberbullying. It was concerned with helping the victims after cyberbullying attacks, mainly from the psychological aspect. Less work was directed towards implementing technical methods to detect and stop an ongoing cyberbullying attack, or even to prevent cyberbullying attacks while or before they happen [6]

After the Introduction, this paper furnishes a background that covers the technologies underlying cyberbullying detection in Section 2. Then in section 3 presents a survey of all existing literature in both cyberbullying detection and multilingual techniques. It concludes at the end of section 3 that there is no work done on detecting cyberbullying attacks in Arabic language. Therefore, in section 4 a system for cyberbullying detection in Arabic content is proposed. This system is designed for preventing cyberbullying attacks, by detecting and stopping them. It uses Natural Language Processing (NLP) to identify and process Arabic words. Then Machine Learning (ML) techniques are used to classify bullying content. In section 5 the results obtained from the system are displayed and analyzed. Conclusions and future works are finally stated in section 6.

2. Background

2.1. Cyberbullying

Cyberbullying is defined as the use of Internet, cell phones, video game systems, or other technologies to send or post text or

* Batoul Haidar: batoul.haidar@net.usj.edu.lb

images intended to hurt or embarrass another person or group of people [7]. Some examples of cyberbullying include sending mean or threatening messages, tricking someone into revealing personal or embarrassing information and sending it to others, sending or forwarding private messages to others, sharing explicit pictures with others without consent, starting rumors via text message or online or creating fake online profiles on websites such as Facebook, Myspace, Twitter, etc. to make fun of people [7]. There are several categories of cyberbullying as stated by [8] and [9] :

- Flaming: starting a form of online fight.
- Masquerade: where there is a bully pretending to be someone else, in order to perform malicious intents.
- Denigration: sending or posting gossip to ruin someone's reputation.
- Impersonation: Pretending to be someone else and sending or posting material to get that person in trouble or danger or to damage that person's reputation or friendships.
- Harassment: Repeatedly sending profane and cruel messages.
- Outing: Publishing someone's embarrassing information, images or secrets.
- Trickery: Talking someone into revealing secrets or embarrassing information for the sake of sharing it online.
- Exclusion: Intentionally and cruelly excluding someone from an online group.
- Cyberstalking: Repeated, intense harassment and denigration that includes threats or creates significant fear.

Bullying and cyberbullying leave mental and physical effects on both the bully (predator) and the victim. Cyberbullying is more severe than physical bullying due to the fact that it is wider, public, and the victim has nowhere to escape. Victims of cyberbullying reported emotional, concentration, and behavioral issues, as well as trouble getting along with their peers. These victims were more likely to report frequent headaches, recurrent stomach pain, and difficulty sleeping. One out of four students revealed that they felt unsafe at school. They were also more likely to be hyperactive, have conduct problems, abuse alcohol, and smoke cigarettes [10].

As stated previously, cyberbullying detection systems mainly use Machine Learning and Natural Language Processing techniques in the course of detection. Thus before dwelling in the previous literature in the area of cyberbullying detection, a thorough background for ML, NLP and other techniques collaborating in the process of cyberbullying detection is displayed.

2.2. Machine Learning

Machine Learning (ML) is defined as the ability of a computer to teach itself how to take a decision using available data and experiences [11]. Available Data is known as *Training Data*. Decisions to be taken in ML might be a classification or prediction for new objects or data. The computer classifies a new piece of data by depending on learning algorithms. When the training data is labeled, i.e. classified by human experts, the algorithms depending on these labeled data are called Supervised Learning algorithms [12].

In cyberbullying detection, there could be a corpus of data manually labeled (or classified) by people as either containing harm or not, as described in Section 3. When the training data is unlabeled, the algorithms depending on these non-labeled data are called Unsupervised Learning algorithms [12]. They teach themselves how to classify the data based on similarities and differences between data. When both supervised and unsupervised learnings are combined together by using labeled and unlabeled data, to get the most out of both ways, the algorithm is known as Semi-supervised Learning algorithm [12].

When ML is used to classify a certain object as belonging or not belonging to a certain category, the machine learner is called *Binary Classifier* [13]; for example in spam email filtering, ML algorithms are used to take decisions against incoming emails and label them as either spam or not spam. A second type is when the task given to the classifier, is to match a certain object against several classes or categories, then it is called *Multi-Class Classifier*. A third type might be predicting a value for an object and is called *Regression*, i.e. predicting a priority level for an incoming email.

There are several ML algorithms available, from which the most frequently used in relevance to the scope of research of this paper will be mentioned.

- Naive Bayes: A probabilistic supervised learning method [14] that mainly calculates the probability of an item belonging to a certain class, depending on metrics obtained from training data. Naïve Bayes algorithm was used in some cyberbullying detection research, such as in [15] and [14]. It was used for sexual predation detection.
- Nearest Neighbor Estimators: A simple estimator [16] that uses distance between data instances, in order to map a certain instance to its closest distance neighbor, thus estimating the class of this instance, this algorithm was used in [17] and [18].
- Support Vector Machine (SVM): Also a supervised algorithm. SVM is a binary classifier that assumes a clear distinction between data samples. It tries to find an optimal hyper plane that maximizes the margin between the classes [12]. SVM was used in many cyberbullying detection systems [19] , [20].
- Decision Tree: Decision tree learners use a set of labeled data, thus they are supervised learners. Decision trees classify data using a command and conquer approach. The Trees compose of leaves and arcs. Each leaf of the tree represents a classification class and each arc represents a feature inspected from training data [21]. The C4.5 algorithm is an implementation of decision trees. It was employed in cyberbullying detection by [22] and [17].

ML algorithms are widely incorporated in cyberbullying detection systems as seen in Section III, due to the huge amount of

data incorporated in social networking platforms, which makes it hard to be processed by human power, thus comes the need for a machine learner.

2.3. Natural Language Processing

Natural Language Processing (NLP) is the collection of techniques employed to make computers capable of understanding the natural unprocessed language spoken between humans by extracting grammatical structure and meaning from input [23].

NLP is a common branch of all of Linguistics, Artificial Intelligence and Computer Science [24]. NLP research started with Machine Translation in the late 1940s [25]. Then it spread to other areas of application, such as information retrieval, text summarization, question answering, information extraction, topic modeling, opinion mining [26], optical character recognition, finding words boundary, word sense disambiguation, and speech recognition [24].

According to Chandhana [27], NLP can be divided into three areas; Acoustic – Phonetic: where acoustic knowledge studies rhythm and intonation of language; i.e. how to form phonemes, the smallest unit of sounds. Phonemes and phones are aggregated into word sounds. Phonetic knowledge relates sounds to the words we recognize. Morphological – Syntactic: Morphology is lexical knowledge which studies sub words (morphemes) that would form a word. Syntactic knowledge studies the structural roles of words or collection of words to form correct sentences. Semantic - Pragmatic: Semantic knowledge deals with the meaning of words and sentences, while pragmatic knowledge deals with deriving sentence meanings from the outside world or outside the content of the document [28].

2.4. Sentiment Analysis

Sentiment analysis is a textual analysis technique. Sentiment analysis is used to define the polarity, subjectivity or features of a certain text. By polarity we mean defining whether a certain content is positive, negative or neutral [29]. Both Machine Learning and Natural Language Processing techniques are incorporated in Sentiment Analysis.

Sentiment analysis is currently used in detecting the opinions of social media users. Opinions are studied in several areas such as marketing and politics

2.5. Common Sense Reasoning/Sentic Computing

Common sense is the knowledge (usually acquired in early stages of life) concerning social, political, economic and environmental aspects of the society we live in. Common sense usually varies among different cultures and is built from layers of learning experiments we acquire throughout life [30].

Computers do not have common sense reasoning by nature, but there is a research field, known by Sentic Computing [31], that aims towards transforming computers into machines that could feel. This field of research is a multidisciplinary approach to opinion mining and sentiment analysis, which uses common sense reasoning and web semantics in order to inspect the emotions, not just the opinions from certain text. The term 'Sentic' derives from the Latin 'Sentire', the root of words like sentiment and sensations.

2.6. Performance Measures

Some evaluation metrics were adapted in Information Retrieval (IR) and then extended to other fields of computer science such as ML. These evaluation metrics are used as measures for the performance of IR and ML systems. The most widely used metrics are Recall, Precision, Area under the ROC and F-Measure.

- Recall is the proportion of returned documents (or values) which are relevant (or correct) $Rl \cap Rt$ [32] out of all relevant documents returned and not returned [33]. The metric is also known as Sensitivity of a system.

$$R=(Rl\cap Rt)/Rl \quad (1)$$

- Precision is the proportion of returned documents (or values) which are relevant (or correct) Rl∩Rt [32]. The metric is also known as Accuracy of a system.

$$P = (Rl \cap Rt)/Rt \quad (2)$$

- F-Measure, proposed by van Rijsbergen in 1979, is a weighted harmonic mean of precision and recall. It is a combination between Recall and Precision metrics, which was introduced to overcome the negative correlation between Precision and Recall [34].

$$F\beta=(1+\beta^2)PR/(\beta^2P+R) \quad (3)$$

- F1 is a special case of F- measure with β =1. β Is a parameter to control the balance between Recall and Precision where $0\leq \beta \leq \infty$. When β is set to 0, it implies giving no importance to recall, when β tends to ∞ then no importance is given to precision, and when $\beta = 1$ then Recall and Precision will be given equal importance [35].

$$F1= 2PR/(P+R) \quad (4)$$

P: Precision
R: Recall
Rt: Returned documents
Rl: Relevant documents

- ROC (Receiver Operating Characteristics) graph [36] –or area- is an analysis technique which had been used originally in medical diagnosis, to be later adopted in Machine Learning evaluation. ROC depicts the tradeoff between True Positive and False Negative rates. ROC areas are categorized roughly according to the values:

0.9-1 = excellent (A)
0.8-0.9 = good (B)
0.7-0.8 = fair (C)
0.6-0.7 = poor (D)
0.5-0.6 = fail (F)

3. Previous Work

As stated previously, the research efforts in cyberbullying covered several areas, including the detection of online bullying when it occurs; reporting it to law enforcement agencies, Internet service providers and others for the purpose of prevention and awareness;

and identifying predators with their victims. No effort was directed towards detecting cyberbullying in Arabic language. This section focuses on the previous work done in the areas of cyberbullying detection, ML, feature extraction and cross language transliteration and translation since those are the main techniques used in the implementation of cyberbullying detection systems.

3.1 Cyberbullying Detection

Most of the research done in detecting cyberbullying constituted of either a filtration software or ML techniques. A filtration software has to be employed by social networking platforms, in order to automatically delete or shade profane words [37] [38] [39]; but the filtration method is first limited by its inability for detecting subtle language harassment and second it has to be manually installed [40].

Most work other than filtration methods employs ML techniques, where old corpora of comments or conversations is crawled, whether from Facebook, Twitter, Formspring (a platform similar to Facebook, popular between teens) or even real conversations of sexual attackers [41]. These corpora are used to feed ML algorithms responsible for detecting cyberbullying attacks by building a classification rule from the training set. The obtained classification rule classifies the testing set comments. Such work was done in [22] where the authors crawled data from Formspring and used the Amazons Mechanical Turk [42] for labelling comments. Then they used the learning methods from the Waikato Environment for Knowledge Analysis (WEKA) toolkit [43] to teach and test the model for classifying comments.

The problem of detecting subtle language cyberbullying attacks was tackled by Dinakar et al [6]. They depended on commonsense reasoning in the detection of cyberbullying content. As an example of the commonsense they used: they considered comments of wearing makeup when subjected on Males might indicate the presence of harassment. They built their datasets from both YouTube and Formspring for both training and testing. They used NB, JRip, J48 and SVM for text classifications. For feature sets they used general features, such as a list of unigrams or profane words, tf-idf weighting scheme, Ortony Lexicon for negative affect, Part-of-speech tags for commonly occurring bigrams, and Label Specific Features including frequently used forms of verbal abuse.

The weighting scheme tf-idf is the product of term frequency and the inverse document frequency in the dataset. It involves multiplying term frequency (tf), that represents the number of times a term occurs in a document, by inverse document frequency (idf), which varies inversely with the number of documents to which a word is assigned [44].

Nahar, Li and Pang [40] employed the tf-idf weighting scheme for building features. In addition they built a network composing of bullies and their victims. The network was used to rank the most active predator and its target. In [45] Dinakar et al. stated that detecting profane language cyberbullying is easier than detecting sarcasm and subtle language attacks. Chayan and Shylaja of [46] enhanced the performance of the cyberbullying detection model by 4%, through looking for comments directed towards peers by using Supervised ML and Logical Regression models. However they didn't detect sarcasm comments. Dadvar et al. [47] state that incorporating user context such as the user's history as a feature for training the cyberbullying detection model increases accuracy of classification; however they didn't include sarcasm detection in their system.

SVM was also used by Yin et al. [48] for classifying posts as containing harassment or not. They used documents from CAW 2 dataset, which included posts from Kongregate, Slashdot and Myspace. For feature selection they incorporated several features;

- *Local features,* they used tf-idf.
- S*entiment features,* such as 1- grams, 2- grams and 3- grams, and they also captured second pronouns.
- *Contextual features* in which they studied both the similarity of a post to other neighboring posts and the cluster of posts neighboring around a certain harassment post.

Capturing sentiment features only didn't perform well so they compared the performance of their system by mixing features. Tf-idf performed better than n-grams and foul words, however, combining tf-idf with contextual and sentiment features achieved an additional enhancement in results in Precision, Recall and F1-measure. A similar work was done by Dadvar et al. [47], who built their feature space from Content-based, User based and Context based features. They also proved that incorporating contextual features such as gender information from the user's profile enhances cyberbullying detection.

Other research efforts were focused around social network profiles, such as [17]. They presented a methodology to detect and associate fake profiles on Twitter social networks to real users. This system had been capable of linking the owners of a fake account on Twitter to a real account for one or more students in a school class; this was a case of a real cyberbullying incident. The system was devised by collecting features from tweets then analyzing the features using various supervised ML techniques included in WEKA. Afterwards the performance among these techniques was compared on True Positive Ratio (TPR), False Positive Ratio (FPR) and Area Under ROC Curve (AUC). Bayzick, Kontostathis and Edwards [49] proposed the BULLYTRACER software which detects cyberbullying in chat rooms 58.63% of the time.

Chen et al. [50] proposed a new model for detection which they named as the Lexical Syntactic Feature-based (LSF) model; it achieved a precision of 98.24% and recall of 94.34%. Their model calculated both a post and a user's offensiveness depending on the ratio of offense appearing in a user's posts. This model detects "strong profanity" in online posts by using lexical analysis methods such as Bag of Words; and subtle language harassment to which the authors referred as "weak profanity". Then the model uses semantic analysis and NLP techniques to analyze the context of sentences by studying the grammatical relations among words. This research was an extension to the work presented in [51] for cyberbullying posts filtration.

Most of the research in cyberbullying did not give importance for the distinction between cyberbullying and cyberagression, but

for Hosseinmardi et al. [52]. They proposed a definition for cyberbullying which is the ***repetition*** of harmful actions using electronic devices over a certain period of time. They stated that most of the work previously done in detecting cyberbullying was actually focusing on detecting cyberagression. They define cyberagression as a single instance of harmful action that if repeated over time would be considered as cyberbullying. They also demonstrated that a Linear SVM classifier can significantly improve the accuracy of identifying cyberbullying to 87%. In addition they incorporated using features other than text such as images for better detection of cyberbullying.

Another study on cyber aggression was done by Nakano et al [53]. They analyzed anonymous and non-anonymous questions and answers from ask.fm. Their study shows that anonymous questions tend to be more aggressive than non-anonymous ones. They also showed that replies to anonymous questions tend to be less aggressive than replies to questions from known profiles.

Potha and Maragoudakis [54] stressed on a window of time in order to study the textual patterns of previous conversations in order of predicting the upcoming actions of a predator. They incorporated time series modelling in their research in addition to SVM for features selection. SVD (Singular Value Decomposition) [55] was used for feature reduction and DTW (Dynamic Time Warping) [56] for matching time series collections.

Fuzzy Logic and Genetic algorithms were also used in cyberbullying detection [57], where a new system was proposed using those two methods. This system's performance was compared across precision, recall and F1 measures. The system achieved better in Accuracy, F1-measure and Recall than previous fuzzy classification methods with 0.87, 0.91 and 0.98 respectively.

3.2 Arabic Language

Work related to Arabic language is scarce due to the complex morphological nature of Arabic. Arabic language is used by around 300 million Arabs around the world, mainly Muslims. It is a script language which is read and written from right to left and it constitutes of an alphabet of 28 letters. Vowels in Arabic are represented by special punctuation marks called Diacritics [58]. There are three variations for Arabic languages going on together. The Classical Arabic which is the language of the Islamic manuscripts -such as the Quran and prayers- and Arab people until Mid-20th century. The Modern Standard Arabic (MSA) which is the formal language used nowadays in schools and news, and it is known by all Arabs. Finally comes the dialects, which are accents for the Arabic language, usually used informally between people. There are around 10 dialects, one for every country - or group of countries [59], [60]. Arabic dialects imply a difference in meanings of words between different countries. We might find some words that are considered profane in one country, while good or ordinary in others, for example the word "Yetqalash" in Yemen is a compliment while in Morocco it is an offensive word [61].

Arabic language is a challenging and complex language due to its nature, where Arabic words do not include capitalization [60]. The morphologic nature of Arabic inflicts a lot of ambiguity, and the Arabic corpus is very scarce. Arabic language is ranked the 7th around the world, and its use over internet is growing vastly [59], thus arose the research interest in Arabic language fields.

An extensive search was performed on available articles and publications and no previous work for cyberbullying detection in Arabic language texts and comments was found. But some papers in the fields of ML and NLP were found. The previous work done in Arabic deals with text preprocessing or text classification.

Ghaleb Ali and Omar [62], proposed a key phrase extraction method that combines several key phrase extraction methods with ML algorithms. The output from the key phrase extraction methods is used as features to the ML algorithms. The ML algorithm in turn classifies the feature as either a key phrase or not. They compared their results by using three ML algorithms: Linear Logistic Regression [63], SVM and Linear Discriminant Analysis [64]. They have proved that SVM gives the best results in key phrases extraction among the three algorithms.

Some work had been done in Arabic named entity extraction, such as Named Entity Recognition for Arabic (NERA) [65] to identify proper names in Arabic documents. NERA used a white-list of named entities and corpora compiled from various sources; its performance was measured across recall, precision and F1. The results were satisfactory; 86.3% 89.2% 87.7% respectively for person named entities.

Filtering for spam emails written in Arabic and English was done by El-Halees [66] on pure English, pure Arabic and mixed collections of emails. Several ML techniques were used, including SVM, NB, k-Nearest Neighbor (k-NN) [67] and Neural Networks. The performance of the system was measured across all three variations and SVM was proved to be best in pure English environment. The system performed less on pure Arabic emails, due to the inflictive nature of Arabic Language. The authors also proved that stemming Arabic words enhances the performance of the classifiers, where NB performed best with 96.78% Recall and 92.42% F1-measure.

Other than emails there are also attempts for detecting spam in social networks, such as Twitter. Such work was done by El-Mawass and Alaboodi in [68]. They elaborated a system to detect spam in Arabic tweets. Their system achieved significant accuracy, precision and recall measures.

Sentiment analysis is one of the text classification categories. Sentiment analysis classifies a certain text as positive, negative or neutral [69]. Sentiment analysis was done by Hamouda [70] on Facebook comments written in Arabic. They built a corpora from 6000 comments sampled from Facebook, preprocessed this corpora, and then applied classifications to determine the sentiment behind the comment. Three classifiers were used: SVM, NB and Decision trees. The best performance was achieved by SVM with 73.4% accuracy. Another attempt for sentiment analysis was done in [71] for Arabic Tweets, their special contribution was in handling Arabizi and dialects. They incorporated NB, SVM and k-NN for classification and the best accuracy was approached by NB.

In [72], Duwairi detected sentiments from dialectical Arabic texts. Two methods were applied for detection. First, by translating dialectical words into MSA, then detecting according to MSA

lexicon. Second method was by detecting dialectical lexicon. NB and SVM classifiers were used to detect both negative and positive polarities. The results obtained showed improvement in Precision, Recall and F-measure upon translating into MSA.

Sentiment analysis was applied on Arabizi also by Duwairi et al [69]. In their system they first converted Arabizi text into Arabic by using their own rule based method. They labeled their data using their crowdsourcing tool [71] and then applied SVM and NB for classification. A comparison between SVM and NB showed that SVM outperformed NB. However better results were achieved when they first eliminated neutral entries from the dataset.

Arabic tweets in Saudi Arabia were analyzed by Alhumoud, Albuhairi and Altuwaijri [73]. They analyzed the tweets using a hybrid approach. Their hybrid approach composed of building a classifier and training it using a one-word dictionary. They compared the results obtained from both their hybrid approach and the supervised learning approach. Two classifiers were used from WEKA, NB and SVM. Their results showed the outperformance of the hybrid approach.

A significant research effort was done on stemming for Arabic language. Stemming is a text preprocessing technique. In stemming words are truncated to obtain their roots [74]. Several stemmers for Arabic are available, including rule-based stemmers such as Khoja's [75] and light stemmers. Light stemmers blindly remove letters from words –affixes and suffices – without prior knowledge of roots [76]. Stemmers are either monolingual or multilingual. Gadri and Moussaoui [77] elaborated a multilingual stemmer. Their stemmer is Language independent and it used the n-gram technique. This stemmer segments words into bigrams, then statistical measures are used to reach the best root. This stemmer was tested against English, French and Arabic. The best success rate (94%) was achieved in small Arabic Datasets. In large datasets, the best results were for English (86, 50%) and the worst for Arabic (67, 66%).

4. Proposed System

As stated previously, one of the purposes of this journal is presenting a solution for the problem of cyberbullying in both English and Arabic languages. As seen in the previous section, there is some work done for detection in English, but none in Arabic Language. Proving the hypothesis that Arabic cyberbullying can be detected was a challenge. Thus in the first stage of the system the focus was on detecting cyberbullying in Arabic language.

Since the proposed system employs ML, a dataset had to be prepared to be used for training and testing the system. Two toolkits were tested for ML, Dataiku DSS and WEKA. The decision was to use WEKA toolkit because it supports Arabic language.

4.1 Data Preparation

In order to train and test the system, a huge amount of data had to be obtained. Thus, the choice was to scrap data from both Facebook and Twitter. This choice was inflicted by the fact that those two social media portals are the most widely used by the Arab nation, especially Arab youth.

For the data acquirement phase two custom tools were built. Those tools were implemented to scrap data from social networks, one for Facebook and the latter for Twitter. Twitter Scrapper was written using PHP while Facebook scrapper incorporated python. Both scrappers connected to a mongo dB server, where all downloaded tweets and messages were stored.

In the process of getting the data from Twitter, the Twitter Scrapper searches for Tweets according to a given location and perimeter. A huge database was accumulated using this tool. In order to get the tweets from the Arab countries we focused on the center of the Middle East Region as a location, and selected an area of 10,000 kilometers radius in each run. Thus the tweets were collected from Lebanon, Syria, Gulf Area and Egypt mainly. The size of the tweets database summed up to 4.93 GB.

Collecting Facebook data was harder, due to the restrictions imposed by Facebook security and privacy measures. In Facebook the data source had to be specified beforehand. It is specified by page IDs. The aim was for pages of public figures and news agency, since those pages include more discussions and interactions between Facebook users. The tool had to be run many times to get data. Each run required that the Facebook Page ID is specified to collect the data from it. Many requests were blocked due to privacy measures on the targeted pages. The Facebook database reached 0.98GB of size, thus the decision was to keep Facebook data for validating the system in a later stage and use the tweets dB for training and testing the system at the current stage.

4.2 Data Labelling and Preprocessing

Before the system training phase, the obtained data had to be cleaned and preprocessed. WEKA was used for this purpose as mentioned above. For the system in hand, version 3.9.1 of WEKA was used since it contains specific packages that will be mentioned later.

The only features included in the first stage were text (the content of the tweet) and language. Several languages were observed in the database, mostly Arabic, English and Turkish. All the tweets of languages other than English and Arabic were discarded. The remaining tweets were separated into two datasets, one having only tweets in Arabic language and the latter English. The dataset with only Arabic language was denoted by ArabicTweets. This Dataset contained 35273 unique tweets after removing all duplicates. While the English Dataset contained 91431 tweets.

Due to the sensitivity of Arabic language and to ensure the correct labelling of sarcastic cyberbullying content, Arabic tweets were labelled manually by adding an extra attribute to the dataset, which is the "bullying". This attribute was assigned to be the class attribute, where labelling will take place and it will be used by WEKA to evaluate the performance of the system. The values of bullying instances are either "yes" or "no". "Yes" is when the tweet contains cyberbullying and "no" otherwise. The server used to preprocess the datasets was a 3.5 GHz CPU and 32GB RAM's windows 10 virtual machine.

4.3 System Training and Testing

Training and classification procedures were done several times for the purpose of reaching the best results. Two models were chosen in the first stage, Naïve Bayes and SVM. Those two models were chosen based on analysis of the previous work done in the field of ML. Researchers reached a conclusion that those are the best two algorithms for text classifications – as mention in section 2 [54, 57, 59, and 61]. In the first model the system was trained using Naïve Bayes and promising results were achieved. The precision was 90.8514 % for the overall system. However this precision is somehow tricky. The dataset contained only 2196 bullying content out of the total 35273. Thus even if in the worst case the system classifies all the tweets as not bullying then it will be still achieving a high precision. Hence the analysis of the performance was shifted towards the actual cyberbullying content which was detected (or in other words classified as "yes"). The system in the first run was capable of detecting 801 out of 2196 actual bullying instances (or tweets) which is a good result and it proves the aim of the paper: Cyberbullying in Arabic is detectable!

In training the second model; the AffectiveTweets package was used [78] and specifically the TweetToSentiStrengthFeatureVector filter. This filter depends on the SentiStrength mechanism developed by Thelwall et al [79]. SentiStrength originally works with English, but it is capable of supporting other languages. SentiStrength works on weights of tweets: positive weights (2 to 5) are given to words denoting positive feelings and negative weights for negative feelings (-2 to -5). 1 and -1 are considered neutral. Then a list of Arabic profane words was collected and weighted manually. The weight given for profane words was (-5) which is the extreme negativity in SentiStrength. Afterwards the English lexicon files used by SentiStrength were replaced with custom built Arabic files including the weighted profane words.

In the process of customizing the SentiStrength files, the phrases added contained the profanity from several Arabic countries. This gave a chance for handling dialectical difference between Arabic countries implicitly. With more effort in including more dialectical varieties of phrases all the dialects in Arabic countries would be covered.

The dataset prepared had to be preprocessed before building the training model. Preprocessing was performed using WEKA. The set of techniques used for preprocessing included applying the TweetToSentiStrengthFeatureVector filter, converting strings to word vectors and normalizing. After that the system was trained using SVM. Training and testing the model is a time-consuming process (it took around 8 hours). As a preliminary stage, the system achieved results somehow close to the results achieved by Naïve Bayes in the first model. The model was capable of detecting 710 bullying instances. Those results are promising and they show that SentiStrength can be customized to be used in cyberbullying detection. The results achieved by the system are summarized in the next section.

5. Results Obtained

Table 1 shows the division of the classified instances. How many instances were classified correctly and how many wrong in each model and class. As seen, in Naïve Bayes model 31245 non-bullying instances were classified correctly. There are also 1832 non-bullying instances classified as containing bullying. As for the bullying instances, it is seen that 801 were classified correctly while 1395 were classified as not containing bullying.

Table 1: Classification Results

	Correctly Classified		Miss Classified	
Class	***no***	***yes***	***no***	***yes***
Naïve Bayes	31245	801	1832	1395
SVM	32479	710	1923	161

Table 2: Summary OF Rates

	Class	Precision %	Recall %	TP %	FP %	F-Measure %
Naïve Bayes	***no***	94.5	95.7	95.7	69.6	95.1
	yes	36.5	30.4	30.4	4.3	33.2
	Overall	90.1	90.9	90.9	64.7	90.5
SVM	***no***	94.4	99.5	99.5	73	96.9
	yes	81.5	27	27	0.5	40.5
	Overall	93.4	94.1	94.1	67.6	92.7

Table 2 displays the results obtained by WEKA after training and testing the model. Precision, Recall, True Positive Rate, False Positive Rate and F-Measure are displayed for both models, Naïve Bayes and SVM. In each model the measures mentioned are displayed for the two classes of the "bullying" attribute which are "yes" and "no". In the third row of every model there is the weighted measure from both classes.

There is a difference between the "yes" precisions of the two models noted in Table 2. SVM achieved a much higher precision for the "yes" class -0.815 for SVM and 0.365 for NB. This difference is due to the lower number of wrong classified instances by SVM as seen in Table 2.

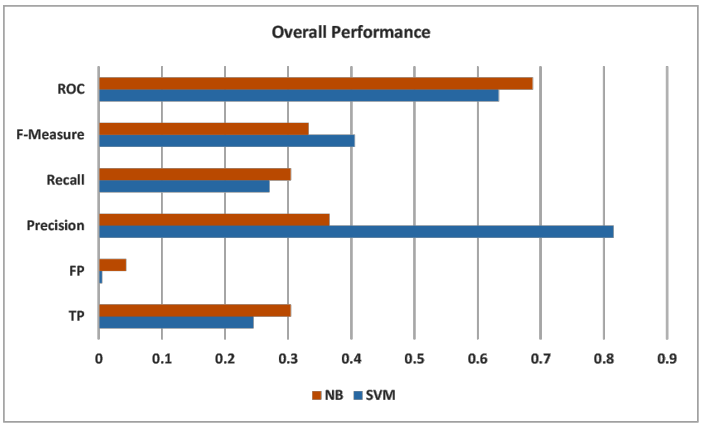


Figure 1: Difference in performance between NB and SVM

Figure 1 displays graphically the difference in performance between NB and SVM for the weighted class, or the overall per-

formance. As seen the precision achieved by SVM was much better than NB, while the other measures were somehow close, such as ROC, TP, FP, etc.

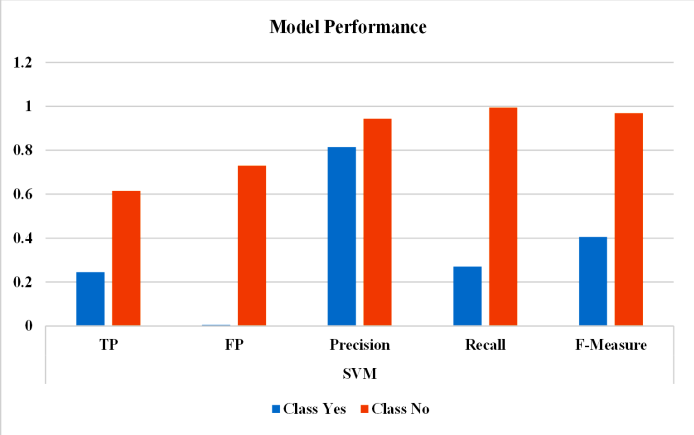


Figure 2: Comparison between the measures achieved by SVM model

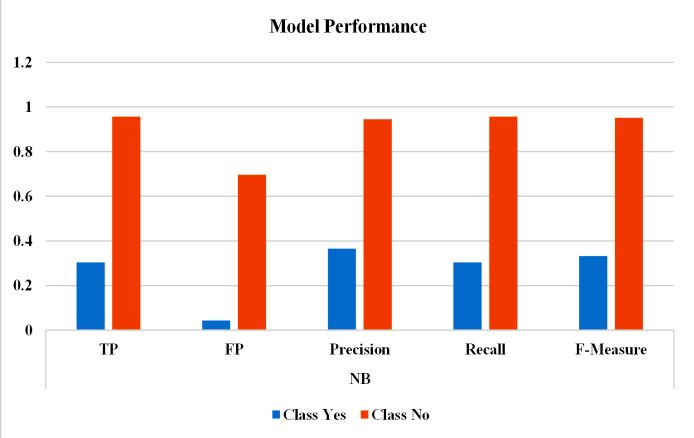


Figure 3: Results of NB Model

Figure 2 compares graphically between the measures achieved in both classes by SVM model, while Figure 3 displays the results of NB. As shown previously Table 1 and Table 2, the low number of False Positives, high precision and recall are significant.

The results obtained by this system are not perfect if compared with previous work done in English cyberbullying detection systems. But the aim of the paper was to prove that Arabic cyberbullying is detectable. Thus, the results displayed in the Tables and Figures above cannot be compared with previous work done for Arabic Language due to the negligibility of such work. Table 2 shows a recall of 30.4 for the "yes" class instances in the NB model, which means at least one third of the bullying in Arabic is detectable by the system.

The overall measures for both models give reasonable results. Although the focus was on the analysis of the "yes" class since it contains the bullying instances which are less than the "no" count. But the high precisions approached by the system denote that the model is not classifying a high number of non-bullying instances as bullying ones. Which is also an important point to consider.

6. Conclusion

The work done in this paper proves that detecting Arabic cyberbullying is possible. However, some effort is yet needed to enhance the performance of the system presented. Therefore, some future plans still need to be completed in later stages.

The first step is to enhance the performance measures achieved by the system through using hybrid training models, such as combinations of Distance Functions, NB and SVM. Treatment of Arabizi content is also in the future plans. This is a crucial point, since a lot of Arabic youth use the chat language on social networks for communication among each other. Thus the proposed system will be upgraded to handle cyberbullying when written in Arabizi.

Further plans incur training the system using deep learning methods instead of machine learning and then comparing the differences of the outcomes from the two schemes.

References

[1] B. Haidar, M. Chamoun and F. Yamout, "Cyberbullying Detection: A Survey on Multilingual Techniques," in UKSIM 2016, Pisa, Italy, 2016.

[2] K. Poels, A. DeSmet, K. Van Cleemput, S. Bastiaensens, H. Vandebosch and I. De Bourdeaudhuij, "Cyberbullying on social network sites. An experimental study into bystanders," Cyberbullying on social network sites, vol. 31, p. 259–271, 2014.

[3] S. S. Kazarian and J. Ammar, "School Bullying in the Arab World: A Review," The Arab Journal of Psychiatry , vol. 24, no. 1, pp. 37 - 45, 2013.

[4] M. Y. Ba-Isa, "Beaten up in cyberspace," 25 8 2010. [Online]. Available: http://www.arabnews.com/node/353552. [Accessed 16 9 2016].

[5] ICDL, "Cyber Safety Report: Research into the online behaviour of Arab youth and the risks they face," ICDL Arabia, 2015.

[6] K. DINAKAR, B. JONES, C. HAVASI, H. LIEBERMAN and R. PICARD, "Common Sense Reasoning for Detection, Prevention,and Mitigation of Cyberbullying," in ACM Transactions on Interactive Intelligent Systems, NY, September 2012.

[7] O. f. V. o. C. National Crime Prevention Council, "Cyberbullying Tip Sheets," National Crime Prevention Council, 2016. [Online]. Available: http://www.ncpc.org/topics/cyberbullying/cyberbullying-tip-sheets/. [Accessed 10 June 2016].

[8] N. Willard, "Educator's Guide to Cyberbullying and Cyberthreats," Center for Safe and Responsible Internet Use, 2007.

[9] N. Samaneh, A. Masrah, M. Azmi, M. S. Nurfadhilna, A. Mustapha and S. Shojaee, "13th International Conrence on Intelligent Systems Design and Applications (ISDA)," in A Review of Cyberbullying Detection . An Overview, 2013.

[10] D. Mann, "Emotional Troubles for 'Cyberbullies' and Victims," WebMD Health News, 6 July 2010. [Online]. Available: http://www.webmd.com/parenting/news/20100706/emotional-troubles-for-cyberbullies-and-victims. [Accessed 24 August 2015].

[11] T. M. Mitchell, "The Discipline of Machine Learning," CMU-ML-06-108, Pittsburgh, July 2006.

[12] P. Kulkarni, Reinforcement And Systemic Machine Learning For Decision Making, New Jersey: IEEE, WILEY, 2012.

[13] P. FLACH, MACHINE LEARNING The Art and Science of Algorithms that Make Sense of Data, Cambridge University Press, 2012.

[14] D. Vilariño, C. Esteban, D. Pinto, I. Olmos and S. León, "Information Retrieval and Classification based Approaches for the Sexual Predator Identification," Faculty of Computer Science, Mexico.

[15] H. José María Gómez and A. A. Caurcel Diaz, "Combining Predation Heuristics and Chat-Like Features in Sexual Predator Identification," 2012.

[16] A. S. a. S. Vishwanathan, Introduction to Machine Learning, Cambridge: Cambridge University Press, 2008.

[17] I. Santos, P. G. Bringas, P. Gal´an-Garc´ıa and J. Gaviria de la Puerta, "Supervised Machine Learning for the Detection of Troll Profiles in Twitter Social Network: Application to a Real Case of Cyberbullying," DeustoTech Computing, University of Deusto, 2013.

[18] I.-S. Kang, . C.-K. Kim, . S.-J. Kang and S.-H. Na, IR-based k-Nearest Neighbor Approach for Identifying Abnormal Chat Users, 2012.

[19] C. M. a. G. Hirst, Identifying Sexual Predators by SVM Classification with Lexical and Behavioral Features, 2012.

[20] D. E. L. a. A. B. Javier Parapar, "A learning-based approach for the identification of sexual predators in chat logs," 2012.

[21] Ron Kohavi and R. Quinlan, "Decision Tree Discovery," 1999.

[22] K. Reynolds, "Using Machine Learning to Detect Cyberbullying," 2012.

[23] S. Ahmad, "Tutorial on Natural Language Processing," Artificial Intelligence (810:161) Fall 2007.

[24] V. Gupta, "A Survey of Natural Language Processing Techniques," vol. 5, 01 Jan 2014.

[25] B. MANARIS, "Natural Language Processing: A Human–Computer Interaction Perspective," vol. 47, no. pp. 1-66, 1998..

[26] E. Cambria and B. White, "Jumping NLP Curves: A Review of Natural Language Processing Research," IEEE ComputatIonal IntEllIgEnCE magazIne, May 2014.

[27] C. Surabhi.M, "Natural Language Processing Future," in International Conference on Optical Imaging Sensor and Security, Coimbatore, Tamil Nadu, India, July 2-3, 2013.

[28] G. G. Chowdhury, "Natural Language Processing," Annual Review of Information Science and Technology, vol. 37, no. 0066-4200, pp. 51-89, 2003.

[29] S. Dinakar, P. Andhale and M. Rege, "Sentiment Analysis of Social Network Content," in IEEE 16th International Conference on Information Reuse and Integration, 2015.

[30] E. Cambria, Application of Common Sense Computing for the Development of a Novel Knowledge-Based Opinion Mining Engine, University of Stirling, Scotland, UK, 2011.

[31] M. Grassi, E. Cambria, A. Hussain and F. Piazza, "Sentic Web: A New Paradigm for Managing Social Media Affective Information," Cogn Comput (2011) 3:480–489.

[32] W. E. Webber, Measurement in Information Retrieval Evaluation (Doctor of Philosophy), The University of Melbourne, September 2010.

[33] C. J. v. RIJSBERGEN, INFORMATION RETRIEVAL, University of Glasgow.

[34] N. Chinchor, "MUC-4 EVALUATION METRICS," in Fourth Message Understanding Conference, 1992.

[35] Y. Sasaki, "The truth of the F-measure," University of Manchester, 26th October, 2007.

[36] T. Fawcett, "An introduction to ROC analysis," in Pattern Recognition Letters 27 (2006) 861–874.

[37] WatchGuard, "Stop Cyber-Bullying in its Tracks - Protect Schools and the Workplace," WatchGuard Technologies, 2011.

[38] "https://blog.barracuda.com/2015/02/16/3-ways-the-barracuda-web-filter-can-protect-your-classroom-from-cyberbullying/".

[39] "Internet Monitoring and Web Filtering Solutions," PEARL SOFTWARE, 2015. [Online]. Available: http://www.pearlsoftware.com/solutions/cyber-bullying-in-schools.html. [Accessed 2 June 2016].

[40] V. Nahar, X. Li and C. Pang, "An Effective Approach for Cyberbullying Detection," in Communications in Information Science and Management Engineering, May 2013.

[41] "Perverted Justice," Perverted Justice Foundation, [Online]. Available: http://www.perverted-justice.com/.

[42] "Amazon Mechanical Turk," 15 August 2014. [Online]. Available: http://docs.aws.amazon.com/AWSMechTurk/latest/AWSMechanical-TurkGettingStartedGuide/ SvcIntro.html. [Accessed 2 June 2016].

[43] S. Garner, "Weka: The waikato environment for knowledge analysis," New Zealand, 1995.

[44] "tf-idf: A single Page Tutorial," [Online]. Available: http://www.tfidf.com. [Accessed 13 May 2016].

[45] K. Dinakar , R. Reichart and H. Lieberman, "Modeling the Detection of Textual Cyberbullying," Cambridge, 2011.

[46] V. S. Chavan and Shylaja S S , "Machine Learning Approach for Detection of Cyber-Aggressive Comments by Peers on Social Media Network," in International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015.

[47] M. Dadvar, D. Trieschnigg, R. Ordelman and F. De Jong, "Improving cyberbullying detection with user context," 2013.

[48] D. Yin, Z. Xue, L. Hong, B. D. Davidson, A. Kontostathis and L. Edwards, "Detection of Harassment on Web 2.0," Madrid, Spain, April 21, 2009.

[49] J. Bayzick, A. Kontostathis and L. Edwards, "Detecting the Presence of Cyberbullying Using Computer Software," Koblenz, Germany, June 14-17, 2011.

[50] Y. Chen, S. Zhu, Y. Zhou and H. Xu, "Detecting Offensive Language in Social Media to Protect Adolescent Online Safety," 2012.

[51] Z. Xu and S. Zhu, "Filtering Offensive Language in Online Communities using Grammatical Relations," Redmond, Washington, US, July 13-14, 2010.

[52] H. Hosseinmardi, S. Arredondo Mattson, R. IbnRafiq, R. Han, Q. Lv and S. Mishra, "Detection of Cyberbullying Incidents on the Instagram Social Network," 2015.

[53] T. Nakano, S. Tatsuya, Y. Okaie and M. J. Moore, "Analysis of Cyber Aggression and Cyber-bullying in Social Networking," in IEEE Tenth International Conference on Semantic Computing, 2016.

[54] N. Potha and M. Maragoudakis, "Cyberbullying Detection using Time Series Modeling," 2014.

[55] K. Baker, "Singular Value Decomposition Tutorial," 2013.

[56] M. Muller, "Dynamic Time Warping," in Information Retrieval for Music and Motion, Springer, 2007, pp. 69 - 84.

[57] B. Nandhinia and J. Sheebab , "Online Social Network Bullying Detection Using Intelligence Techniques," 2015.

[58] M. A. Attia, Handling Arabic Morphological and Syntactic Ambiguity within the LFG Framework with a View to Machine Translation, Doctor of Philosophy in the Faculty of Humanities, 2008.

[59] K. Darwish and W. Magdy, "Arabic Information Retrieval," vol. 7, no. 4, 2013.

[60] A. FARGHALY and K. Shaalan, "Arabic Natural Language Processing:Challenges and Solutions," vol. 8, December 2009.

[61] "12 Arabic Swear Words and Their Meanings You Didn't Know," [Online]. Available: http://scoopempire.com/swear-words-meanings-around-middle-east/#.V0fdjPl96M9. [Accessed 2 June 2016].

[62] N. Ghaleb Ali and N. Omar, "Arabic Keyphrases Extraction Using a Hybrid of Statistical and Machine Learning," in International Conference on Information Technology and Multimedia (ICIMU), Putrajaya, Malaysia, 2014.

[63] T. Haifley, "Linear Logistic Regression: An Introduction," IEEE, 2002.

[64] G. J. McLACHLAN, "Discriminant Analysis and Statistical Pattern Recognition," Wiley InterScience, New Jersey, 2004.

[65] K. Shaalan and H. Raza, "Arabic Named Entity Recognition from Diverse Text Types," Berlin Heidelberg, GoTAL 2008.

[66] A. El-Halees, "Filtering Spam E-Mail from Mixed Arabic and English Messages: A Comparison of Machine Learning Techniques," The International Arab Journal of Information Technology, vol. 6, no. 1, 2009.

[67] T. M. COVER and P. E. HART, "Nearest Neighbor Pattern Classification," IEEE TRANSACTIONS ON INFORMATION THEORY, vol. 13, no. 1, 1967.

[68] N. EI-Mawass and S. Alaboodi, "Detecting Arabic Spammers and Content Polluters on Twitter," in Sixth International Conference on Digital Information Processing and Communications (ICDIPC), 2016.

[69] R. M. Duwairi, M. Alfaqeh, M. Wardat and A. Alrabadi, "Sentiment Analysis for Arabizi Text," in 7th International Conference on Information and Communication Systems (ICICS), 2016.

[70] A. E.-D. A. Hamouda and F. E.-z. El-taher, "Sentiment Analyzer for Arabic Comments System," (IJACSA) International Journal of Advanced Computer Science and Applications, vol. 4, no. 3, 2013.

[71] R. M. Duwairi, R. Marji, N. Sha'ban and S. Rushaidat, "Sentiment Analysis in Arabic Tweets," in 5th International Conference on Information and Communication Systems (ICICS), 2014.

[72] R. M. Duwairi, "Sentiment Analysis for Dialectical Arabic," in 6th International Conference on Information and Communication Systems (ICICS), 2015.

[73] S. Alhumoud, T. Albuhairi and M. Altuwaijri, "Arabic sentiment analysis using WEKA a hybrid learning approach," in 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K), Lisbon, Purtogal, 2015.

[74] A. Al-Zyoud and W. A. Al-Rabayah, "Arabic Stemming Techniques: Comparisons and New Vision," in Proceedings of the 8th IEEE GCC Conference and Exhibition, Muscat, Oman, 2015.

[75] S. Khoja and R. Garside, "Stemming arabic text," Computing Department, Lancaster University, Lancaster, UK, 1999.

[76] L. S. Larkey, L. Ballesteros and M. E. Connell, "Light Stemming for Arabic Information Retrieval," in Arabic Computational Morphology, book chapter, , , Springer, 2007.

[77] S. Gadri and A. Moussaoui, "Information Retrieval: A New Multilingual Stemmer Based on a Statistical Approach," in 3rd International Conference on Control, Engineering & Information Technology (CEIT), 2015.

[78] S. M. Mohammad and F. Bravo-Marquez, "Emotion Intensities in Tweets," in Joint Conference on Lexical and Computational Semantics (*Sem), Vancouver, Canada, August 2017.

[79] M. B. K. P. G. C. D. &. K. A. Thelwall, "Sentiment strength detection in short informal text," Journal of the American Society for Information Science and Technology, vol. 61, no. 12, p. 2544–2558, 2010.

[80] Hewlett-Packard Development Company. L.P., 2013. [Online]. Available: http://www.autonomy.com/html/power/idol-10.5/index.html. [Accessed 2 June 2016].

[81] "Arabic chat alphabet," 23 May 2016. [Online]. Available: https://en.wikipedia.org/wiki/Arabic_chat_alphabet. [Accessed 2 June 2016].

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 285-289 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Applying Machine Learning and High Performance Computing to Water Quality Assessment and Prediction

Ruijian Zhang*[1], Deren Li[2]

[1]*Purdue University Northwest, Computer Science, 46323, USA*

[2]*Indiana University-Purdue University Indianapolis, Computer Science, 46202, USA*

ARTICLE INFO

Article history:
Received: 13 November, 2017
Accepted: 03 December, 2017
Online: 23 December, 2017

Keywords:
Machine Learning
Classification
Decision Tree
High Performance Computing
Water Quality Assessment and Prediction

ABSTRACT

Water quality assessment and prediction is a more and more important issue. Traditional ways either take lots of time or they can only do assessments. In this research, by applying machine learning algorithm to a long period time of water attributes' data; we can generate a decision tree so that it can predict the future day's water quality in an easy and efficient way. The idea is to combine the traditional ways and the computer algorithms together. Using machine learning algorithms, the assessment of water quality will be far more efficient, and by generating the decision tree, the prediction will be quite accurate. The drawback of the machine learning modeling is that the execution takes quite long time, especially when we employ a better accuracy but more time-consuming algorithm in clustering. Therefore, we applied the high performance computing (HPC) System to deal with this problem. Up to now, the pilot experiments have achieved very promising preliminary results. The visualized water quality assessment and prediction obtained from this project would be published in an interactive website so that the public and the environmental managers could use the information for their decision making.

1. Introduction

This paper is an extension of work of "Applying Parallel Programming and High Performance Computing to Speed up Data Mining Processing", originally presented in the 16th IEEE international Conference on Computer and Information Science, 2017 [1]. Water quality is an important issue. Not only can the drinking water directly influence our body as about 70% of the human body is made up of water, the natural water is also an important part as people like to have some entertainment activities near lakes or rivers. Nowadays the prediction of river or lake water qualities is more and more important. The traditional way to assess water quality is analyzing the water attributes' value. If the value is in a certain range, the quality of water can be classified. This is an easy and fast way, but it can only be used for assessment. To predict water quality, scientists need to use the chart to show the changes of an attribute in a period of time. This is a statistic way and can predict water quality in a relatively accurate way. The problem is it takes lots of time to analyze the water attributes and it can only predict a long period of time's water quality. As the results are analyzed by people, increasing water attributes for analyzing will greatly increase the workload [2, 3, 4, 5, 6, 7].

In this research, we use machine learning algorithm on water quality prediction. By applying machine learning algorithms, the increasing water attributes and cases are no longer a big problem. Compared with the traditional way, using machine learning can deal with a great number of data in a short time. It can also slightly reduce the human error on the analysis phase. The idea of implementation is not very complex. First, we use the K-MEANS algorithm to classify the data. The K-MEANS algorithm is a very basic, easy and important clustering algorithm. In this research, we specify K is 5, which means we want to classify the water data into five clusters to represent five water qualities from good, relatively good, medium, relatively bad, and bad. By analyzing each cluster's center; we will know which cluster represents what water quality. After that, we use the C5.0 algorithm to generate a decision tree [8]. In this research, we use five or seven water attributes for classification.

*Ruijian Zhang, 2200 169th Street, Hammond IN USA 46323, Phone: (219)9892706, zhangrj@pnw.edu

2. Significance

As water quality is more and more important, not only scientists but also the general publics are interested in the future natural water quality. Thus, the request to predict natural water quality like weather forecast is getting stronger. Unlike weather forecast which can use satellites to get, we can't use several satellites to cover a large area to collect water data. The data needs to be collected by observations in different places. It also takes far more time to analyze the data so that people can make a water quality prediction. The United States is a country with many rivers and lakes. The government has already established many water observations to monitor the water system of the country [9]. Collecting data is not so difficult. Making predictions in an efficient way is the key problem.

For this research, the most significant part is it completely uses the data from currently operational observations. This means the hardware cost of prediction is greatly reduced. As long as an observation station can provide enough data, this method can be used to predict water quality. By using machine learning technology, we can deal with a great amount of data in a short time, which can make water quality analyze very quickly. Although these kinds of water quality analysis may not be as accurate as biology analysis, it is accurate enough for water quality forecast to the generate public [10]. Compared with the traditional way for water quality research, the human factors in this research is relatively low. This can slightly reduce the possibility of human error.

3. Design and Implementation

There are four steps in this research: data selection, data preprocessing, classification, and assessment and prediction. The first two steps are preparing a suitable data for classifying. Step three is to make the classifications. It will be implemented by the K-MEANS algorithm. Step four will generate different decision trees for assessment and predicting by using the C5.0 algorithm.

Data selection is very important. This step is the root of this research. As this is a machine learning research, a "good" data definitely can greatly affect the results. The "good" here means it needs to have enough cases, enough attributes, and less missing data. Enough cases can ensure the number of training data. The accuracy of classifications is based on the number of training data. The more cases we have, the better the classifications will be.

The data we used was from the USGS website. USGS keeps the daily water data from 10472 observation sites all over the United States. Many of the sites keep about four attributes of the water. After checking all the sites keeping at least seven attributes, we finally choose the site numbered 08070200. This site is in San Jacinto River in Texas and the data is still updating till now. It keeps 7 attributes of the water, and has few missing values. In this research, we choose the data between 11/22/2005 and 09/08/2012 which has 2483 records [9].

The second step is data preprocessing. We made the data more meaningful for use. It included resetting the extreme data, and normalizing the data. To deal with the extreme data, we put the entire data of one attribute on a scatter chart and find the extreme value. Fig. 1 is an example of original "discharge" value shown in a chart. The x-axis represents different cases and the y-axis is value. It is very clear that the maximum value of discharge is about 8000. If we use these values directly and normalize them, the problem is most normalized value will be 0. This is definitely not good for classification as it cannot reflect the data distribution. To solve this problem, the idea is set all values greater than an "extreme value" to be equal to this "extreme value". In Fig. 2, we use the same data in Fig. 1 but set all values greater than 700 to be equal to 700.

Fig. 2 looks better than Fig. 1. It now can reflect the changes and distribution of the discharge. A problem here is how to set the "extreme value". In Fig. 2, we can see that there is not too much difference if we set the extreme value to 600 compared with 700, as most value are under 200. The key point is we do not want to change too much value as changing data is a kind of human factor. Making the charts have better data distribution with the least changes is the role of the extreme value. In this attribute, totally 136 values have been affected by setting the extreme value to 700, which is about 5% of the total 2483 values.

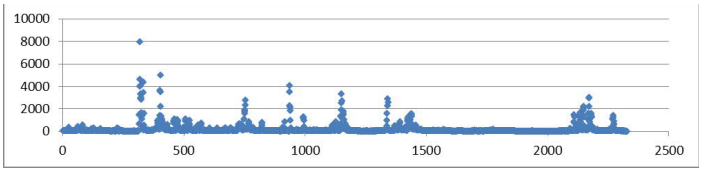


Figure 1, the original scatter chart of the discharge in this research. All values are either the original values or calculated values.

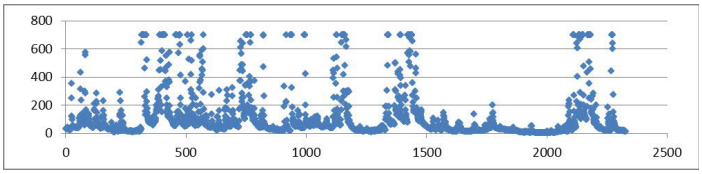


Figure 2, the new discharge chart which set all values greater than 700 to be equal to 700.

The third step of this research is classification. We want to classify the data into five clusters (classifications). Each cluster represents a certain water quality range from very bad, bad, medium, good, to very good. The method for classification is to implement the K-MEANS algorithm. The K-MEANS algorithm is a very basic but important algorithm for classification. It treats every node in an N-dimensional space. Each dimension represents an attribute. In this research, we used seven attributes, so each record was a node in 7-dimension spaces. We first gave five initial centers. Then based on the distance of each node and each center, we assigned every node a group. After that, we recalculated the distance of every node inside one group to move the center. So we repeated the iteration of assigning groups and moving centers until the distance between the new center and the previous center was smaller than a "quit distance" or the new center moved to a previous center node.

The last step of this research was assessment and prediction. Before we could make the assessment and prediction, the very first thing was to analyze which group represents what kind of water quality. This can be done by analyzing the center node's attributes. We needed to look up some references and some criteria of water. This was not easy but also not hard. The results of classification actually were the results for water quality assessment. We already classified the data into five groups. We could make a water quality assessment by generate a decision tree.

The main topic was how to make prediction. Before we could make the decision tree, we needed the training data. The format of the training data we used to generate the decision tree was "attributes" plus "classification results". If the training data was "day 1's attributes" plus "day 1's classification result", it was used to generate a decision tree for assessment. So, we "shifted" the classification results by 1day. The training data would be like "day 1's attributes" plus "day 2's classification result". This data can generate a decision tree for prediction. Fig. 3 shows how to "shift" the data. In this way, we lose one day's case as the last day only has attributes. This is a basic way for prediction. It uses previous 1day's attributes for prediction. An improved way is using more than 1 day's attributes. We first created an additional attribute which made the first line look like "day 1's attributes" plus "day 1's attributes" plus "day 2's classification result". Then we shifted it to "day 1's attributes" plus "day 2's attributes" plus "day 3's classification result". This was using previous 2 days' attributes to predict the third day's water quality. In our research, we used up to previous 8 day's attributes to predict the ninth day's water quality to see if the results have any improvements. This does not mean the more the better. In fact, adding more day's attributes is the same with adding water attributes. It can increase the data interference, which will increase the difficulty to generate the decision tree and reduce the accuracy of prediction.

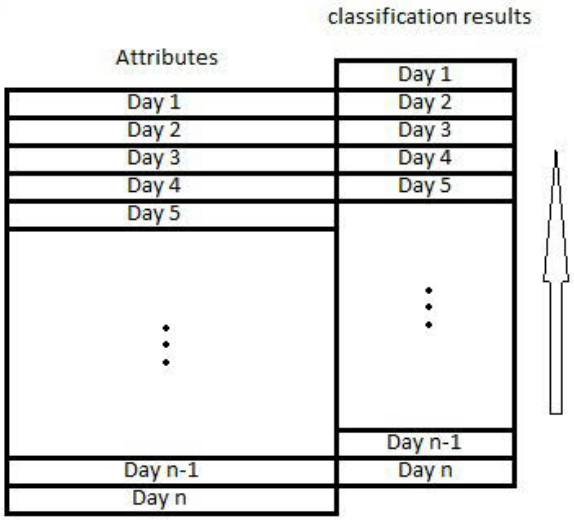


Figure 3, the idea of how to "shift" the classification result.

After preparing the training data, we used See5.0 to generate a decision tree. See5.0 is a software tool which implements the C5.0 algorithm. It can deal with a great amount of data in a short time and generate the decision trees by the classic method, the winnow method, and the boost method. It generates a decision tree and gives the importance of every attribute. Most times all attributes are used. If one attribute has not been used by the C5.0 algorithm, very often this attribute may have a linear relationship with some other attributes [9].

4. Result and Analysis

The results analysis of this research is based on the prediction accuracy. Among the 2483 records, the even records are used as training data to generate decision tree, the odd records are used as testing data. We tested the decision tree by testing data to get the prediction error rate.

The accuracy for assessment was quite good. Using the winnow method, the prediction accuracy was 93.5% by using seven attributes and 93.6% by using five attributes. Dissolved oxygen, turbidity and PH were mainly used attributes for assessment. The usage of temperature was relatively low, which shows the way we normalized the temperature reasonable. Using boost method, the prediction accuracy was 96.1% by using 7 attributes and 95% by using five attributes. The results of boost method show increasing attributes can increase the prediction accuracy.

The prediction has 32 results into 2 groups. One group used five attributes to predict, another group used seven attributes to predict. In this way, we can know if increasing attributes can improve the prediction accuracy. Both the winnow and the boost methods are used for test. One group have 8 tests differed by using different number of day's attributes.

Table 1 is the result for group 1. It shows the error rate for prediction the next day's water quality by using five attributes to classify the data into five classifications. Table 2 is the result for group 2. It shows the error rate for prediction the next day's water quality by using seven attributes to classify the data into five classifications. All the results above are the test error rate assigned by See5.0. Based on these results, we can change the data from error rate to prediction accuracy as an alternative.

Table 1, the error rate of predict the next day's water quality by using 5 attributes and 5 classifications. W stand for winnow method and B stand for Boost method.

Predict 2nd Day		Predict 3rd Day		Predict 4th Day		Predict 5th Day	
W	B	W	B	W	B	W	B
19.1%	16.5%	19.4%	17.1%	19.6%	16.9%	20.7%	17.0%
Predict 6th Day		Predict 7th Day		Predict 8th Day		Predict 9th Day	
W	B	W	B	W	B	W	B
22.8%	16.1%	21.2%	17.0%	21.4%	17.0%	22.6%	16.9%

Table 2, the error rate of predict the next day's water quality by using 7 attributes and 5 classifications. W stand for winnow method and B stand for Boost method.

Predict 2nd Day		Predict 3rd Day		Predict 4th Day		Predict 5th Day	
W	B	W	B	W	B	W	B
19.6%	18.1%	21.2%	16.8%	20.2%	16.9%	20.3%	16.9%
Predict 6th Day		Predict 7th Day		Predict 8th Day		Predict 9th Day	
W	B	W	B	W	B	W	B
21.5%	16.6%	21.1%	16.0%	22.1%	17.7%	20.7%	16.3%

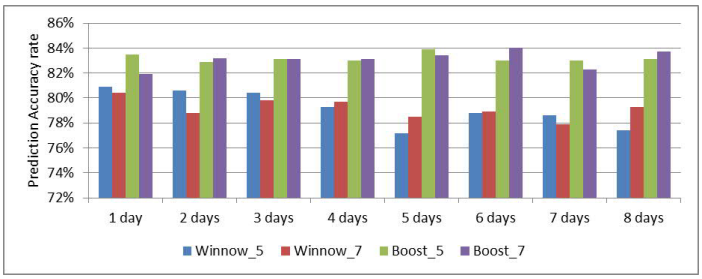


Figure 4, Prediction accuracy of two different methods and using different number of attributes on predict the next day's water quality to 5 classifications.

Figure. 4 shows the prediction accuracy for each method using different attributes. It proves the great performance of boost method. We can see using seven attributes has better performance. The winnow method seems to have no great difference. The boost method, however, shows the advantage of using seven attributes. Out of eight tests, five tests show using seven attributes can reach a better performance than using five attributes. Also using the boost method reaches the highest prediction accuracy, which is 84% by using previous six day's attributes to predict the next day's water quality.

This figure also shows some interesting results. As we mentioned before, increasing more day's attributes may not improve the prediction accuracy. Actually, these eight tests are all used to predict the next day's water quality. With the increase number of attributes, the prediction accuracy of the winnow method, generally speaking, is reducing. This is because data interference. No matter we increase the number of single day's attributes or we increase the number of days we use for prediction; they have same effect for the winnow method. When the attributes' number is greatly increased, the difficulty to generate the decision tree also increasing tremendously. The boost method, however, doesn't take too much effect from data interference. As each decision tree make a slightly improvement based on the previous decision tree's error, the "vote" mechanism can greatly reduce the data interference and has a better performance.

5. High Performance Computing Speed up the Processing

In this project, we were running machine learning program on Purdue University's High Performance Cluster – *Falcon*, which has 8 nodes, each with 4 CPUs, and/or *Miner*, which has 512 nodes, each with 4 CPUs. This high performance computing (HPC) system is supported by the Northwest Indiana Computational Grid (NWICG) program funded by the U.S. Department of Energy grant ($4.9 million). Taking advantage of the NWICG's HPC system, it is practical for us to use complexity to exchange the water quality prediction accuracy. HPC dramatically decreases the running time of the enumeration based algorithm, which is time consuming but more accurate.

Facilitated perfectly by the current existing Northwest Indiana Computational Grid, the project has applied the HPC system which dramatically decreases the running time of the enumeration based algorithm, so that makes this very time-consuming algorithm for water quality prediction becoming practical.

Up to now, the experiments of this innovative approach obtained very promising results. The sixteen-processor HPC system achieved more accurate water quality prediction and reduced the execution time by more than ten times. The experiment results showed applying machine learning models on HPC system is a very encourage direction for Northwest Indiana's water quality assessment and prediction.

In light of almost 3000 cases, the optimal enumerative algorithm is very time-consuming (running in days) when executing on a single machine. In order to improve the performance, we applied parallel programming on the enumerative algorithm. The program run on multi-processors in parallelism. The parallel computing assigned one process as the master process and others as slave process. The master process read in the data, broadcasted the data and other information to the slave processes. The parallel programming split the clustering calculation into threads for slaves. Each slave process worked on calculations for clustering. Finally, the master processor received the results from each slave process and chose the global optimized clustering [11, 12, 13].

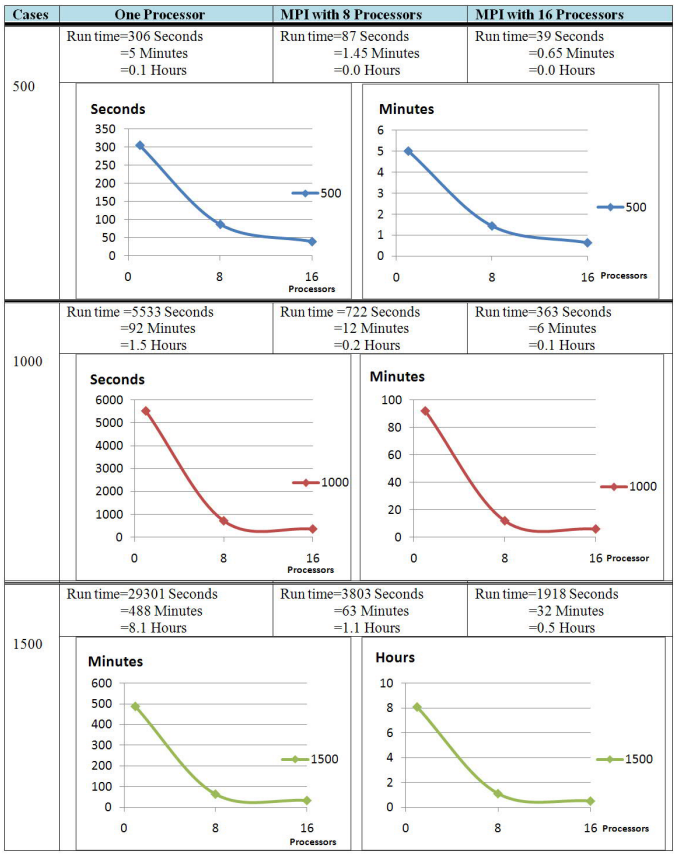


Figure 5, Experiment results with MPI on the HPC

Our experiments compared the running time of the enumerative algorithm using a single processor with that using HPC with 8 processors and 16 processors. The number of the cases was 500, 1000, 1500, 2000, 2500, and 2700 respectively. Figure 5 and Figure 6 below show the execution time of our preliminary experiments. It exposed the critical role of HPC for the time consuming enumerative algorithm to process large amount of data. For example, for 1500 cases, running on a single processor, it took more than eight hours to finish the clustering; whereas running on HPC with 16 processors, it only took about 32 minutes. This improvement of execution time is critical, because it is important to have a reasonable amount of indicative and representative data samples (cases) in order to discover patterns in data samples using machine learning technologies. It is not unusual to have a data set of 1500 cases by this approach.

The results of these experiments produced a solid fundamental for future developments in our water quality assessment and prediction project. We created decision tree employed C5.0 for the water quality assessment and prediction. Using the produced decision trees to predict unseen cases, the prediction accuracies reach 82 percent, two percent better than applying k-means algorithm and about the same improvement than applying the

mechanistic simulation models. It is anticipated that if we use more attributes of the water quality in the future, the accuracy rate could be further improved.

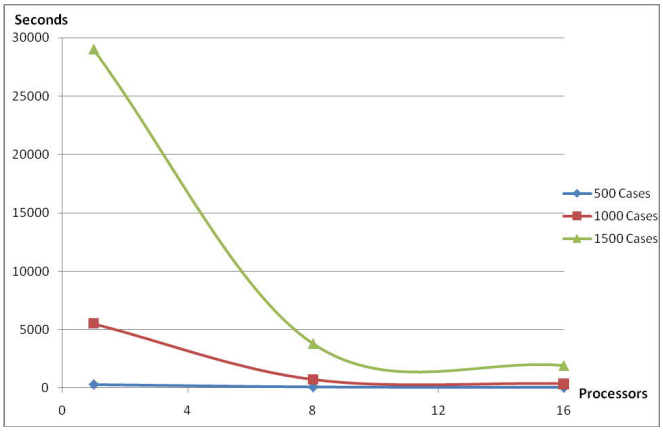


Figure 6, Execution time in seconds for comparison of various processors in HPC

6. Summary and Conclusion

In this research, by applying machine learning algorithm to a long period time of water attributes' data; we can generate a decision tree so that it can predict the future day's water quality in an easy and efficient way. The idea is to combine the traditional ways and the computer algorithms together. Using machine learning algorithms, the assessment of water quality will be far more efficient, and by generating the decision tree, the prediction will be quite accurate. The drawback of the machine learning modeling is that the execution takes quite long time, especially when we employ a better accuracy but more time-consuming algorithm in clustering. Therefore, we applied the HPC System to deal with this problem. Up to now, the pilot experiments have achieved very promising preliminary results. The visualized water quality assessment and prediction obtained from this project would be published in an interactive website so that the public and the environmental managers could use the information for their decision making.

References

1. Ruijian Zhang, "Applying Parallel Programming and High Performance Computing to Speed up Data Mining Processing", Proceedings of the 16th IEEE international Conference on Computer and Information Science, May 2017.
2. Faruk, D. "A hybrid neural network and ARIMA model for water quality time series prediction" Eng. Appl. Aritf. Intel., 23, 586-594, 2009. https://doi.org/10.1016/j.engappai.2009.09.015
3. L. M. He, Z. L. He, "Water quality prediction of marine recreational beaches receiving watershed baseflow and stormwater runoff in southern California", Water. Res., 42, 2563-2573, 2008. https://doi.org/10.1016/j.watres.2008.01.002
4. W. Labiosa, J. Leckie, R. Shachter, D. Freyberg, J. Rytuba, "Incorporating Uncertainty in Watershed Management Decision-Making: A Mercury TMDL Case Study" in Watershed Management Conference, Williamsburg, Virginia, United States, 2005. https://doi.org/10.1061/40763(178)125
5. A. Said, "The Implementation of a Bayesian Network for Watershed Management Decisions" Int. Ser. Prog. Wat. Res., 20, 591–605. 2005. https://doi.org/10.1007/s11269-006-3088-z
6. S. Shrestha, F. Kazama, "Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan" Environ. Modell. Softw., 22, 464-475, 2006. https://doi.org/10.1016/j.envsoft.2006.02.001
7. C. A. Stow, M. E. Borsuk, K. H. Reckhow, "Nitrogen TMDL development in the Neuse River watershed: An imperative for adaptive management", Water Resources Update, 122, 16–26, 2002.
8. See 5.0 [Computer software]. Australia: Rulequest Research, 2012.
9. USGS Water-Quality Data for the Nation. USGS Water-Quality Data for the Nation from U.S. Geological Survey, U.S Department of the Interior, 2012. http://waterdata.usgs.gov/nwis/qw
10. S. Tong, W. Chen, "Modeling the relationship between land use and surface water quality". Environ. Manage., 66, 377-393, 2002. https://doi.org/10.1006/jema.2002.0593
11. C. Severance, K. Dowd, High performance computing, 2nd Edition, 1998.
12. G. Stellner, "Cocheck: Checkpointing and Process Migration for MPI", Proceedings of International Conference on Parallel Processing, Hawaii, USA, 1996. https://doi.org/10.1109/IPPS.1996.508106
13. P.L. Vaughan, A. Skjellum, D.S. Reese, F.C. Cheng, "Migrating from PVM to MPI, Part I: The Unify System", Frontiers of Massively Parallel Computation, McLean, Virginia, United States, 1995. https://doi.org/10.1109/FMPC.1995.380476

Advances in Science, Technology and Engineering Systems Journal Vol. 2, No. 6, 290-301 (2017)

www.astesj.com

Special issue on Advancement in Engineering Technology

ASTESJ

ISSN: 2415-6698

Flexible lengthening-shortening arm mechanism for fishery resource management

Yoshiki Iwamochi[1], Motoki Takagi[2], Tasuku Miyoshi[1,3*]

[1]*Department of Mechanical Engineering, Iwate University, Morioka-city, Iwate 020-8551, Japan*

[2]*Department of Information and Electronic Engineering, Teikyo University, Utsunomiya-City, Tochigi 320-8551, Japan*

[3]*Soft-Path Science and Engineering Research Center (SPERC), Faculty of Science and Engineering, Iwate University, Morioka-city, Iwate 020-8551, Japan*

ARTICLE INFO

Article history:
Received: 31 October, 2017
Accepted: 16 December, 2017
Online: 23 December, 2017

Keywords :
Underwater robot arm
Flexible arm
Sliding screw mechanism
Obstacle avoidance
Path-tracing algorithm

ABSTRACT

The goal of this study was to use underwater robots instead of a diver's observations to monitor and record the condition of an obstructed seabed in a shallow area. It is difficult to investigate marine resources that exist in deep water shaded by rock due to large and/or small rocks on the seabed. To solve these problems, we newly constructed a flexible lengthening-shortening arm with a small camera unit for an underwater robot to assist in the management of fishery resources. In this paper, we describe the concept and configuration of the newly developed arm mechanism using a sliding screw mechanism to overcome obstacles by changing arm posture in a two-dimensional plane, and we demonstrate the experimental results of a path-tracing controller for the rear links. The results were that the maximum deviations between the target path and the tracing path were less than 4.0% of the total width of the arm mechanism. These results suggest that the newly developed path-tracing algorithm is effective for our flexible lengthening-shortening arm mechanism.

1. Introduction

This report is an extension of work that was originally presented at the 2017 IEEE International Conference on Mechatronics and Automation (ICMA) [1]. Extended points are as follows; (a) the geometric configurations and control parameters for an unique slider-screw mechanism of our system in Section 2.2.3, (b) waterproof design for the slider-screw mechanism in Section 2.2.4, (c) an operational test on land to clarify the validity of the flexible lengthening-shortening arm mechanism in Section 3.1, (d) an explanation of the details of the path-tracing algorithm in Section 4.2.3 and Appendix, and (e) improvements of the movement error occurred in previous work [1].

Underwater robots, such as autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs), were developed to reduce the big burden on fishermen in managing fishery resources and/or the cost of submarine operations and risk to human life (e.g.; AUV for work in shallow waters and coastal areas [2], reduce the burden on divers [3], achieving the core sampling at shallow seabed [4], and achieving the video transect method instead of divers [5]). Generally, underwater robots are equipped with cameras to monitor and record the condition of the seabed rather than relying on a diver's observations. In the case of shallow sea areas, such as surf zones, there are many rocks and marine algae. Although these areas are useful as growth environments of fishery resources, it is assumed that the view of the camera is hindered and/or the robots cannot approach the target objects closely. In addition, it is difficult to investigate rock shadows, which cannot be penetrated by underwater robots, and the shellfish living deep within them.

There is a possibility that a survey could not be achieved using current robot arms (e.g.; dual arm operation [6] and tele-operated manipulators installed to ROV [7]). Instead of using robot arms, developing snake-type robots is one solution. Snake robots can search for and investigate targets, avoiding multiple obstacles by their serpentine motion in narrow environments where humans cannot intrude. Previously, many researchers had been developing

*Corresponding Author: Tasuku Miyoshi, Department of Mechanical Engineering, Iwate University, Morioka-city, Iwate 020-8551, Japan, Email: tmiyoshi@iwate-u.ac.jp

snake-type robots for investigating unknown environments and searching for disaster sites (e.g.; review of snake-type robots [8], tutorial of snake-type robot [9], and carrying out the inspection work [10]). Also, an amphibious snake-type robot, which moves in a serpentine motion for generating the forward propulsion in water, has been studied [11]. With this unique motion, however, it is necessary for the robot to keep thrusting in order to stay in place in shallow areas where wave disturbances occur from all directions. Without any propulsion force for forward movement by a serpentine motion, a mobile robot equipped snake-type arm mechanisms had been studied (e.g.; surveying robot equipped with snake-type arm on wheel robot [12], assembling an aircraft using a snake-type arm mounted on a traditional industrial robot arm [13], and linear motion mechanism mounted on base of snake-type arm [14]). When entering the camera at the tip of the arm into the narrow space, the snake-type arm changes its attitude so as to avoid obstacles and the intrusion operation (forward approach action) is in charge of the base unit. In order to investigate a narrow area in the shallow sea using AUV or ROV equipped an arm with snake-type mechanism, it is necessary to take forward motion while maintaining posture and position against the tidal current near the seabed. Therefore, we assumed that a snake-type arm, which had a forward movement mechanism and installed to AUV and/or ROV, is effective for investigating narrow areas in shallow sea. Hence, the purposes of this study were: 1) to construct a new robot arm that utilizes the features of a snake-type robot's obstacle avoidance, and 2) to evaluate the autonomous path-tracing algorithm to achieve obstacle avoidance at the part of the rear-link mechanism. In this paper, Section 2 outlines our newly developed arm mechanism and the results of the operation test. In Section 3, we propose a path-tracing algorithm of the rear-link mechanism required for obstacle avoidance and show the effectiveness of this algorithm experimentally.

2. Flexible lengthening-shortening arm mechanism (FLSAM)

2.1. Design specifications

We selected the following design specifications for the newly developed flexible lengthening-shortening arm mechanism (FLSAM):

1. The FLSAM can avoid obstacles,

2. The FLSAM can reach into rock shadows and penetralia,

3. The FLSAM can approach target objects closely, and

4. The FLSAM can obtain photos and movies for further analysis.

To satisfy the required specifications of the rotational and translational degrees of freedom, the FLSAM's design considers the following points:

A1. To avoid obstacles, the FLSAM should be equipped with rotational degrees of freedom,

A2. To reach into rock shadows and go near the target, the FLSAM should be able to expand its arm length by translational degrees of freedom, and

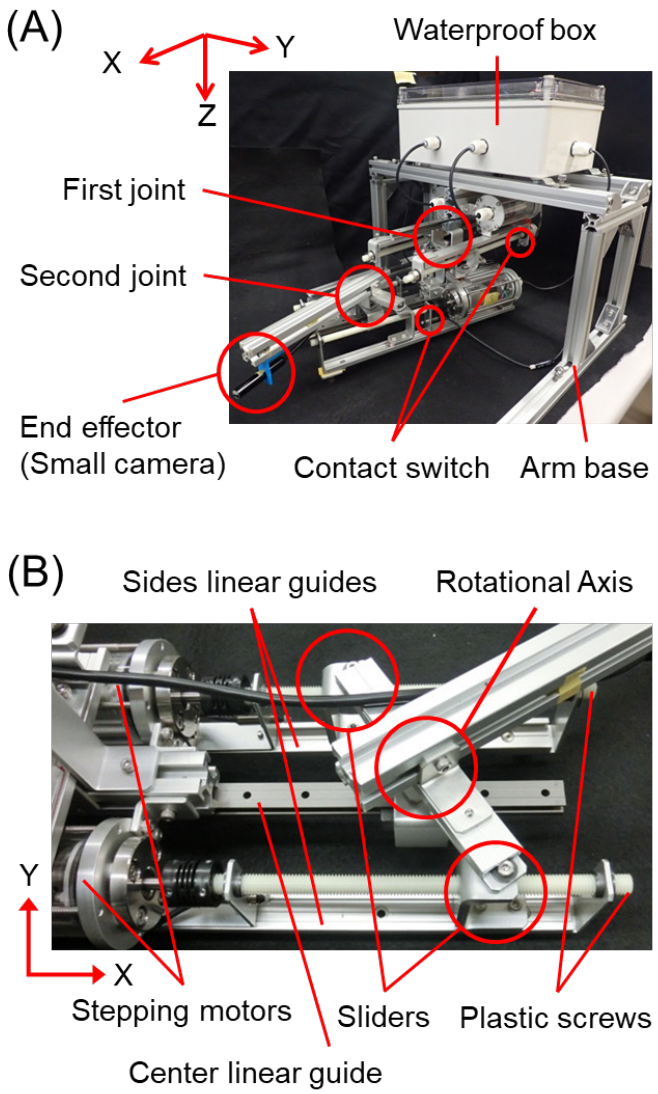


Figure 1: (A) Overview of the newly developed flexible lengthening-shortening arm mechanism, "FLSAM"; (B) mechanical details around the second joint of the FLSAM

A3. To obtain photos and movies, a small video camera should be installed on the FLSAM.

2.2. Configuration of the FLSAM

2.2.1. Overview of the FLSAM

Figure 1A shows an overview of the FLSAM, and the specifications are denoted in Table 1. Each joint consists of two stepping motors (PKP213D05A, ORIENTAL MOTOR Co., Ltd.), two plastic screws (MB-RENY-M8-300, MISUMI Corporation), and three linear guides (for both sides: NK-02-17-1-220-LLYZ, igus k.k.; for the center: TK-04-12-1-250, MISUMI Corporation). Each linear guide was in parallel, and sliders were installed on both sides of the linear guides. To achieve rotational degrees of freedom, the rotational axis was set at the center linear guide. As to the detailed structure of the joint, the appearance of the second joint is shown in Figure 1B. Contact switches (PTP5M3CB-L, MISUMI Corporation) were installed to detect the origin of the

Table 1: Specifications of the FLSAM

Total height [mm]	190
Shortened length [mm]	290
Maximum lengthened condition [mm]	560
Total width [mm]	140
First joint stroke [mm]	90
Second joint stroke [mm]	180
Range of motion [deg]	Plus/minus 45
Total weight [kg]	3.9
Sensors	Small camera Contact switch

sliders and to prevent an overrun. A small waterproof camera (WPMISC7M, THANKO Co., Ltd.) was equipped to record photos and movies from the end effector. Electric circuits with an AVR microcomputer (Arduino Uno R3, Arduino AG) were placed in a waterproof box (SPCM13-18-10-T, MISUMI Corporation).

2.2.2. *Sliding screw mechanism to achieve rotational and translational degrees of freedom of the FLSAM*

When constructing an arm with the simultaneous motion of the translational and rotational degrees of freedom, it is natural to couple the translational and rotational mechanisms. However, the arm mechanism increases in size due to the increased number of joints when the translational and rotational joints are arranged separately. Therefore, we adopted a sliding screw mechanism to achieve the translational and rotational degrees of freedom in one joint. The sliding screw mechanism could be expected to reduce the size and number of joints. Moreover, it could be said of this mechanism that (1) power consumption is low because it is not necessary to maintain the posture, and (2) it can operate in water without a lubricant [15].

Figure 2A–C shows the schematic of the sliding screw mechanism installed in the FLSAM. We arranged two sliding screws for one joint, and the slider part moves forward and backward with respect to the rotational volume of the first and second actuators. The FLSAM installed two sliding screw mechanisms; in total, four degrees of freedom were achieved. The changes in the stroke length of each sliding screw mechanism were defined by the number P_k [pulse] and frequency A_k [pulse/s] of pulses' input to the stepping motor, as

$$\begin{bmatrix} P_k \\ A_k \end{bmatrix} = \frac{R}{L_e} \begin{bmatrix} X_k \\ V_k \end{bmatrix} \quad (1),$$

where, k is the number of the slider (1 and 2 for the first joint, 3 and 4 for the second joint), X_k [mm] is the stroke length of each slider, V_k [mm/s] is the moving velocity of each slider, L_e [mm] is the lead of the screw, and R [pulse] is the number of pulses per

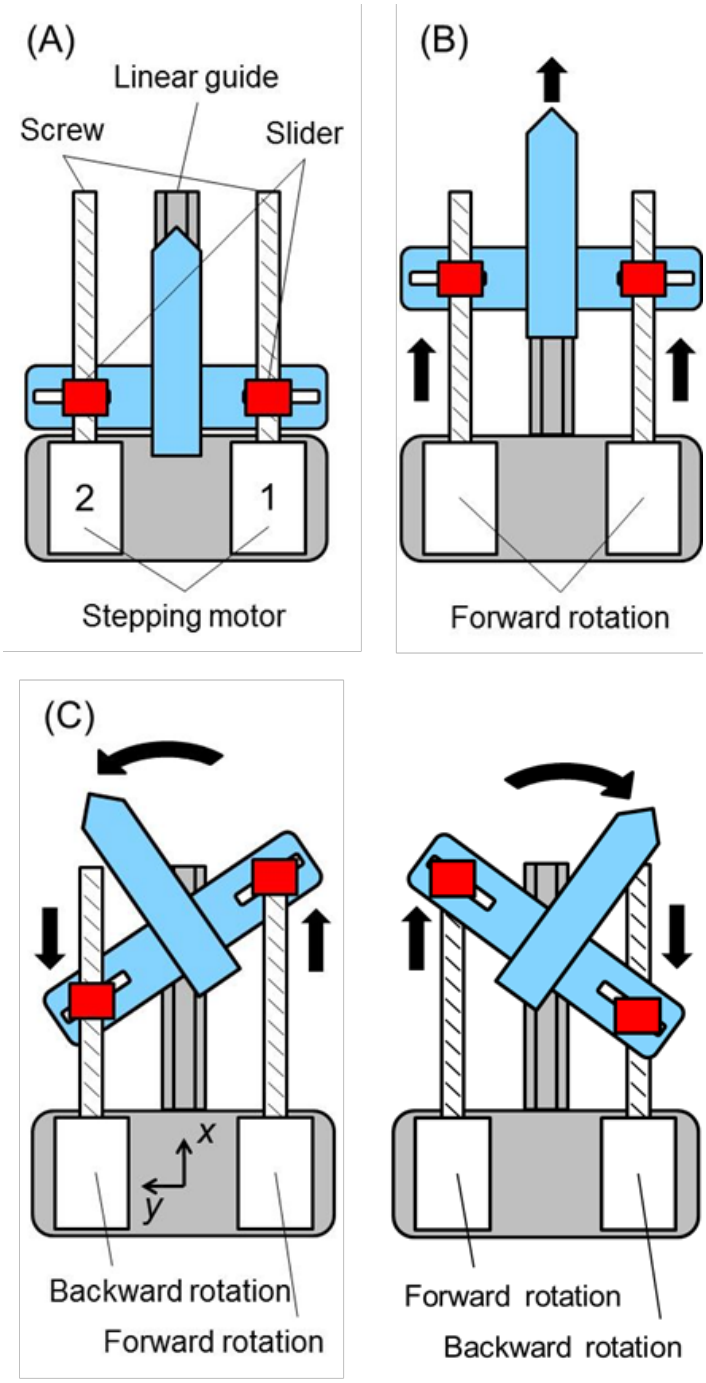


Figure 2: Schematic figure of the sliding screw mechanism installed in the FLSAM; (A) initial condition, (B) lengthened condition, (C) flexible condition

revolution of the stepping motor. In the FLSAM, the L_e is set at 1.25, and the R is 200.

2.2.3. *Geometric configuration of the sliding screw mechanism*

Figure 3 is the schematic of the geometric constraints and the definition of the control parameters by the sliding screw mechanism for the FLSAM. The $O_{i-1}X_{i-1}Y_{i-1}$ joint coordinate system was set at the base of the sliding screw mechanism. The

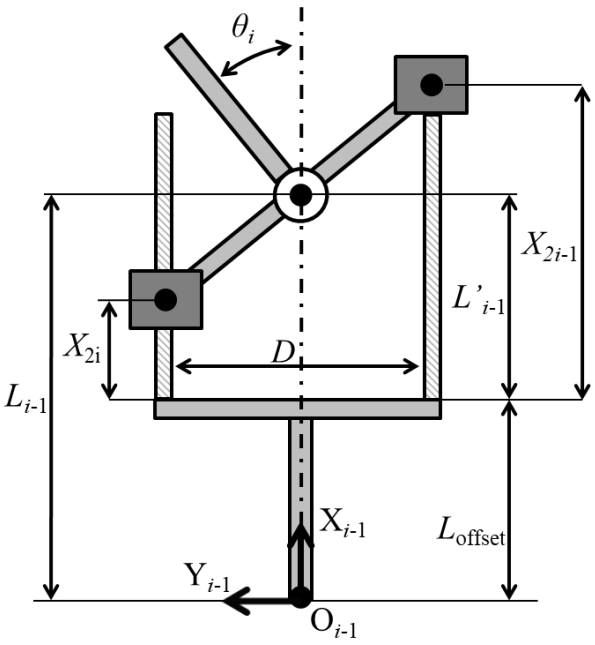


Figure 3: Schematic figure for the parameters of the sliding screw mechanism

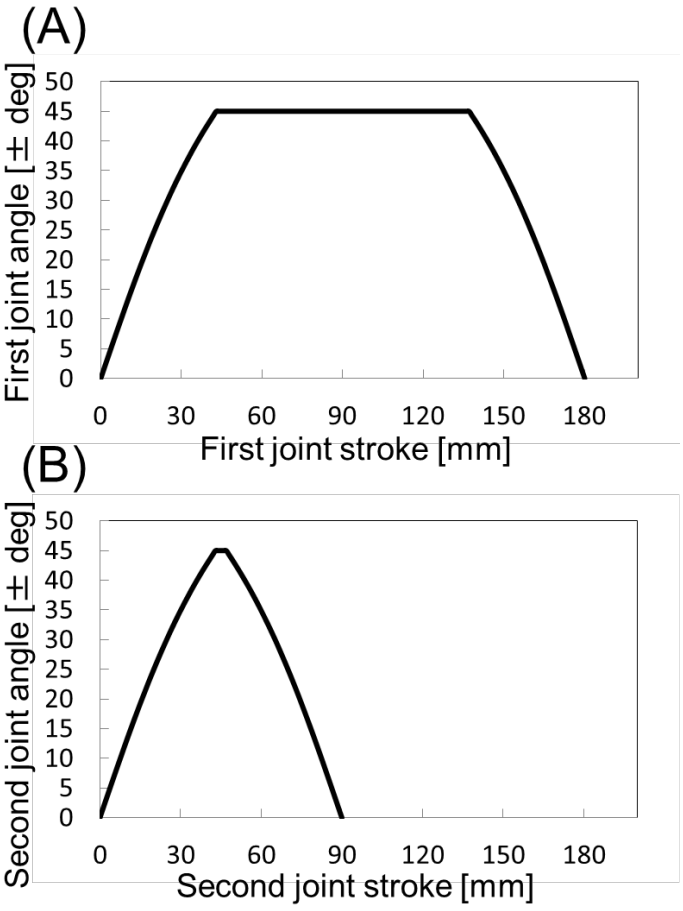


Figure 4: Geometric relationships between the joint range of motion and the joint stroke length; (A) first joint, (B) second joint

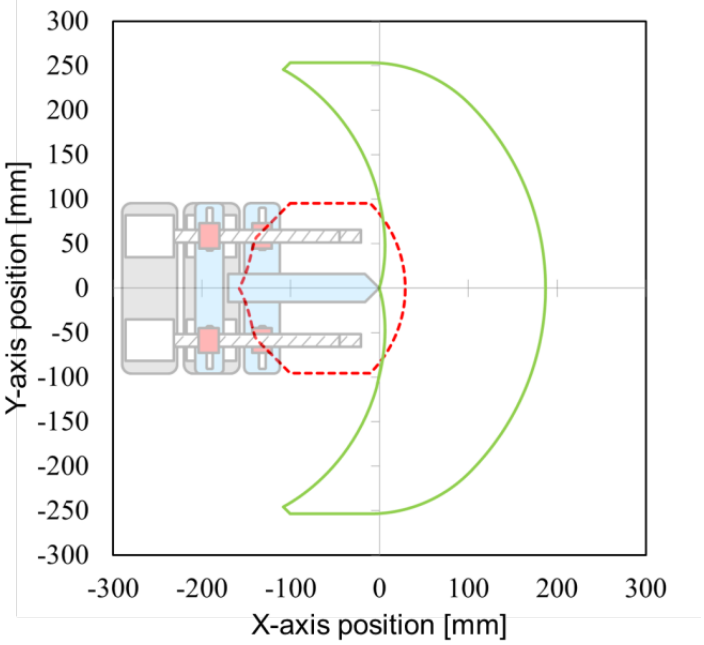


Figure 5: Schematic figure of the range of motion of the FLSAM; red broken line denotes the range of motion for maintaining the maximum flexibility, green line denotes the maximum range of motion

link length, L_{i-1}, and the joint angle, θ_i, were controlled by right- and left-moving sliders. The L'_{i-1}, L_{i-1}, and θ_i were calculated as

$$L'_{i-1} = \left(X_{2i-1} + X_{2i}\right)/2 \quad (2),$$

$$L_{i-1} = \left(X_{2i-1} + X_{2i}\right)/2 + L_{\text{offset}} \quad (3),$$

$$\theta_i = \tan^{-1}\left(\left(X_{2i} - X_{2i-1}\right)/D\right) \quad (4),$$

where L'_{i-1} (i=1, 2) is the link length while lengthening, shortening, and/or rotating. X_{2i-1} is the slider position at the side of the positive in the y-axis, X_{2i} is the slider position at the side of the negative in the y-axis, and D is the length between the two sliders. The position of the end effector of the FLSAM was operated by combinations of the L_{i-1} and θ_i. The movement speed of both sides of the slider is determined so that the operation time is the same according to the L_{i-1} and θ_i, since only the slider on one side works as a singular point.

The sliding screw mechanism has a unique characteristic, in which the maximum range of motion of a joint angle varies with the changes in the link length, as shown in Figure 4. Both maximum joint ranges of motion were set at plus/minus 45 degrees; however, the maximum joint range of motion is 0 degrees when the left and right sliders are at the minimum and/or the maximum position of the link length simultaneously. In addition, these maximum joint ranges of motion were independent parameters, so that it is possible to take the most flexible posture from 43 mm to 47 mm in the first joint and from 43 mm to 137 mm in the second joint. Figure 5 shows a range of motion of FLSAM. In this study, the range of motion was set at the inside area depicted by red line, which maintain the maximum flexibility of FLSAM (joint ranges of motion was at ± 45 deg). When

operating beyond this range, the maximum joint ranges of motion become small and the flexibility of the FLSAM decreases.

2.2.4. *Waterproof structure of the sliding screw mechanism*

The main parts of the FLSAM (sliding screws, sliders, guide rails, and a small camera) are highly water resistant. Cable entry and exit of the waterproof box equipped with an electrical system were waterproofed by cable grounding. Four stepping motors were waterproofed using custom-designed cylindrical housings, as shown in Figure 6. Basically, the housing is made watertight by careful fitting of the acrylic tube, aluminum lid, and O ring. The rotary shaft was waterproofed by combining stainless steel bushes, resin bushes, and a sealing material.

2.2.5. *System configuration of the FLSAM*

Figure 7 shows the system configuration of the FLSAM. The FLSAM's control program was constructed using commercialized software, LabVIEW 12.0.1 (National Instruments Japan Corp.). All control signals from the controller were communicated to Arduino via a serial port, and Arduino then calculated the control signals for the motor drivers (SFE-ROB-12779, SWITCH SCIENCE). Contact switch signals are the only feedback signals to determine the initial and/or the maximum lengthened condition of the FLSAM.

3. Operational tests of the FLSAM

3.1. *Operational test on land*

In order to check whether lengthening-shortening and flexion motions by a sliding screw mechanism of design requirement for image acquisition by a small camera are possible, a ground motion test was carried out. As targets for image acquisition, models of aquatic organisms (turtle, shark, dolphin, and seahorse) and an obstacle were placed randomly. The seahorse model was placed in the shadow of an obstacle that would be a blind spot from the initial posture of the FLSAM. One of the objects of this investigation is abalone, and the catch size is defined as more than 90 mm. For this reason, an imaging target model in this operational test was also the same size of abalone. Each mode size is shown in Table 2. Hence, the FLSAM avoided obstacles due to the lengthening-shortening and flexion motion of the arm and eventually was able to discover the seahorse model. The amount of the stroke length and velocity of each slider of the FLSAM was given as a numerical input from a PC, and the operator checked the posture of the FLSAM and the camera image visually. Figure 8 shows a typical image acquisition and a birds-eye view of the FLSAM operation. In the process of finally photographing seahorses, images were acquired, in order, of the turtle, shark, dolphin, and then the seahorse in this arrangement. When the base unit is stable, FLSAM is not affected by posture change due to disturbance in water because of the characteristics of the sliding screw mechanism. Therefore, when the FLSAM achieves the ideal motion on land, the FLSAM can perform the same operation even in water as on land. It was confirmed that the FLSAM can achieve lengthening-shortening and flexion motions as demanded of the design. Also, image acquisition with a small camera was achieved.

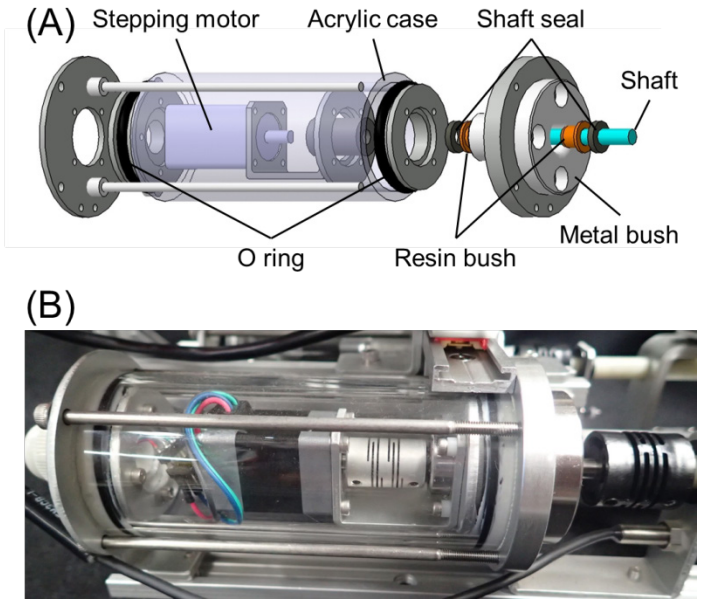


Figure 6: Waterproof motor housing; (A) exploded view of the designed housing, (B) photo of the manufactured housing

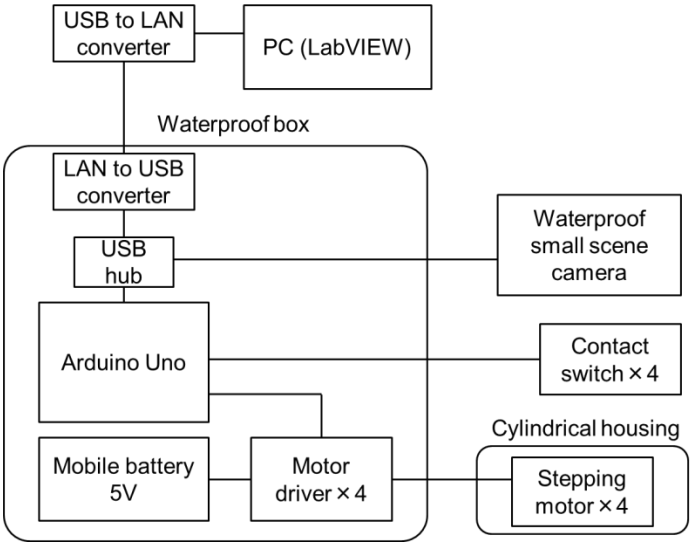


Figure 7: System configuration of the FLSAM

Table 2: Each mode size

Models of aquatic organisms	High [mm]	Length [mm]	Width [mm]
Turtle	40	90	70
Shark	60	170	90
Dolphin	50	110	40
Seahorse	60	40	30

3.2. Operational tests in water

Although it was confirmed that FLSAM could operate without any problems during the ground test, FLSAM was developed to be used in water. Therefore, it was necessary to conduct verification tests to confirm whether it can function properly underwater. Hence, this experiment was planned (1) to check the underwater

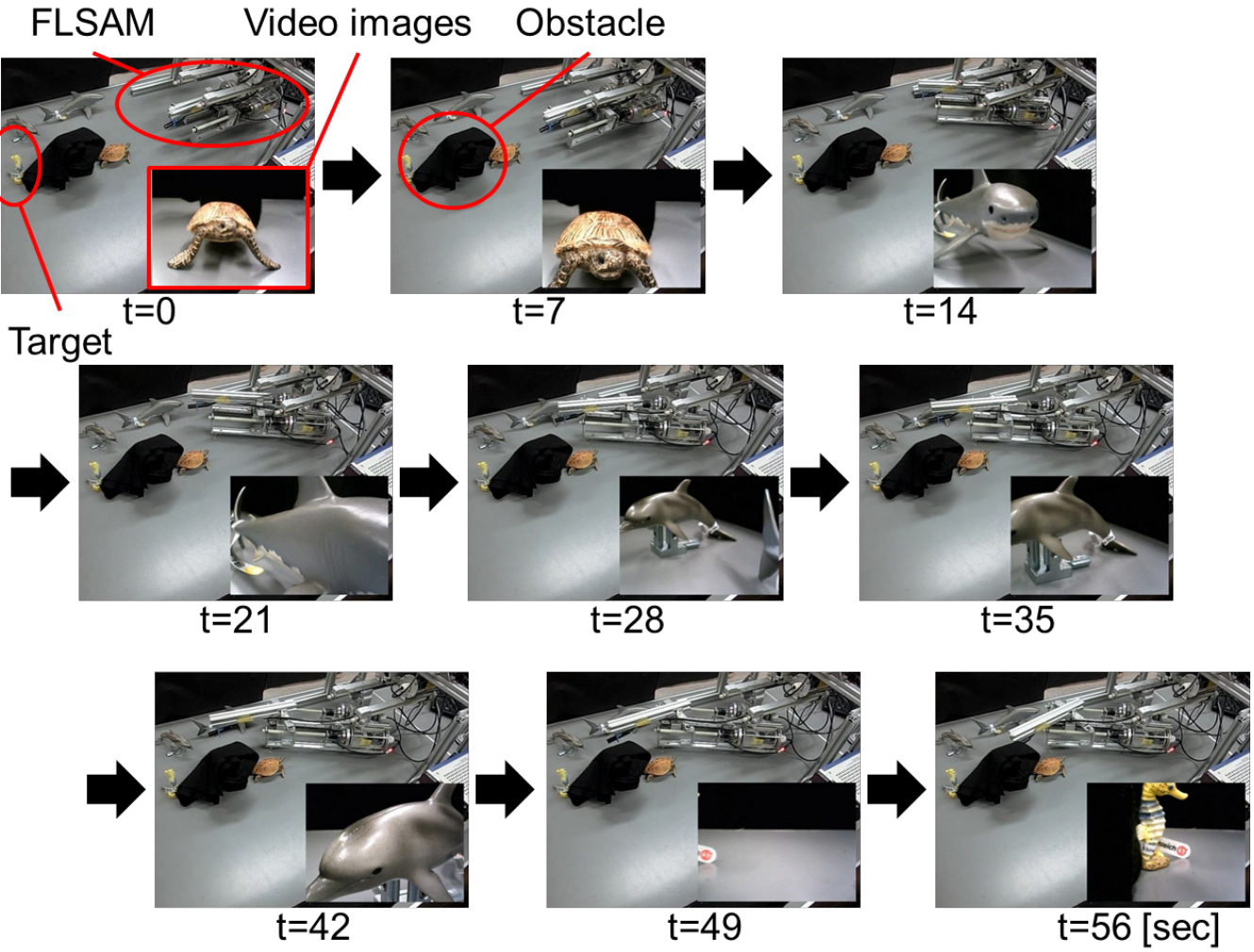


Figure 8: Time series of the state of the operational test conducted on land

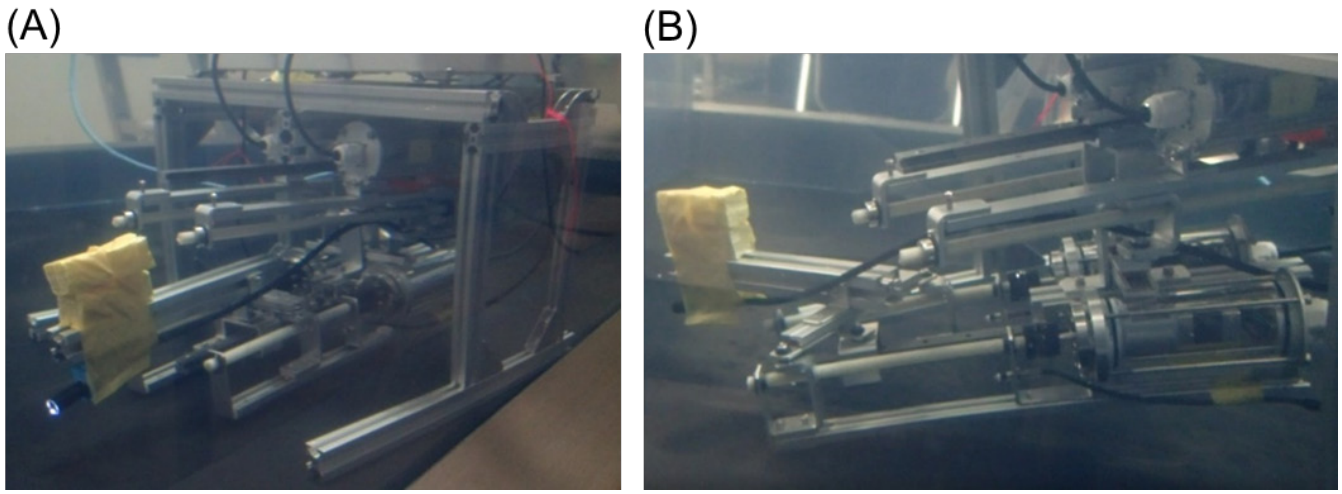


Figure 9: Typical states of the operational test in water; (A) the initial condition, (B) left-side flexed condition

motion, (2) to obtain camera images, and (3) to check the waterproof performance of the FLSAM.

The experiment was carried out in static water conditions (the water tank was 1.57 m deep × 1.2 m wide × 1.27 m high). The FLSAM was placed on the floor of the water tank in its initial condition, and then any arbitrary lengthening/shortening and/or flexion motion was operated manually. To make sure that the sliding screw mechanism of the FLSAM functions properly in water, lengthening-shortening and flexion motions were tested multiple times with reciprocated motions. Further, in order to ascertain whether there is a possibility of water leak into the motor housing due to vibration of the motor and/or the rotation of the shaft, the operation speed is changed while the underwater test continues. Also, video images were checked on the PC simultaneously to determine whether it is possible to acquire images via the small underwater camera installed. The operation of FLSAM was confirmed from the outside of the water tank.

From outside the water tank, it was confirmed that, for 30 minutes, the FLSAM performed every operation on the floor of the water tank. In addition, no water leaked into the housing. Figure 9 shows the typical condition during the operation tests of the FLSAM in water: (A) is the initial condition, and (B) is in the left-side flexed condition. In water, the FLSAM achieved the desired behavior commanded by the operator. However, as the arm lengthened, the FLSAM unexpectedly became tilted, and the small camera at the tip of the arm contacted the floor. This was due to the movement of the center of gravity that occurred when the arm was lengthened. Therefore, it is suggested that attitude control by the support base is required to acquire a stable image using the FLSAM.

4. Autonomous path-tracing experiment for rear-link mechanisms

4.1. Purpose of the path-tracing experiment

Path planning and tracing are key issues for investigating the shadows of rocks. In the case of the FLSAM, path planning and its tracing algorithm should also be installed to achieve obstacle avoidance and to reach into the depths of rock shadows to enable surveying. Methods of controlling a flexible robot arm with a retractable mechanism have been studied (e.g.; series studies of path-tracking control of a moray-type robot arm by same group [16–19], expands and/or contracts arm mechanism [20], and path tracking control of underwater snake-type robot [21]); however, it is impossible to plan a path in environments where the camera cannot obtain data due to blind sports and/or deep rock shadows.

Previous studies have developed methods of manipulating the tip of the arm using a camera image to intrude into an unknown environment (e.g.; surveying robot to avoid obstacles based on screw principle [22], double headed snake-type robot [23], and remotely operated snake-type robot [24]). The FLSAM that can enter unknown environments, such as rock shadows in water, is considered to be effective. Therefore, we adapted to control the FLSAM manually using only the video data obtained from the small installed camera. Figure 10 shows the schematic figure entering the rock shadows by changing the posture of the FLSAM. The operator decides the traveling direction of the tip of the arm using the obtained video data and then manipulates the traveling direction forward/backward and left/right so that only

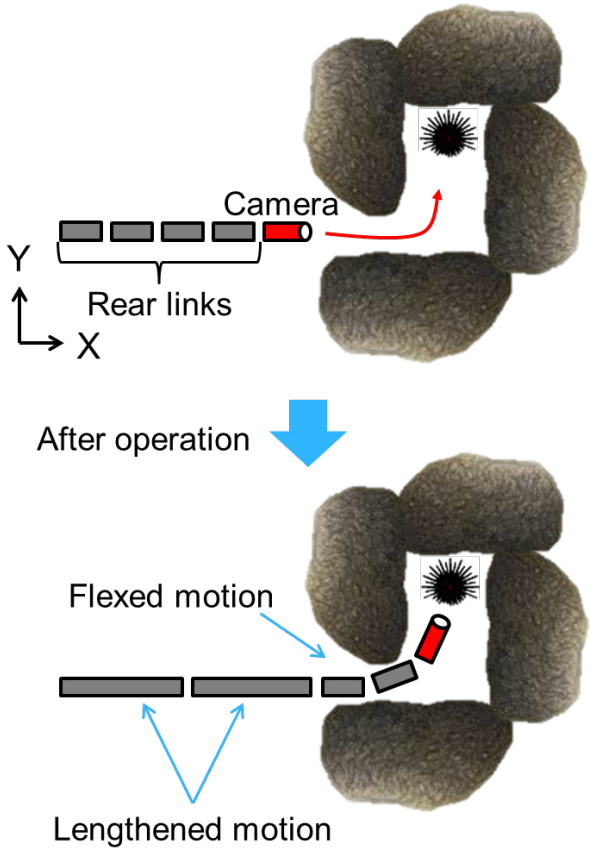


Figure 10: Schematic figure of entering the rock shadow by changing the posture of the FLSAM

the tip of the arm should be controlled manually. This means that the rear links would autonomously follow the movement path of the arm tip. Hence, the purpose of this experiment was to evaluate the accuracy of path tracing of the rear links using the FLSAM's control path-tracing algorithm.

4.2. Path tracing in the rear-link parts of the FLSAM

4.2.1. Definition of the rear-link in FLSAM

FLSAM has two sliding screw mechanism joints. That means that if the second link of the FLSAM equipped with the small camera is defined as the arm tip, the rear link is only the first link. However, it is difficult to evaluate the validity of the path-tracing algorithm by measuring the movement of one link. Hence, as shown in Figure 11AB, the virtual arm tip is defined as existing at the tip of the second link of the FLSAM, and the first link and the second link of the FLSAM are set as the rear links.

4.2.2. Target path defined by the Spline interpolation curve

The first and second rear links of the FLSAM autonomously trace the tip's moving path. In this research, we prepared the target path under the assumption that the tip's moving path is generated

by the movement of the virtual tip, and the first and second links of FLSAM trace this virtual path. For the target path, a cubic Spline interpolation curve capable of creating a two-dimensional curve in space was used. The details are described in Appendix-1; however, the first derivatives of the start point of the Spline curve are set to 0 in the boundary condition (A5). With respect to this definition, the slope of the Spline curve at the start point becomes 0. Therefore, the Spline curves of the target path and the X axis are continuous at the origin of the global coordinate system, and discontinuity of the arm joint angle can be avoided.

4.2.3. *Development of a path-tracing algorithm in the rear-links of the FLSAM*

Figure 11C shows a schematic of path tracing in a rear link of the FLSAM. In this study, the rear links of the FLSAM trace the Spline curve created as the target path to evaluate the validity of the path-tracing algorithm. In order to trace the target path, it is necessary to calculate each joint angle as output with the target path and each link length as input. Thus, as shown in Figure 11C, links L_0 and L_1 should lengthen and/or shorten to control the joint angles for tracing the target path.

At first, we defined each joint coordinate and link length as follows: the origin of the arm base as $J_0=[(x_0, y_0) = (0, 0)]$, the position of the first joint as $J_1=[(x_1, y_1) = (x_0 + L_0, 0)]$, the position of the later joints as $J_{j+1} = [x_{j+1}, y_{j+1}]$ (j = 1 to n); n is the number of total joints, the position of the arm tip is Jc (= J_{n+1}) = $[x_{n+1}, y_{n+1}]$, and the link lengths are L_j. J_0 and J_1 are determined by the initial arrangement and the link length of L_0. The outline of the procedure for deriving each joint position from each link length is as follows:

(1) let J_{j+1} be the point of intersection with the circle C_j of radius L_j centered on J_j and the Spline curve,

(2) let J_{j+2} be the point of intersection with the circle C_{j+1} of radius L_{j+1} centered on J_{j+1} derived in procedure (1) and the Spline curve,

(3) then repeat procedure (2) until Jc.

Here, we will describe procedures (1) and (2) in greater details. Where the C_j intersects with the interpolation interval, $S_i(x)$ is determined by the following conditional expression, since the Spline curve has a number of ith interpolation intervals:

$$(x_i - x_j)^2 + (y_i - y_j)^2 < L_j^2 < (x_{i+1} - x_j)^2 + (y_{i+1} - y_j)^2 \quad (5).$$

Then, solve the following simultaneous equations to find x_{j+1}:

$$C_j : (x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2 = L_j^2 \quad (6),$$

$$S_i(x_{j+1}) = a_i x_{j+1}^3 + b_i x_{j+1}^2 + c_i x_{j+1} + d_i (= y_{j+1}) \quad (7),$$

where x_{j+1} satisfies $x_i < x_{j+1} < x_{i+1}$. By repeating these procedures, each joint position is derived. Finally, each joint angle is derived by the following formula:

$$\theta_j = \tan^{-1} \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \theta_{j-1} \quad (8).$$

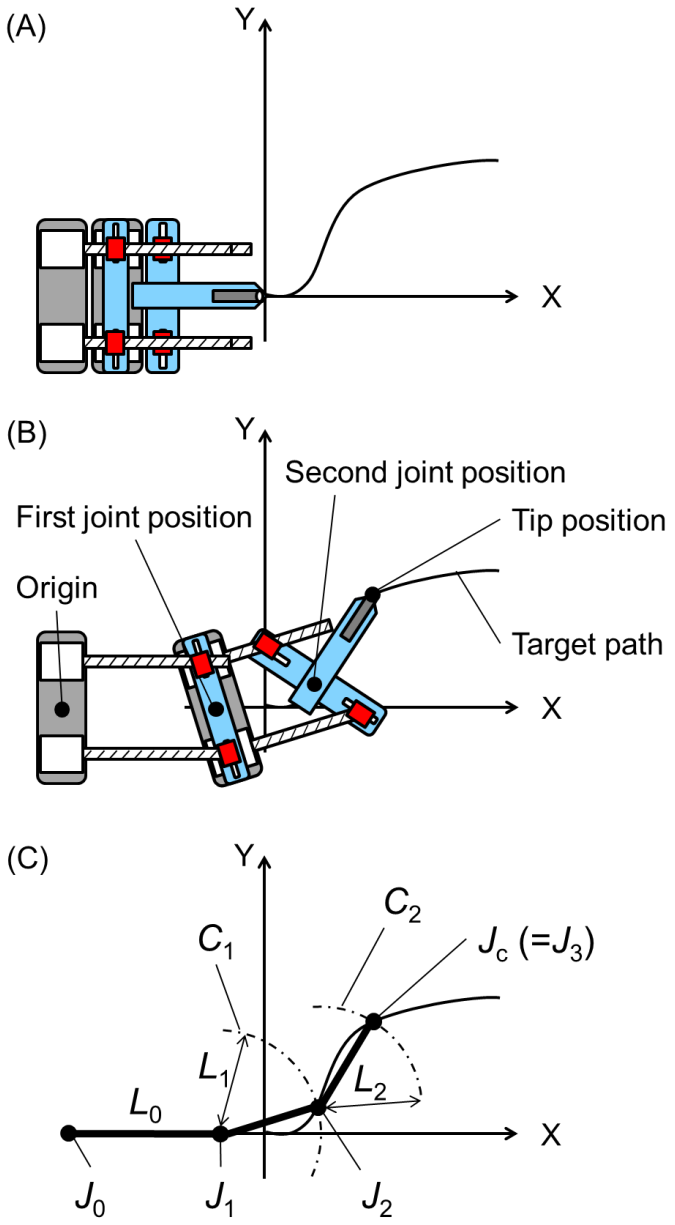


Figure 11: Schematic figure of the definition of the rear links in the FLSAM; (A) state of the arm tip on the ordinate of the global coordinate system, (B) definition of the rear links in FLSAM, (C) definition of each joint coordinate and link length

In this study, using each link length, L_j, and joint angle, θ_j, as input values, the path-tracing algorithm for the rear links calculates the output value as the number of pulses input to each stepping motor of the sliding screw mechanism. Thus, in the case of the FLSAM, $J_0=[(x_0, y_0) = (0, 0)]$ is the base position, the position of the first joint is $J_1=[(x_1, y_1) = (x_0 + L_0, 0)]$, the position of the second joint is $J_2 = [x_2, y_2]$, and the position of the arm tip is Jc (= J_3) = $[x_3, y_3]$, since the number of joints is two.

4.2.4. *Simulation of a path-tracing algorithm in the rear links of the FLSAM*

Computational simulation was performed to evaluate whether the proposed rear-link path-tracing algorithm works precisely enough to trace the target path (an arbitrary Spline curve). This simulation was done with a custom-designed graphical user

interface using LabVIEW 12.0.1. Figure 12 shows the simulation results of the trajectories of the tip and the second joint calculated by using each link length and joint angle output from the path-tracing algorithm. Thus, we confirmed that each joint is moving on the Spline curve completely, and that the path-tracing algorithm of the rear link of the FLSAM is correct.

4.3. Methods

The accuracy of the FLSAM motion was measured using a real-time three-dimensional motion capture system (VENUS3D, Nobby Tech., Ltd.). The conceptual motion pattern for path tracing in the FLSAM is shown in Figure 12, which shows that the initial posture was determined, in which the origin of the global coordinate system of the FLSAM was harmonized with the position of the arm tip. The tip of the end effector traced a predetermined target path, and then the rear links followed the tip motion, as shown in Figure 11B. For each slider of the FLSAM, the maximum moving speed was set to 5 mm/s. In addition, the L_{i-1} (i=1, 2) was lengthened by a step of 2 mm. Three target paths (a figure "S" pattern, an exponential figure pattern, and a sigmoid curve figure pattern) and three mirrored target paths were defined before the experiment and were started from the origin of the global coordinate system. Each pattern was followed in ten trials. The motion trajectory of the arm tip and the second joint were measured to compare the accuracy of the target paths and the motion of the FLSAM. All experiments were conducted on land.

4.4. Results and Discussion

Figure 13 shows the results of all ten trials of tracing (A) the figure "S" and its mirrored patterns, (B) the exponential figure and its mirrored patterns, and (C) the sigmoid curve figure and its mirrored patterns. Time series of the position of the tip in X and Y axis are shown in Fig. 14. It was observed that all ten path trials are highly reproducible, the tip of the FLSAM traced the target paths, and the tracking results and errors were also symmetrical in mirrored paths. These results confirm the validity of the path-tracing algorithm. The maximum and the root mean square of the deviations were (2.25 mm, 1.7 mm) for pattern (A), (1.86 mm, 0.66 mm) for pattern (B), and (4.7 mm, 1.48 mm) for pattern (C), respectively. All of these deviations were under 4.0% (2.8 mm) of the total width of the FLSAM (140 mm). Although the same tendencies had found in the previous studies (e.g.; series studies of path tracing of a snake-type robot using screw driving theorem by same group [25-26]) that the deviations are approximately 4.0%, it is suggested that the path-tracing algorithm for the FLSAM while path tracing is effective, and minimizing the deviations is a task for future studies. In addition, the movement error of 4% (2.8 mm) is below 4% or less of the survey target size of 90 mm or more. In this respect, the validity of the path-tracing algorithm of the rear links of the FLSAM proposed in this paper was confirmed.

We would like to construct an algorithm to achieve three-dimensional motion by changing the arrangement and the number of sliding screw mechanisms of the FLSAM. Further studies should be needed at least the following points; (1) to improve the accuracy of path tracing, since the tracing deviations would accumulate if the FLSAM would be expanded to increase the number of joints, making it possible for a large error to occur at the tip position and (2) to generate a following path of rear links from the movement path of the tip instead of the Spline curve used in this study.

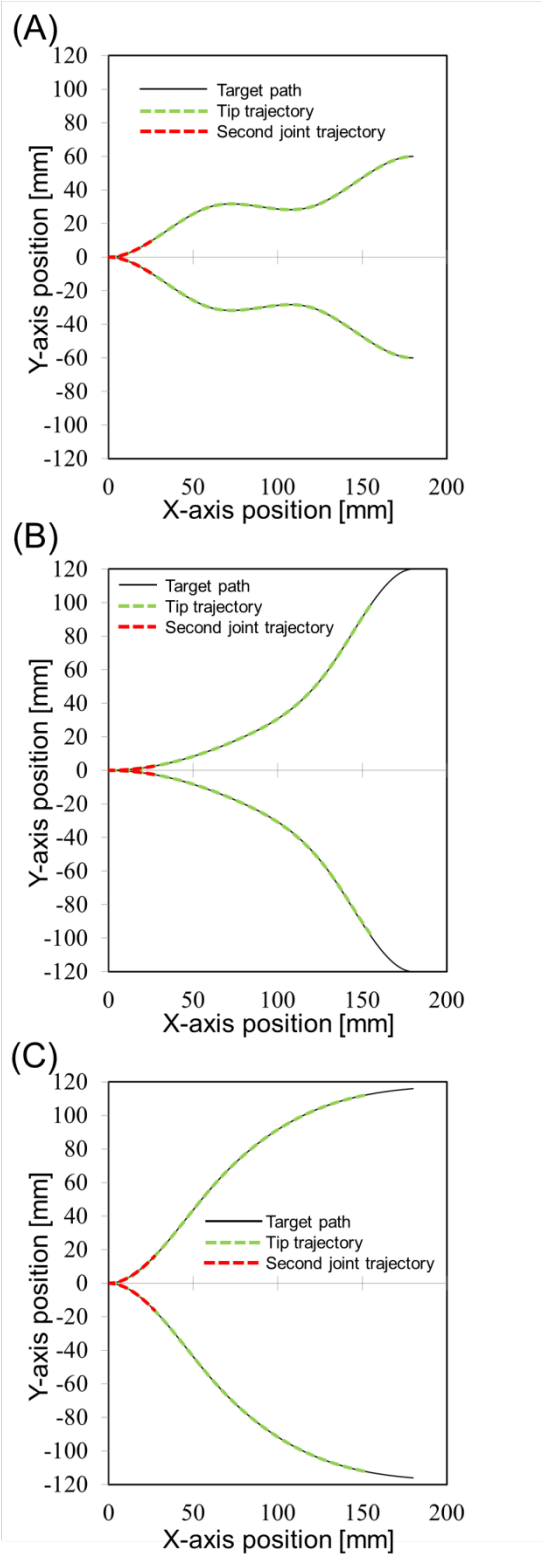


Figure 12: Simulation results of the path tracing algorithm; (A) figure "S" and its mirrored pattern, (B) the exponential figure and its mirrored patterns, and (C) the sigmoid curve figure and its mirrored patterns. Black lines are the target paths, green lines are the tip trajectories, and red lines are the trajectories of the second joint position.

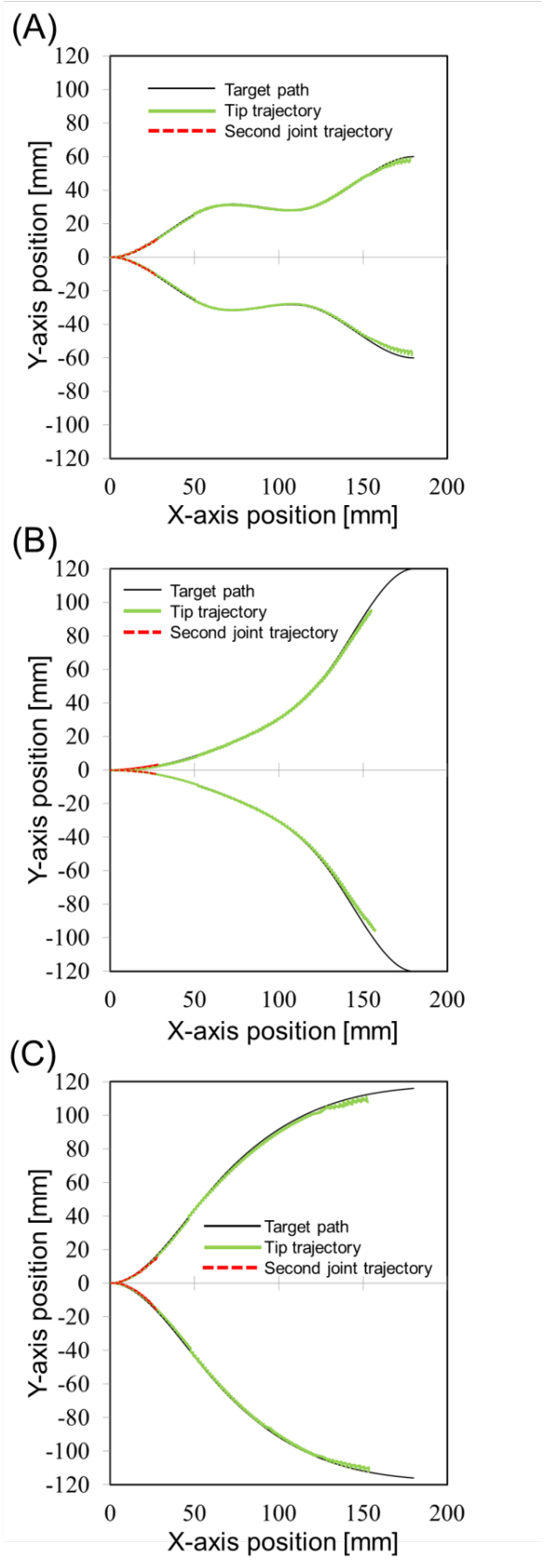


Figure 13: Results of path tracing by the FLSAM; (A) figure "S" and its mirrored pattern, (B) the exponential figure and its mirrored patterns, and (C) the sigmoid curve figure and its mirrored patterns. Black lines are the target paths, green lines are the tip trajectories, and red lines are the trajectories of the second joint position.

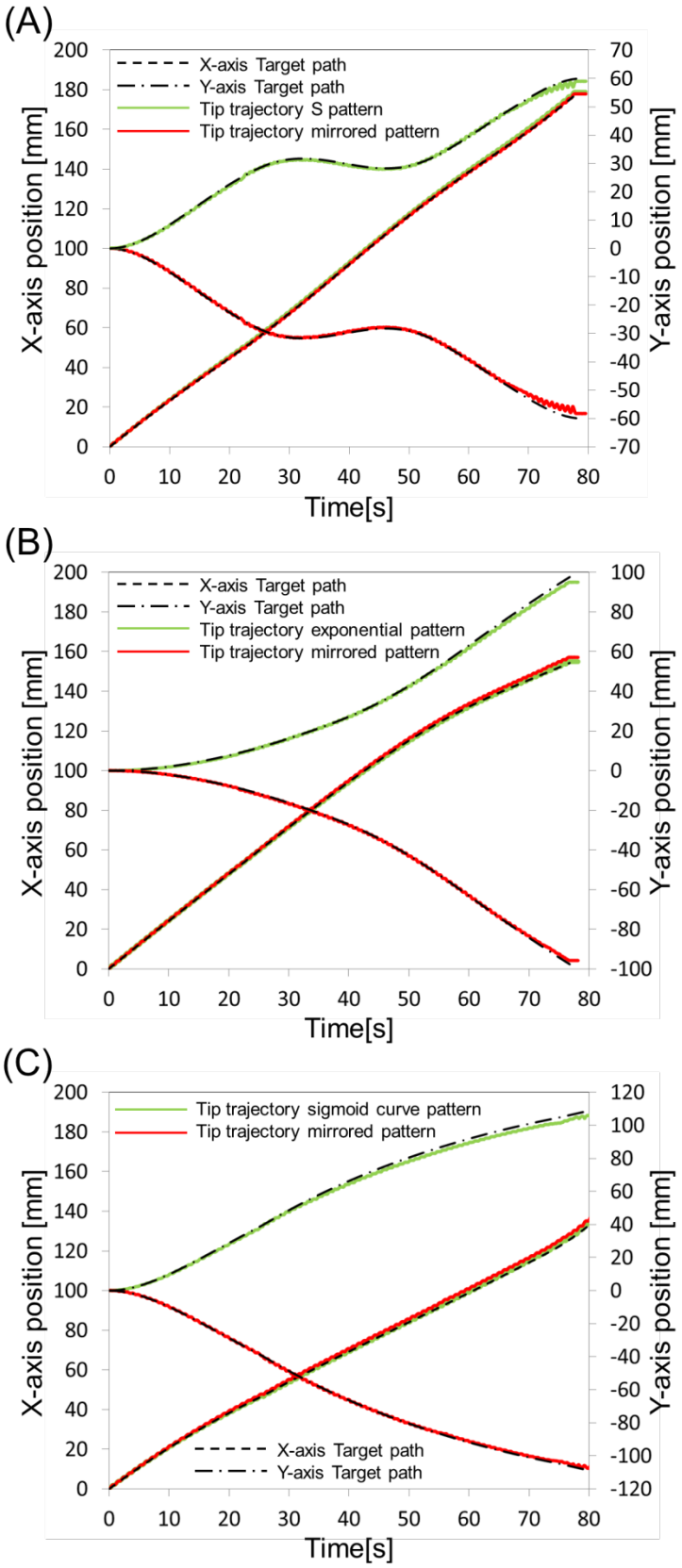


Figure 14: Time series of the position of the tip in X and Y axis; (A) figure "S" and its mirrored pattern, (B) the exponential figure and its mirrored patterns, and (C) the sigmoid curve figure and its mirrored patterns.

5. Conclusion

In order to realize a flexible underwater robot arm, we adopted a sliding screw mechanism and developed a prototype device, a "FLSAM." Path tracing in the rear-link mechanisms and underwater operation tests were executed, and the results suggested that the sliding screw mechanism is sufficient for construction of a flexible underwater robot arm. However, to obtain stable video data, further modification, such as using an autonomous postural control base, is necessary. Further study is also needed to improve the path-tracing control performance using kinematic and dynamic analysis of the FLSAM.

Appendix-1: Calculation of the cubic Spline interpolation curve

For the target path, a cubic Spline interpolation curve capable of creating a two-dimensional curve in space is used. The *n*th Spline interpolation is continuous to the (*n*-1)th derivative function, and oscillations between the representative points hardly occur. In addition, it is effective for generating a target path simulating an obstacle-avoidance route, since it is possible to produce a complicated shape in accordance with the given interpolation points. In this study, a cubic Spline curve was calculated by setting the first interpolation point as the origin of the global coordinate system and sequentially setting the next interpolation point at an arbitrary position in two-dimensional space.

When (*n*+1) points are given, there are *n* interpolation divisions, and the polynomials exist in each division as follows:

$$S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \qquad \text{(A1)} ,$$

where $S_i(x)$ is the cubic Spline curve function in the *i*th division; a_i, b_i, c_i, and d_i are the coefficients in *i*th division; and *i* (*i*=1 to *N*) equals the number of deviations. The coefficients of the polynomial (a_i, b_i, c_i, and d_i) are derived using the continuous condition (A2–4) at the interpolation point and the two boundary conditions (A5–6), as follows:

$$S_i(x_{i+1}) = S_{i+1}(x_{i+1}) \qquad \text{(A2)} ,$$

$$S_i'(x_{i+1}) = S_{i+1}'(x_{i+1}) \qquad \text{(A3)} ,$$

$$S_i''(x_{i+1}) = S_{i+1}''(x_{i+1}) \qquad \text{(A4)} ,$$

$$S_1'(x_1) = 0 \qquad \text{(A5)} ,$$

$$S_N'(x_{N+1}) = const \qquad \text{(A6)} ,$$

where the ith interpolation point coordinate is taken as (x_i, y_i), and $S_N'(x_{N+1}) = 0$ in this study. In the boundary condition (A5), the first derivatives of the start point of the Spline curve are set to 0, so that the Spline curves of the target path the and X axis are continuous at the origin of the global coordinate system.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 17K08029.

References

[1] Y. Iwamochi, M. Takagi and T. Miyoshi "Flexible Lengthening-Shortening Arm Mechanism for an Underwater Vehicle", Proceedings of 2017 IEEE International Conference on Mechatronics and Automation, MP1-6(6), 2017. http://doi.org/10.1109/ICMA.2017.8015921

[2] J. Pyo. H. Cho, H. Joe, T. Ura and S. Yu, "Development of hovering type AUV "Cyclops" and its performance evaluation using image mosaicking", Ocean Engineering, vol.109, pp.517-530, 2015. https://doi.org/10.1016/j.oceaneng.2015.09.023

[3] S. Kawamura, "Underwater robot development for manipulation task and their uses in Biwa lake", IFAC-PapersOnLine, vol.48, No.2, pp.14-19, 2015. https://doi.org/10.1016/j.ifacol.2015.06.003

[4] P. Gao. B. Zgeng, J. Liang and C. Ren, "Development of an Underwater Robot for Sediment Soil Sampling", Proceedings of the 3rd International Conference on Mechatronics and Industrial Informatics, pp.6-11, 2015. http://doi.org/10.2991/icmii-15.2015.2

[5] K. Satoh, M. Takagi, H. Mori and T. Miyoshi, "Development of an Underwater Robot for Researching of Underwater Environmental", The Japan Society of Mechanical Engineers, The Proceedings of JSME annual Conference on Robotics and Mechatronics (Robomec), Session ID: 1A2-16b5, 2016. http://doi.org/10.1299/jsmermd.2016.1A2-16b5

[6] N. Sakagami, M. Shibata, H. Hashizume and Y. Hagiwara, "Development of a Human-Sized ROV with Dual-Arm", The proceedings of the Oceans 2010 IEEE, 2010. http://doi.org/10.1109/OCEANSSYD.2010.5603897

[7] H. Shim, B.H. Jun, P.M. Lee, H. Beak and J. Lee, "Workspace control system of underwater tele-operated manipulators on an ROV", Ocean Engineering, vol.37, No.11-12, pp.1036-1047, 2010. https://doi.org/10.1016/j.oceaneng.2010.03.017

[8] P. Liljeback, O. Stavdahl, K. Pettersen and J. T. Gravdahl, "A review on modelling, implementation, and control of snake robots", Robotics and Autonomous Systems, vol.60, No.1, pp.29-40, 2012. https://doi.org/10.1016/j.robot.2011.08.010

[9] S. Hirose ans H. Yamada, "Snake-like robots [Tutorial]", IEEE Robotics & Automation Magazine, vol.16, No.1, pp.88-98, 2010. http://doi.org/10.1109/MRA.2009.932130

[10] G. Granosik, J. Borenstein and M. G. Hansen, Serpentine Robots for Industrial Inspection and Surveillance, Industrial Robotics: Programming, Simulation and Applications, 2006. http://doi.org/10.5772/4921

[11] S. Stromsoyen, "Propulsion Methods for Under Water Snake Robots - Investigation and Simulation Using Foil for Propulsion of a Snake Robot", Master thesis, Norwegian University of Science and Technology, 2015.

[12] A. Wolf, H. B. Brown, R. Casciola, A. Costa, M. Schwerin, E. Shamas and H. Choset, "A Mobile Hyper Redundant Mechanism for Search and Rescue Tasks", The proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003. http://doi.org/10.1109/IROS.2003.1249309

[13] R. Buckingham, V. Chitrakaran, R. Conkie and G. Ferguson et al, "Snake-Arm Robots: A New Approach to Aircraft Assembly", SAE Technical Paper, 2007-01-3870, 2007. https://doi.org/10.4271/2007-01-3870.

[14] D. Palmer, S. C. Guzman and D. Axinte, "Real-time method for tip following navigation of continuum snake arm robots", Robotics and Autonomous Systems, vol.62, No.10, pp.1478-1485, 2014. https://doi.org/10.1016/j.robot.2014.05.013

[15] H. Nakagawa and K. Yoshimura, "Super-multi Joint Manipulator Composed of Linking 3 DOF Joint Unit", The Japan Society of Mechanical Engineers, The proceedings of the JSME annual meeting, Session ID: 834, May 2005. http://doi.org/10.1299/jsmemecjo.2005.5.0_7

[16] S. Hirose and S. Ma, "A moray drive of multi-joint manipulator", Journal of the Robotics Society of Japan, vol. 8, No.1, pp.9-16, 1990. http://doi.org/10.7210/jrsj.8.9

[17] S. Ma, S. Hirose and K. Yokoshima, "2 DOF moray drive for multijoint manipulators", Journal of the Robotics Society of Japan, vol.14, No.3, pp.436-443, 1996. http://doi.org/10.7210/jrsj.14.436

[18] S. Ma and I. Kobayashi, "Path-tracking control of a moray-type robot arm with consideration of actuator's actuation characteristics", Journal of the Robotics Society of Japan, vol.18, No.1, pp.135-141, 2000. http://doi.org/10.7210/jrsj.18.135

[19] S. Ma and Y. Ohmameuda, "Obstacle avoidance control of moray-type robot arm", Transactions of the Japan Society of Mechanical Engineers Series C, vol.68, No.670, pp.1791-1797, 2002. http://doi.org/10.1299/kikaic.68.1791

[20] Y. Cao, V.J. Modi, C.W.de Silva, M. Chen and A.K.Misra, "Trajectory tracking experiments using a novel manipulator", Acta Astronautica, vol.52, pp.523-540, 2003. https://doi.org/10.1016/S0094-5765(02)00123-6

[21] E. Kelasidi, P. Liljeback, K. Y. Pettersen and J. Tommy Gravdahl, "Integral line-of-sight guidance for path following control of underwater snake robots: Theory and Experiments", IEEE Transactions on Robotics, vol.33, No.3, pp.610-628, 2017. http://doi.org/10.1109/TRO.2017.2651119

[22] M. Hara, S. Satomura and F. Matsuno et al, "Development and Test of Fundamental Action of Snake-Like Robot with Driving Mechanism Based on Screw Principle", The Japan Society of Mechanical Engineers, The Proceedings of JSME annual Conference on Robotics and Mechatronics (Robomec), Session ID: 2P2-D27, 2006. http://doi.org/10.1299/jsmermd.2006._2P2-D27_1

[23] T. Kamegawa and F. Matsuno, "Development of a remote-controlled double headed snake-like rescue robot KOHGA", Journal of the Robotics Society of Japan, vol.25, No.7, pp.1074-1081, 2007. http://doi.org/10.7210/jrsj.25.1074

[24] K. Ito, R. Murai and K. Nakamichi, "Development of snake-like rescue robot designed for ease of use ", Journal of the Robotics Society of Japan, vol.27, No.4, pp.419-426, 2009. http://doi.org/10.7210/jrsj.27.419

[25] M. Hara, S. Satomura, H. Fukushima, T. Kamegawa, H. Igarashi and F. Matsuno, "Control of a Snake-like Robot Using the Screw Drive Mechanism", Proceedings of 2007 IEEE International Conference on Robotics and Automation, Roma Italy, 2007. http://doi.org/10.1109/ROBOT.2007.364074

[26] H. Fukushima, M. Tanaka, T. Kamegawa and F. Matsuno, "Front-Unit-Following Control of a Snake-like Robot Using Screw Drive Mechanism", Journal of the Robotics Society of Japan, vol.28, No.6, pp.707-714, 2010. http://doi.org/10.7210/jrsj.28.707