ASTES

Advances in Science, Technology & Engineering Systems Journal

VOLUME 4-ISSUE 3|MAY-JUNE 2019

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 35 accepted papers in Telecommunication domain.

Editor-in-chief
Prof. Passerini Kazmersk

ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 4 Issue 3 **May-June 2019**

CONTENTS

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3,01-09 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Detection the Places of the Heat Energy Leak on the Underground Thermal Pipelines Using the Computer System

Iryna Zhuravska*, Dmytro Lernatovych, Oleksandr Burenko

Petro Mohyla Black Sea State University, Mykolaiv, 54003, Ukraine

A R T I C L E I N F O

Article history:
Received: 28 March, 2019
Accepted: 16 April, 2019
Online: 04 May, 2019

Keywords:
Thermal energy loss
Underground heat pipelines
Pyrometer
GPS coordinates
Mobile application
Heat map
Google FireBase

A B S T R A C T

The purpose of this paper is to improve the thermographic method for detection the places of heat energy losses on the urban underground thermal pipelines using the computer system. This goal was achieved by building a heat map of the surveyed heating main with data layering directly on a computer map of the area, which can be accessed from a mobile device by authorization via the Internet (profile on the social network Google, Facebook, Twitter, etc.). An analytical model was proposed for converting the temperature values measured over the underground heating main by a non-contact digital pyrometer to an image of the RGB color model using the parameters of the HSV color model. The implementation of the integrated Google technologies complex (Google Firebase, Firebase Authentication, and Google Maps) in the technological process of urban housing and public services was realized. The information-measuring computer system (IMCS) was developed. The proposed improvements in the thermographic method provided independently from each other work of different service teams due to the exchange of data between structural units through the cloud data storage. The results of the practical research allowed detecting the place of the underground breakthrough of the heat pipe quickly, to position the repair equipment over the accident section precisely using the GPS module on the mobile device. Thus, the size of the dug road section was minimized, material resources were saved. The experimental studies carried out confirmed the correctness of the theoretical assumptions.

1. Introduction

The urgency of the searching problem and the reduction of heat loss via the construction elements is due to the need for a rigorous energy saving. This problem is exacerbated by the work of heat power engineering in many countries because of limited natural fuel resources. Therefore, significant attention is paid to reducing operating costs, and timely detection and elimination of emergencies on the heat mains. The detection of heat energy losses on underground thermal pipelines is the most problematic.

In the conditions of the modern urban economy, losses on heating lines are enormous. Increased heat losses along the pipelines and massive coolant leaks, in addition to economic damages, cause excessive thermal pollution.

In order to comply with environmental requirements, it is necessary to implement a set of measures to eliminate the causes of heat leaks on heating mains. It is necessary to improve the methods allowing receiving thermal energy on the main highways.

*Corresponding Author: Iryna Zhuravska, 68 Desantnykiv str., Mykolaiv, 54003, Ukraine. Email: iryna.zhuravska@chmnu.edu.ua

2. Related Works

2.1. Overview of methods for determining losses in heat pipes

Methods for detecting and eliminating defects in main pipelines have been described by many researchers and scientists. [1, 2]. All over the world, common attention is paid to non-destructive techniques and technologies like acoustic leak detection, thermal imaging, tracer gas, radiotracers, etc. [1–4].

Scientifically based methods for non-destructive testing of leakages in pipelines are methods based on analysis of transient processes in pipes, in particular, the impedance method [5]. However, these methods have substantial limitations in the representation of system characteristics such as pipeline length (no more than 90 m) and location of leakage.

The detection and recovery time of the carrying capacity of pipeline sections with local damage can be significantly reduced through the measuring and statistical analytical procedures with using of the information-measuring computer system (IMCS).

The method of the negative pressure wave (NPW) is one of the most modern. Using this method, it is possible to track the appearance of a leak on the heating main without leaving the pipeline zone, while remaining at the central control station [6, 7]. At the time of occurrence of leakage of fluid in the pipeline, dilution waves appear. The pressure wave reaches the measuring points on either side of the leak and is recorded at the central dispatch post. The computing system processes the results of the incoming information and promptly calculates the leak location (within 1 minute). However, the accuracy of this method does not exceed ± 200 m.

Currently, sophisticated methods that combine acoustic and correlation principles of leak detection with remote registration on servers over the air are used on heating mains. However, the authors themselves have determined that such practices are ineffective when the heating main is located deeper than 2.5 m [8]. Also, the instruments used in the implementation of these methods do not work in the presence of increased noise (for example, moving vehicles). Considering that the majority of urban heating pipelines are laid under highways, these factors have a negative impact on the research results. The harmful properties of these methods also include the fact that to obtain the measurement results it is necessary to insert the peak of the sensor a good swift kick in the road ground (thus, destroying the highway). Also, there is a danger of the sensor damage by a stone, metal object, the ice, etc. on its way.

2.2. *The advantages and features of using the thermographic method*

From the analysis of the above methods, we can conclude that the thermographic method is the most promising. Infrared is a proven noninvasive technology, which locates leaks in the systems of heat pipes without any destruction owing to the digging of highways, under which heating mains in cities are usually laid.

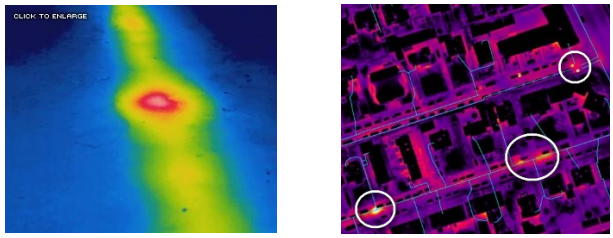
Thermal infrared cameras can see leaking heating pipes, water/steam leaks, etc. If used properly, infrared technology is one of the most highly useful diagnostic tools available today.

It's available to detect places of the local heat leak on the underground thermal pipelines using the IMCS. Such a system should use infrared thermographic devices which can transmit data through wireless ports. This solution helps to determine thermal energy loss location to eliminate unnecessary digging, destruction and expense.

Figure 1, a shows how infrared technology is used to detect a heat loss from steam pipes located under the streets of New York and attributed as a breakdown [9]. These thermal mains demand immediate repair.

Infrared technology in combination with automatic analysis of aerial thermal images allows localizing leaks on district heating pipes immediately in several neighborhoods as shown in Figure 1, b.

In this case, it is essential that the diagnostic equipment moves strictly within the spatial corridor (street) along the heating main [10]. For example, such methods as remote sensing or airborne thermography are widely used in Sweden and Norway [11].



a. b.

Figure 1. Infrared detection the local (a) and large-scale (b) places of heat loss following an accident

Figure 2 shows the location of a non-leaking, but un-insulated unions on the heat pipelines (Mykolaiv, Ukraine). Thanks to the thermal imager, it is determined that this is a non-emergency situation and, in this case, a planned repair is needed. In addition, the thermal imaging method has helped not to unearth the entire section of the roadway over the heating main, but to carry out repairs through point pits.

Figure 2. Detection of the places technological heat energy losses through un-insulated unions

Results of heat pipeline leak detection have statistically analyzed using the methods of variance analysis and the pairwise comparison methods. The researches have shown that a higher going speed of IR camera would reduce the effect of noise on the collected thermal contrast and, therefore, would give better results of prediction the leak location [12].

2.3. *On lineaments and permits for the use of UAVs for moving thermographic equipment*

Usually, the leak imaging system is placed on rotor-wing and fixed-wing UAVs [13, 14]. One of the best modern hardware solutions for studying energy losses on the thermal line is the thermal imaging cameras mounted on a quadrocopter. By applying a UAV, it is possible to specify its operation area accurately [15]. But this solution is relatively expensive and costs about USD 20,000.

Moreover, most often such systems don't determine accurate GPS coordinates of the accident and are therefore useless for the repair team. At present, the leak detection team should wait for repairers at the identified accident section or create written and photo records.

It is positive that modern leak detection methods locate heat leak accurately to 1 sq. m, so any excavation reduces to a minimum when the pipes are repaired. However, such systems also don't have an accurate reference to the location by GPS coordinates. After detection, the place of thermal energy leaks, the written report with photos of thermogram locations and recommendations are made [3, 16, 17].

Due to the lack of the identification of technological objects in these methods, a manual inspection with accordance to the route was visualized on the map is done, after which a heat map is created [18].

In the Czech Republic, the system Workswell WIRIS has been developed, which can be fitted with a GPS sensor for storing location information on the position of the drone when producing a record about the detection of hot water leaks [14]. This system is designed to be mounted on a drone (UAV) too.

But there are two problems for the implementation of this method of study of energy losses. First, the above means of the thermal measurements are too expensive. Second, on June 11, 2018, the State Aviation Service of Ukraine banned flights of remotely manned aircraft (drones or UAVs) over roads of state importance, streets of cities and villages, etc. [19]. Similar prohibitions on the use of drones on the territory of settlements exist in other countries – Italy, Austria, Czech Republic, Great Britain, etc. [20, 21]. In USA the FAA rule will not allow operation of a small unmanned aircraft over a moving vehicle because impact of a small unmanned aircraft may distract the driver of a moving vehicle and result in an accident [22].

Thus, the problem under consideration is relevant not only for Ukraine but also for the majority of countries of the world using underground heat mains.

3. Problem Formulation

The purpose of the work is to improve the thermographic method for detection the places of heat energy losses on the urban underground thermal pipelines using the IMCS.

The object of the work is the data receiving and handling process for detection heat loss using a combination of terrain maps and results of heat measurements.

The subject of the work is a system for creating heat maps of energy leak on thermal pipelines according to the data of pyrometric non-contact measurements over the surface of underground heat mains.

The purpose of the work is achieved by solving the following research tasks:

- Analysis of existing algorithms and methods implementing heat map construction on the basis of collected data.
- Analysis of the existing base of devices for research and analysis the current temperature of objects.
- Development of an analytical model for forming a heat map and calculating its statistical indicators.
- Development of a software application for detection of heat energy leaks on thermal pipelines and visualizing the results of measurements on heat maps of the area.

In this article, it is proposed to collect thermal data using an on-duty car equipped the IMCS based on tablet and manual pyrometer with wireless transmission of measured temperatures to achieve the goal.

Unlike the above examples, this measuring device (that costs about $100 USD) has several advantages such as:

- Measuring the temperature of the object with the help of non-contact infrared beam.
- The ability to transmit data through Bluetooth port 4.0 to the computer equipment.
- Laser sight in the form of a circle with a point in the middle, that well visible at a sufficiently large distance (from the on-duty car).
- It is compact, convenient and not harmful for the operator health, unlike the equipment described in [5–9].

The practical value of the work is:

- Detection and localization of the breakthroughs of the heating main for the planning of repairing works, maintenance, etc.
- The possibility of identifying objects which emit heat.

4. Methods and techniques used for the solution of the problem

4.1. Development of the information-measuring computer system for detecting heat leaks

Thus, due to the inability to use drones for the inspection of heating mains in urban areas, it is proposed to analyze the thermal pipe state using a non-contact pyrometer. For example, the Testo 805i model can be used. The pyrometer and the tablet are the part of the IMCS on board the on-duty machine moving the diagnostic team.

Bluetooth infrared thermometer Testo 805i allows transmitting the measurement data to the mobile device (smartphone, tablet, etc.) via Bluetooth port 4.0 at a distance of 15–20 m. It requires iOS 8.3 (or later) / Android 4.3 (or later).

Unfortunately, the Testo 805i doesn't measure automatically, so, to receive data, it is necessary to press the button on the device each time. Doing every 10 seconds, the temperature will be measured every 100 m at a speed of 40 km/h (10 m/s).

The measured temperature is recorded into the database of the mobile device and uploaded to the cloud storage.

The structure of such a database should include separate fields for displaying different types of data in the processed request (dates, ambient temperature, temperature above the surface of the heating main, geolocation data of the break section, etc.) [23].

Then measured data can be plotted on the terrain map as values or as a heat map.

The structural scheme of the improved thermographic method of detecting leak places on heat pipes is shown in Figure 3.

Black color indicates the main blocks, hardware and software components and the primary connections between them, used in currently applied thermographic methods for detecting the breakthroughs of the heating main. The first proposed improvements to existing methods are highlighted in red.

It should be noted that thanks to the proposed modifications, the repair team can operate autonomously, regardless of the diagnostic team, and receives information about the place of the

necessary repairs directly from the cloud storage. Also, the presence of the geo-coordinates of the accident section allows you to accurately position the repair equipment over the place of the heating main breakthrough. Replacing equipment from thermal imaging to pyrometric not only significantly reduces the cost of the solution, but also allows you to obtain digitized data associated with GPS coordinates.

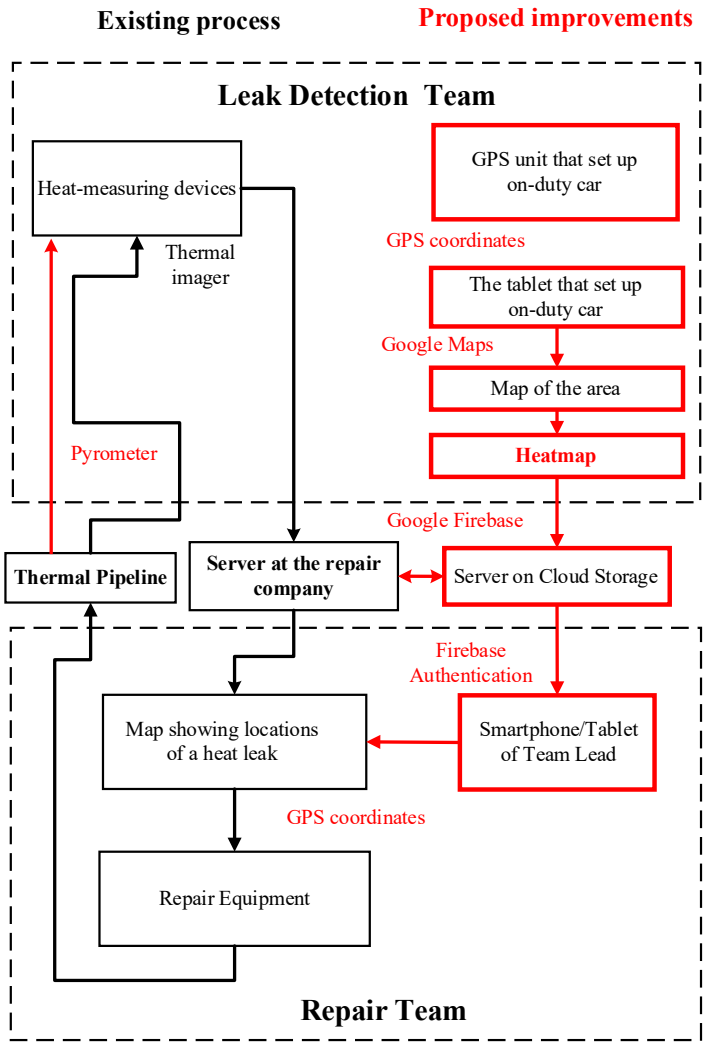


Figure 3. Structural scheme of the improved method for detecting leak places on heat pipes

Figure 4 shows a flowchart of the heat map formation.

Initially, temperature data is measured with a pyrometer, and if these data have been measured, they are recorded in the database, if not – the heat map will not be built, further measurements are crucial.

Further, when the temperature data have been received to the phone (or tablet), there is a process of tracking the coordinates, namely latitude, and longitude, using a GPS module of mobile phone or tablet. If the data is received – it is also entered in the database, and in the case of its absence – the heat map will not be built. The on-duty car equipped with IMCS moves further, to a point with new geolocation, where subsequent measurements would be made.

The final stage is the very construction of a heat map. To do this, you need to add GPS coordinates with measured temperature values into Google Maps [24]. To do this, it is necessary to use the library *com.google.maps.android:android-maps-utils.*

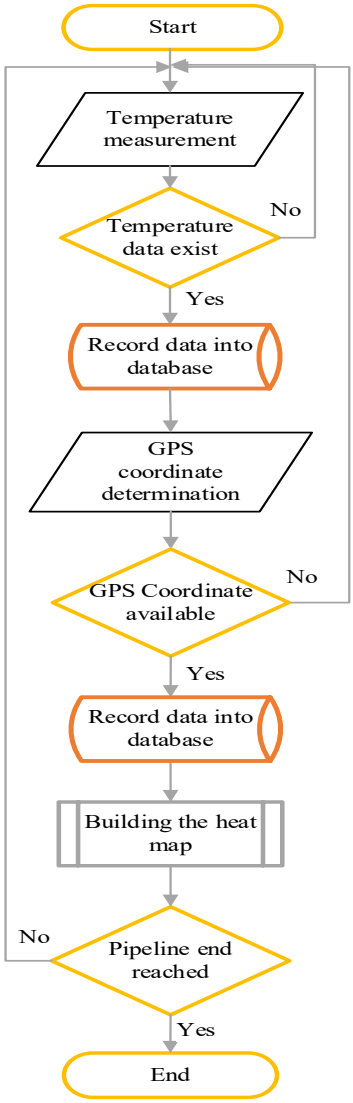


Figure 4. Flowchart of the heatmap building algorithm

The distinctive feature of the proposed method is the application of cloud technologies when obtained data (temperature and GPS coordinates) is uploaded to Google Firebase [25]. In this case, we propose an extended technological process shown in Figure 3.

4.2. Results of measurements and input data for research

During the practical research, measurements of the temperature over the heating main were carried out in preparation for the heating season (in October) on the segment of the thermal pipeline limited by the geocoordinates given in Table 1, along one street.

As a result of the temperature measurements, a breakthrough of the heat pipe was detected. The temperature over the heat main that differs significantly from others evidenced that.

Table 1. Temperature measurements on the thermal pipe in preparation for the heating season (in October)

Temperature (°C)	East longitude (degree)	North latitude (degree)	Conclusion on the state of the thermal pipe
+13.8	32.011009	46.970312	holistic
+14.2	32.012510	46.967368	holistic
+14.1	32.013154	46.965597	holistic
+22.6	32.013060	46.962030	breakthrough
+25.2	32.012781	46.961381	breakthrough
+7.2	32.012688	46.958510	holistic
+7.0	32.011994	46.953942	holistic
+7.1	32.012019	46.949474	holistic

Table 2. Temperature measurements on the thermal pipe during the heating season (in January)

Temperature (°C)	East longitude (degree)	North latitude (degree)	Conclusion on the state of the thermal pipe
+3.6	32.011009	46.970312	holistic
+4.1	32.012510	46.967368	holistic
+4.4	32.013154	46.965597	holistic
+5.2	32.013060	46.962030	holistic
+5.2	32.012781	46.961381	holistic
+3.4	32.012688	46.958510	holistic
+3.0	32.011994	46.953942	holistic
+4.1	32.012019	46.949474	holistic

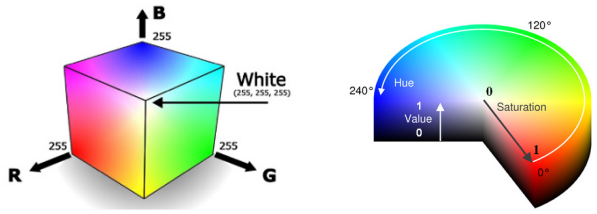
After carrying out repairs, another temperature measurement of the heating main was made at the same points. Measurements were made during the heating season (in January), then the environment temperature was lower (Table 2). Therefore, the average temperature over the surface of the thermal pipelines has decreased compared to the previous measurements made in October.

However, as it can be seen, the temperature along the heating main does not differ significantly from one another, which is why we conclude that the breakthrough of the heating main in this place has been eliminated.

4.3. An analytical model for the formation of a heat map

It should be noted that tabular data is very inconvenient for perception and rapid analysis. Therefore, it was decided to build the heat map with visualizing of found heating main emergency sections on the area map on the basis of the measurements. For this purpose, an analytical model for the formation of a heat map was developed.

Two color models, RGB and HSV, are selected for conversions [26]. The RGB model describes the color space based on Red, Green and Blue colors (Figure 5, a). The HSV model operates with parameters Hue, Saturation and Value (Figure 5, b).



a. b.

Figure 5. RGB (a) and HSV (b) color models

However, if the colors are supposed to be interpreted using the human visual system, it is advisable to integrate the RGB space and the HSV space [27]. Besides, diagnostic and repair work on heating mains can be carried out at different times of the day, and the HSV color space is more resistant to changes in lighting. Therefore, it is advisable to formulate the mathematical equivalent of the projection of the HSV model vectors in the RGB color space.

In contrary to the studies about the usage of color models mentioned above [26–27], in this article, we propose using parameters Hue and Value of the HSV color model to construct the color gradient in the RGB color model. In this case, parameter Saturation equals a constant value. This approach reduces the computing power requirements concerning an IMCS set in an on-duty car. Thanks to that building a heat map for the thermal mains and combining it with a terrain map can be performed even on a tablet instead of a laptop with higher performance, but also more cumbersome and inconvenient to use in a moving car.

The pyrometer readings, corresponding to the seasonal temperature value of the section of the pipeline, are successively added to the array of temperatures $\{t_{i,j}\}$, where $i = \overline{0,N}$ is the number of all N rows of temperature values for each from N seasons, $j = \overline{0,25}$ – the number of columns of each temperature value in the matrix of the input file data to build a heat map.

To recount the parameters of the HSV color model in the value of each spectrum of the RGB color model, the temperature coefficient k of the HSV model should be calculated according to (1). This will allow scaling the temperature values to the color of a certain tone. To switch to the parameters of the color model, it is also necessary to calculate the basic statistical parameters, namely the maximum and minimum values of the temperature (2) – (3) from the set of temperatures (4):

$$k = \frac{(sp_{max} - sp_{min})}{(t_{max} - t_{min})}, \quad (1)$$

$$t_{max} = max\{t_{i,j}\}, \quad (2)$$

$$t_{min} = min\{t_{i,j}\}, \quad (3)$$

$$i = \overline{0,N}; j = \overline{0,25}, \quad (4)$$

where $sp_{max} = 360°$ – selected the maximum value of the color tone (Hue) of the HSV model, which corresponds to the red color; $sp_{min} = 120°$ – selected the minimum value of Hue of the HSV model, which corresponds to the green color. On the basis of the set minimum (green) and maximum (red) Hue values, a range is formed in the middle of which the value of 240° corresponds to the blue color, which in turn is the spectrum of the RGB color model.

Then the color tone parameter $H_{i,j}$ for each heat map segment using the temperature coefficient for the conversion the temperature values to the color values are calculated on the basis of the formula (5):

$$H_{i,j} = sp_{min} + (t_{i,j} - t_{min})\cdot k. \quad (5)$$

In a heat map, to draw a small segment of a certain color of the RGB palette, you need to convert the temperature values using the HSV color parameters. In the proposed analytical model (6) – (9) the saturation " S" is taken as a constant ($S = 1$). This parameter does not appear in further calculations to reduce the computing costs of a mobile device with limited computing capabilities, which is used in the described IMCS for thermal pipelines inspection. $V_{min} = 0.1$ – the minimal digit of the parameter "Value".

Therefore, the next step in forming a thermal image in the proposed analytical model is the conversion of temperature data into a color RGB model based on the rated values of HSV model parameters.

The values of the red, green, and blue components of the RGB spectrum are labeled $R_{i,j}$, $G_{i,j}$, and $B_{i,j}$, respectively and calculated as follows:

$$R_{i,j} = \begin{cases} V_{min}, H_{i,j} = 0 \ or \ H_{i,j} = 1; \\ |V_{min} - ad_{i,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{min} + ad_{i,j}|, H_{i,j} = 4. \end{cases} \quad (6)$$

$$G_{i,j} = \begin{cases} V_{min}, H_{i,j} = 0 \ or \ H_{i,j} = 1; \\ |V_{min} - ad_{i,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{min} + ad_{i,j}|, H_{i,j} = 4. \end{cases} \quad (7)$$

$$B_{i,j} = \begin{cases} V_{min}, H_{i,j} = 0 \ or \ H_{i,j} = 1; \\ |V_{min} - ad_{di,j}|, H_{i,j} = 2; \\ V, H_{i,j} = 3; \\ |V_{min} + ad_{i,j}|, H_{i,j} = 4, \end{cases} \quad (8)$$

$$ad_{i,j} = (V - V_{min})\cdot(H_{i,j}). \quad (9)$$

In this case, the higher the temperature the more the color is close to the red, and vice versa – the colder the temperature the more the color is closer to the blue. As shown in Figure 6, a, in the described model, the temperature scale is implemented in the temperature range from 0 to +25°C, which corresponds to the season when the heat pipelines in the region under consideration are prepared for the heating season.

Figure 6. Temperature scale for the season

It should be noted that thermal measurements can be carried out in different seasons when the temperature of the environment varies considerably over the seasons. When researching at another time of year, the visual temperature scale should be replaced by another scale reflecting temperature values that coincide with the seasonal ones.

When constructing temperature scales, it should be noted that the operating temperature range for the pyrometer is from minus 10°C to +50°C. But with negative values of the ambient temperature, the measurement error is significantly increased. Therefore, further temperature scales were built only for ranges in which the pyrometer measurement error did not exceed 1.5 %.

4.4. The software development

The software was developed in Android Studio 3.3.1, an application created for the Android platform with JDK ver. 11.0.

Let's consider the life cycle of the application. On the initial phase of the application, the task should be sent to diagnostic teams for inspection of the area.

Using the Firebase Cloud Messaging (FCM), the system server sends messages to clients about current tasks. There is an ability to send notifications (up to 2K bytes) and data messages (up to 4K bytes).

Upon receiving the assignment, the diagnostic teams on a duty car are headed to places of information gathering. After their arrival, the system reads its GPS coordinates and sends a message to the server about the diagnostic team location.

The diagnostic team of specialists measures the temperature values using pyrometers. Data via Bluetooth 4.0 technology are transmitted to a tablet/notebook storing a local database in SQLite. When the tablet is connected to the Internet, information is sent to the cloud database.

After completing the inspections by the diagnostic team, other services (mobile clients) will receive a message through the FCM for further works on the territory.

Using the web version of the application helps to create reports based on the data collected. The *JasperReports* library is used to generate reports on the server.

The UML Class Diagram of the developed application is given in "the help file" in the code repository. The shared code that would be good for further researching by other scientists is also placed there.

The following classes were developed:

CurrentApplication – the class responsible for the life cycle of the application. It performs one or another action depending on the state of the application.

PointModel is the class responsible for representing the single point information that is displayed on the map.

PointPackModel – a class serving as a container for the PointModel list.

BaseActivity is the base class (parent) for all Activity entity. This class is responsible for the elements with visual content.

BaseMapActivity is a class with the basic functionality of Google Map.

SplashActivity is the class responsible for displaying the application loading screen.

MainActivity is the class responsible for displaying a map with functional points.

5. Experimental Results and Analysis

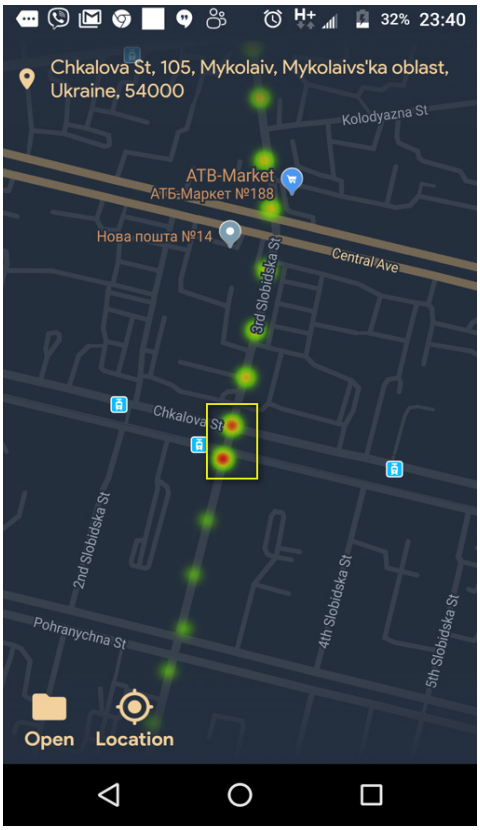
Figure 7 shows screenshots of the interface of the developed Android application.

In Figure 7 you can see the map and two buttons. The button "Location" is responsible for the current location of your mobile device (or tablet). The button "Open" is responsible for opening a file from the database needed to build a heat map.

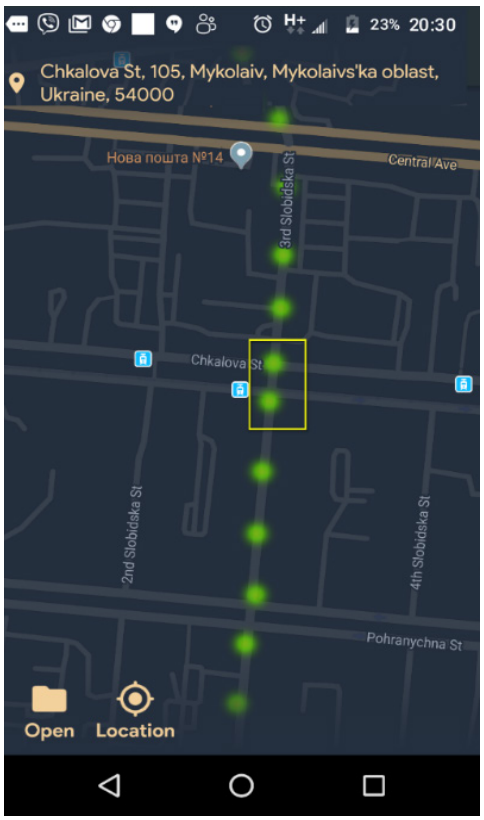
The file was filled via a Bluetooth-port with results of temperature measurements made a pyrometer Testo 805i.

From the analysis of the heat map in Figure 7, the place of heat energy leak from the underground pipeline can be seen (in the area highlighted by a yellow square). It is possible to speak about the presence of heat leaks based on the temperature gradient constructed on the scale in fig. 6. The red color inside the selected two spots (in the color spots above on the map there is no such color at all, or it is represented in an insignificant proportion to the area of the whole spot) indicates the presence of heat leaks in these places of the heating main. In the developed system, accurate geo-coordinates (latitude and longitude) of the section of the break of the underground heating main are registered and available to the repair team (see Table 1). These coordinates can be refined by more frequent pyrometric measurements at a previously diagnosed area and marked on a Google map of a larger scale. Then it will be possible, with higher sampling, to perform image segmentation of the spot of the color scale for sensors of the automated system of technological control [28]. The maximum spot area, indicated in the red inside, will show a more accurate accident section. This will allow the positioning of repair equipment directly above the area of the heating main break. The accuracy of diagnosis is limited only by the measurement interval of the device (3 seconds) and the speed of movement of the operator of the diagnostic team.

Automated fixation of the geo-coordinates of the accident site with their plotting on the terrain map is performed in none of the methods described in the literature.



a.



b.

Figure 7. The heatmap showing the breakdown (a) on the thermal pipeline and its eliminated (b)

With the data of the developed IMCS, it is also possible to conclude that after the detected leakage point along the pipe main, the temperature over the surface of the underground heating pipe has been reduced to the ambient temperature on the date of measurement (green labels in Fig. 7, but without a color gradient according to the developed temperature scale for the studied season).

After the computer diagnostics were carried out, the received data was transferred to the corresponding repair company. The unearthing of the heating pipeline by the found coordinates of the breakthrough has confirmed the fact of the accident precisely in the detected place of the heating main. It was discovered as a result of temperature measurements and reflected in Figure 7, a. The red color in the center of the spot corresponds to the raised temperature. And Figure 7, b shows the same place, but already with the eliminated breakthrough (see the yellow rectangle).

The advantages of the proposed method are that the repair team can operate independently of the work schedule and the location of the diagnostic team specialists. When the repair team arrives at the place of heat energy leakage, the brigadier is authorized on Firebase by any mobile device (smartphone, tablet, etc.). In this case, his profile in any social network can be used – Google, Facebook, Twitter, etc.

When the access to the heat map of the thermal pipeline has given, brigadier of the repair team can compare the own GPS coordinates with the coordinates received from the cloud service. Then it is able to combine the repair equipment accurately (excavator, pump, etc.) with detected places of the heat energy leak from the heat map.

6. Conclusion

Based on the analysis of the methods the detection of the thermal energy leak places on thermal pipelines, the relevance and necessity of developing new analytical models and software for the search and mapping leaks on heating mains with the help of heat maps have been established.

An analytical model for forming a heat map and calculating its statistical parameters was proposed. The convert of the temperature values of a non-contact digital pyrometer from an underground heating main to the image of the RGB color model using the parameters of the HSV color model was substantiated.

A software application for building heat maps based on measured temperature data for the Android platform with JDK 11.0 version was developed. Introduced Firebase Authentication, integrated with popular social services. Due to the heat map layered on Google Maps, the repair team got the opportunity access to detected places of the heat energy leak regardless of the work schedule and the location of the diagnostic team.

Experimental researches with the developed IMCS for the needs of urban housing and communal services for detection places of heat energy leak in the city underground thermal mains due to pipe accident have been carried out. The conducted studies confirmed the sufficient accuracy of the results for the optimal location of repair equipment on detected places of the pipe breakthrough.

The perspective of further development of the thermographic method to detect the places of heat energy loss using the IMCS and cloud services has shown.

As a development of the method, it can be supplemented, besides the technology of processing surface temperatures, also with information technologies of visual information processing on purpose more accurate geo position of heat leaks [29].

In the future, the proposed method for the diagnosis of thermal outflows can be used for both technical and biological objects. For each of the goals, it will be necessary to build separate temperature scales in the appropriate temperature range based on the proposed analytical model. Also, the pyrometric equipment should also be selected with the accuracy of measurement and the price of dividing the scale corresponding to the constructed temperature scale.

The developed system has a lower cost (up to 50 times) compared with similar thermographic systems based on thermal imagers. Also, the proposed system is safe for the operator compared to diagnostic systems based on ultrasound, radiation and electromagnetic non-invasive methods for monitoring heat leaks described in the analyzed literature.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The research was carried out with the support of the Ministry of Education and Science of Ukraine within the framework of the state-funded scientific research works at the Petro Mohyla Black Sea National University.

References

[1] Non-destructive Testing and Repair of Pipelines. E. N. Barkanov, A. Dumitrescu, I. A. Parinov (eds), Springer, Cham, 2018. doi: 10.1007/978-3-319-56579-8.

[2] M. Fahmy, O. Moselhi, “Detecting and locating leaks in Underground Water Mains Using Thermography” in 26th International Symposium on Automation and Robotics in Construction (ISARC 2009), 61–67, 2009. https://www.irbnet.de/daten/iconda/CIB14819.pdf.

[3] “Central heating leaks” [L.D.S Leak Detection Specialists Ltd.]. https://www.leakdetectionspecialists.co.uk/central-heating-leaks.htm.

[4] “Leak Detection in Heat Exchangers and Underground Pipelines Using Radiotracers” [Material for education and on-the-job training for practitioners of radiotracer technology], Vienna, International Atomic Energy Agency, 2009. https://www-pub.iaea.org/MTCD/Publications/PDF/TCS-38_web.pdf.

[5] S. Kim, “Holistic diagnosis of pipeline system with impedance method” PROCEDIA ENGINEER, 70, 924–933, 2014. doi: 10.1016/j.proeng.2014.02.103.

[6] J. Liua, H. Sua, Ya. Maa, G. Wanga, Yu. Wangb, K. Zhang, “Chaos characteristics and least squares support vector machines based online pipeline small leakages detection” CHAOS SOLITON FRACT, 91, 656–669, Oct. 2016. doi: 10.1016/j.chaos.2016.09.002.

[7] M. Valincius, M. Vaisnoras, A. Kaliatka, “Study and demonstration of pressure wave-based leak detection in a district heating network” STRUCT INFRASTRUCT E, 14 (2), 151–162, 2018. doi: 10.1080/15732479.2017.1330892.

[8] X. H. Tang, Y. B. Liu, L. J. Zheng, C. B. Ma, H. Wang, “Leak Detection of Water Pipeline Using Wavelet Transform Method” in International Conference on Environmental Science and Information Application Technology (ESIAT 2009), Wuhan, PEOPLES R CHINA, 2009, 2, 217–220. doi: 10.1109/ESIAT.2009.57.

[9] “Infrared Inspection Services – Thermography and Thermal Imaging in New York, NYC, Long Island and New Jersey for Commercial Industry and

Property Management" [web-site of Infrared Imaging Services LLC]. http://www.infraredimagingservices.com/commercial.
[10] M. P. Musiyenko, I. M. Zhuravska, I. V. Kulakovska, A. V. Kulakovska, "Simulation the behavior of robot sub-swarm in spatial corridors" in 2016 IEEE 36th International Conference on Electronics and Nanotechnology (ELNANO–2016), Kyiv, Ukraine, 2016, 382–387. doi: 10.1109/ELNANO.2016.7493090.
[11] O. Friman, P. Follo, J. Ahlberg, S. Sjokvist, "Methods for large-scale monitoring of district heating systems using airborne thermography" IEEE T GEOSCI REMOTE, 52 (8), 5175–5182, Aug. 2014. doi: 10.1109/TGRS.2013.2287238.
[12] A. Hawari, M. Khader, W. Hirzallah, T. Zayed, O. Moselhi, "Integrated sensing technologies for detection and location of leaks in water distribution networks" WATER SCI TECH-W SUP, 17 (6), 1589–1601, Dec. 2017. doi: 10.2166/ws.2017.044.
[13] G. R. Stockton, "Leak Detection – Aerial Thermal Mapping Reveals Leaks in All Types of Systems" in CampusEnergy2017 [International District Energy Association]. Miami, FL, Feb. 20–24, 2017. https://www.districtenergy.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=5ae4b471-ee47-9f2f-4443-94a8fe505273.
[14] "Inspection of hot-water and warm-water pipes" [web-site of the Workswell s.r.o.]. https://www.workswell-thermal-camera.com/inspection-hot-water-warm-water-pipes/.
[15] I. Zhuravska, I. Kulakovska, M. Musiyenko "Development of a method for determining the area of operation of unmanned vehicles formation by using the graph theory" Eastern-European Journal of Enterprise Technologies, 2, 3 (92), 4–12, 2018. doi: 10.15587/1729-4061.2018.128745.
[16] "Thermal Mapping – Inspection of District Heating" in Case Study for Drone Systems – with Thermal Capture Recorder 2.0 640, Copenhagen, Denmark. https://thermalcapture.com/thermal-mapping-inspection-of-district-heating/.
[17] P. M. Bach, J. K. Kodikara, "Reliability of infrared thermography in detecting leaks in buried water reticulation pipes" IEEE J SEL TOP APPL, 10 (9), 4210–4224, 2017. doi: 10.1109/JSTARS.2017.2708817.
[18] E. Cornelis, F. Meinke-Hubeny, "The STRATEGO Project No: IEE/13/650 (Multi-level actions for enhanced Heating & Cooling plans)", European Union, May 29 2015. http://stratego-project.eu/wp-content/uploads/2014/09/D3_7a-STRATEGO-WP3-Mapping-methodology-and-data-sources.pdf.
[19] "Rules for Use of Airspace of Ukraine", State Aviation Administration of Ukraine, 430/210, May 11 2018. https://avia.gov.ua/wp-content/uploads/2018/05/TYMCHASOVYJ-poryadok-VPPU_31.05.18-2.doc (In Ukrainian).
[20] "Regulation of the European Parliament and of the Council on common rules in the field of civil aviation", Brussels, July 12 015. http://ec.europa.eu/smart-regulation/impact/ia_carried_out/docs/ia_2015/swd_2015_0262_en.pdf.
[21] "Airspace restrictions for unmanned aircraft and drones" [Civil Aviation Authority – CAA in the United Kingdom]. https://www.caa.co.uk/Consumers/Unmanned-aircraft/Our-role/Airspace-restrictions-for-unmanned-aircraft-and-drones/.
[22] "Unmanned Aircraft Systems (UAS)" [U.S. Department of Transportation – Federal Aviation Administration, FAA]. https://www.faa.gov/uas/.
[23] M. Fisun, M. Dvoretskyi, A. Shved, Y. Davydenko, "Query parsing in order to optimize distributed DB structure" in 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS 2017), Bucharest, Romania, 2017, 1, 172–178. doi: 10.1109/IDAACS.2017.8095071.
[24] E. Price, "Create a heat map of your Google location history with this tool", Apr. 11, 2018. https://lifehacker.com/create-a-heat-map-of-your-google-location-history-with-1830215410.
[25] K. S. Ashok, Mastering Firebase for Android Development: Build real-time, scalable, and cloud-enabled Android apps with Firebase, Packt Publ., 2018.
[26] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, J. Jatakia, "Human skin detection using RGB, HSV and YCbCr color models" [ICCASP/ICMMD-2016] ADV INTEL SYS RES, Atlantis Press, 137, 324–332, 2017.
[27] M. Loesdau, S. Chabrier, A. Gabillon, "Hue and saturation in the RGB color space" [Lecture Notes in Computer Science], 8509, 203–212, Springer, Cham, 2014. doi: 10.1007/978-3-319-07998-1.
[28] M. Fisun, W. Smith, A. Trunov, "The vector rotor as instrument of image segmentation for sensors of automated system of technological control" in the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT 2017), Lviv, Ukraine, 2017, 1, 458–463. doi: 10.1109/STC-CSIT.2017.8098828.
[29] G. Shcherbakova, S. Antoshchuk, A. Sachenko, M. Gerganov, M. Polyakova, V. Krylov, " " in the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing (DSMP 2018), Lviv, Ukraine, 2018, 1, 605–608. doi: 10.1109/DSMP.2018.8478551.

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3,10-15 (2019)

www.astesj.com

Special Issue on Advanced Electrical and Communication Technologies

ASTESJ

ISSN: 2415-6698

Microstrip Patch Antenna for Ultra-Wideband Applications

QADDI Mohamed Hamza*[1], SRIFI Mohamed Nabil[1] and MHARZI Hassan[2]

[1]*Telecommunication Systems Research Group, National School of Applied Sciences, Ibn Tofail University Kenitra, Morocco*

[2]*Laboratory of Electric Engineering, Computing and Mathematical Sciences National School of Applied Sciences, Ibn Tofail University Kenitra, Morocco*

ARTICLE INFO

Article history:
Received: 31 January, 2019
Accepted: 17 April, 2019
Online: 04 May, 2019

Keywords:
Ultra-wideband
Micro-strip antenna
Directivity
Gain
Slots

ABSTRACT

A new microstrip patch antenna for ultra-wideband applications is presented in this paper. The design and performance of the antenna component are discussed. The propounded antenna is mounted on a compact FR-4 substrate having dimensions 20 x 30 x 1.6 mm^3 with relative permittivity εr=4.3. The rectangular patch antenna is slotted with two types; rectangular and semi-circle slots so as to ensure a broad bandwidth. The results reveal that the antenna covers the frequency range of 3.1–7.5 GHz with a reflection coefficient reduced to -55 dB and a maximum gain of 5.9 dB. The details of the simulated and measured results for reflection coefficient are presented, showing a good agreement between them. To analyze the effect of the slots, the surface current distribution is investigated. The performances of good impedance matching are achieved within the operating band. Simulations are performed using CST Microwave Studio. The propounded antenna can be deployed for UWB applications and other radio communication services systems such as high-resolution radar, military communication, communications and sensors, position location and tracking.

1. Introduction

UWB technology has recently attracted significant attention in wireless communications thanks to its main features in particular high bandwidth and data rate, ensuring secure communication in military applications. The world of ultra-wideband (UWB) has changed in very recent history. In February 2002, the FCC (Federal communications commission) issued a ruling that UWB could be used for commercial applications on unlicensed basis as well as for radar and safety applications. The decision authorized very low power spectral density emission in a bandwidth ranging from 3.1 to 10.6 GHz [1-3]. UWB appears to be a technology very promising for wireless communications at very high levels data-rate, high-precision radar and imaging systems. UWB systems use short pulses (of the order of the picosecond), repeated at a certain rate that can be up to several giga-impulses per second, offering thus a very high bandwidth with a very low transmission power level. This offers UWB systems the possibility of coexisting with other systems.

*Corresponding Author: QADDI Mohamed Hamza, 27 Bir Rami EST, Kenitra, Morocco, Cell No; +212 661634417, qaddihamza@gmail.com

To fulfill the demand in this type of communications, many research on UWB antenna design have been published. Most of them focused on patch antennas because of their simple structure, small size, low cost, low complexity, light weight, and high-speed data rate. They are considered as good candidates for UWB applications among many versions of UWB antennas. Nevertheless, micro-strip patch antenna has narrow bandwidth.

To resolve this problem and obtain the desired performances, many methods are used. The typical technique focus on cutting slots of several forms in the radiating element, the ground plane, and the feed line. Nowadays, many slot shapes are available and discussed by various researchers [4-6]. Slot antennas have proven to be useful in that context. However design and placement of these slots on the radiating element is challenging as compared to the other traditional techniques. The purpose of this project is to design a micro-strip patch antenna for UWB applications that covers the frequency range of 3.1–7.5 GHz.

2. Design and Configuration of the Antenna

The propounded antenna is a rectangular patch antenna which is fed by a micro-strip line and mounted on a dielectric substrate named FR4-lossy, with thickness h=1.6mm and relative

permittivity εr=4.3, having dimensions of 20 x 30 x 1.6 mm^3. These specifications for the substrate are a part of material parameters window on CST Microwave Studio. The material used in the patch and the ground is lossy copper with a thickness of Mt=0.035mm as given in Tab .1. The width of the line feeding is 3 mm and the length is 14.5 mm. The partial ground plane is also used with the propounded antenna geometry as shown in Figure 1. To investigate the impact of various antenna's characteristics (different slots), a meticulous parametric study was performed. The parametric values of the propounded antenna are shown in Tab.1. The antenna is excited by a 50-Ω micro-strip feed line. The parameters measurement of the patch are calculated using the following formulas [7]:

Width formula of rectangular:

$$W_p = \left(\frac{c}{2 f_r}\right)\left(\frac{\varepsilon_r + 1}{2}\right)^{-0.5}$$

Where c = 3×108 ms-1 and $\varepsilon_r = 4.3$

Formula of effective dielectric constant:

$$\varepsilon_{eff} = \left(\frac{\varepsilon_r + 1}{2}\right) + \left(\frac{\varepsilon_r - 1}{2}\right)\left(1 + \frac{12\ h}{W_p}\right)^{-0.5}$$

At h=1.6mm

Formula of length extension:

$$\Delta L = 0.412\ h\left(\frac{\varepsilon_{eff} + .3}{\varepsilon_{eff} - 0.258}\right)\left(\frac{\left(\frac{W_p}{h}\right) + 0.264}{\left(\frac{W_p}{h}\right) + 0.8}\right)$$

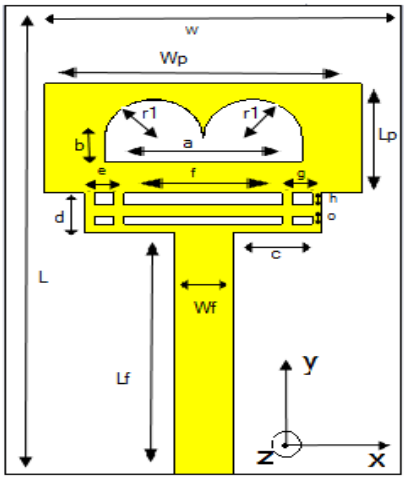
Length formula of rectangular:

$$L_p = \left(\frac{c}{2 f_r \sqrt{\varepsilon_{eff}}}\right) - 2\Delta L$$

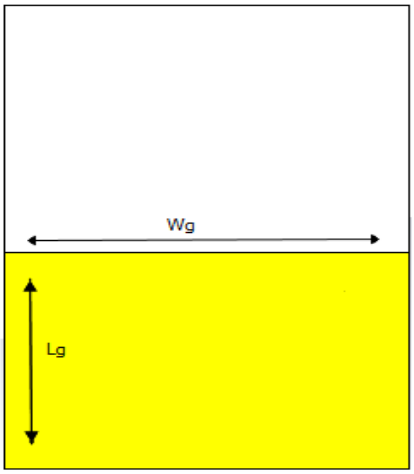
Table 1: Parameter values of the propounded antenna

Parameters	Values (mm)
L	30
W	20
Lp	7
Wp	16
Lg	14
Wg	20
h	1.6
Mt	0.035
Wf	3
Lf	14.5
a	10
b	1.5
c	4.5
d	2.5
e	1
f	8
g	1
h	0.75
o	0.5
r1	2.5

To amplify the coupling between the feed line and the slots, multiple slots are introduced in order to broaden the operating bandwidth of the antenna. Having a quite large slot is logical transition to obtain wide bandwidth from an aperture-coupled patch. The first two slots are designed by combining rectangular and semi-circle slots placed in a juxtapose way [8-10]. In addition, different other rectangular slots are included in the radiating element. The structure of the antenna design is illustrated in Fig .1.



(a)



(b)

Figure 1 Design of the propounded UWB antenna (a) Top (b) Bottom

To improve the performance of micro-strip patch antenna, it is advisable to design a combination of two or more slots on the radiating element. This can be observed from the discussed literature survey that a micro-strip patch antenna with slots can be designed for wireless applications [11]. The technique employed in this paper retains the same elementary structure of the antenna with different slots included in the radiating element. In order to obtain a wide bandwidth, a rigorous investigation on different parameters in terms of feed width (Wf), feed length (Lf) and different slots width and length was performed. From the analysis we found the optimum dimensions of above parameters to be 3 mm, 14.5 mm respectively for the feed line, to meet the

requirement. The slots dimensions and other antenna's parameters are listed in Tab .1 on the basis of undertaken parametric analysis.

Figure 2 presents the simulated reflection coefficient of the antenna, the antenna without rectangular slots and the antenna without semi-circle slots. When different slots are included, a wide bandwidth is achieved from 3.1 to 7.5 GHz with three resonant frequencies at 3.9, 5.95 and 7.3 GHz. Effect of using different slots is shown. We notice that the propounded antenna is the best candidate to meet the requirements in terms of UWB bandwidth. The flow chart of the propounded antenna's design methodology is presented in figure 3.

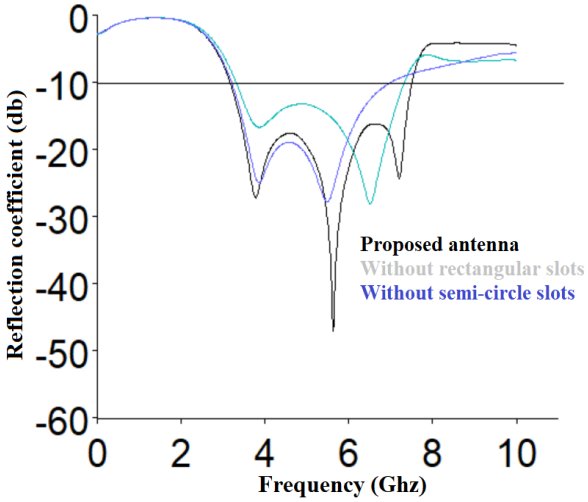


Figure 2 Simulated reflection coefficient of the propounded antenna, the antenna without rectangular slot and the antenna without semi-circle slots

3. Results and Discussions

The propounded UWB antenna is designed by opting for the optimal values of above parameters in order to cover the frequency range 3.1- 7.5 GHz. Both numerical and experimental tests have been achieved to assess the effectiveness of the propounded antenna and the results are presented in this section. Toward this end, the antenna was fabricated on FR4-lossy substrate (30 x 20 x 1.6 mm^3) with dielectric constant of Ɛr = 4.3. Figure 4 depicts the fabricated antenna for experimental verification.

The reflection coefficient of the antenna has been analyzed using CST Microwave STUDIO simulator and verified using another electromagnetic simulator HFSS as shown in Figure 5. A 3.1–7.5 GHz frequency range below -10 dB of return loss S11 is obtained. The reflection coefficient (S11), measured in decibel (dB), can be calculated using the formula:

$$S_{11} = -20log_{10}\left|\frac{Z_{in} - Z_0}{Z_{in} + Z_0}\right|$$

Where Z0 is the characteristic impedance of the 50 Ω SMA port and Zin is the driving point impedance of the antenna. Figure 5 shows a comparison of the simulated reflection coefficient using both CST and HFSS. The results corroborate that there is a good agreement between the two simulations with a small variation as the two softwares use different numerical techniques, finite integration technique, a relative of FDTD for CST Microwave Studio and Finite Element Method (FEM) for HFSS [12].

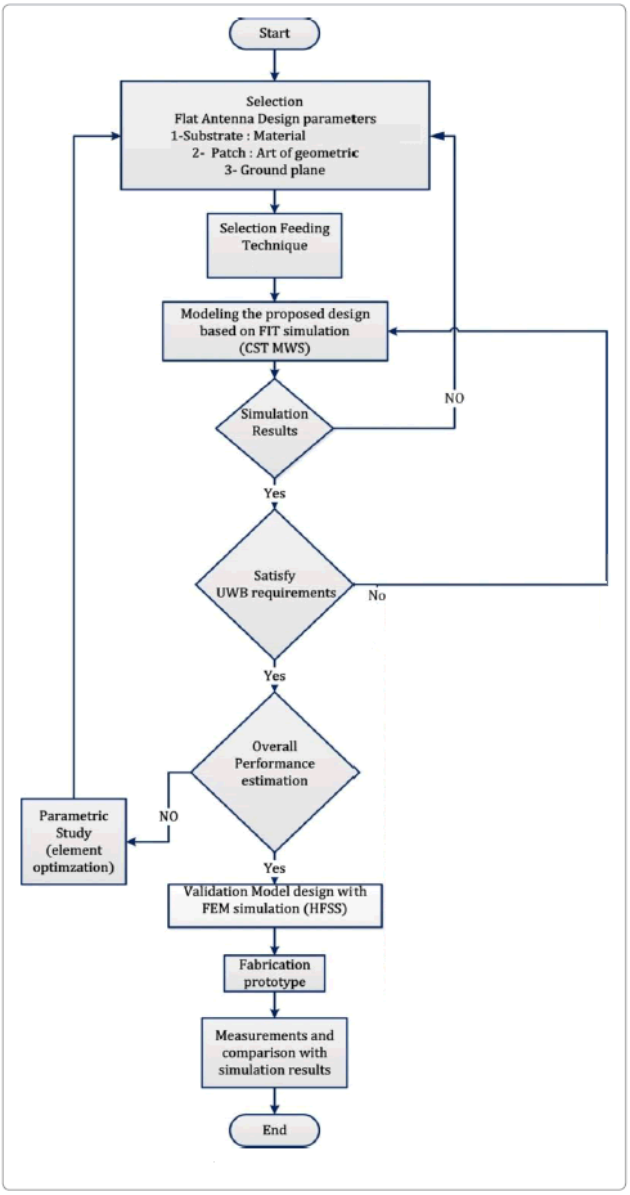


Figure 3 The flow chart of the propounded antenna's design methodology.

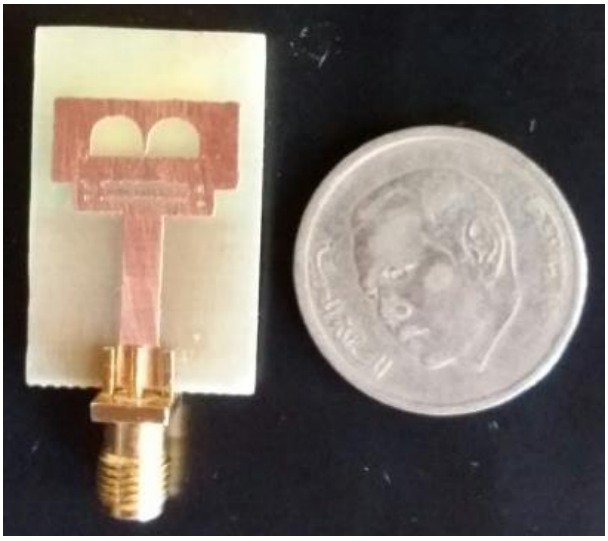


Figure 4 Photograph of the fabricated UWB antenna

The UWB antenna was measured after fabrication to examine the performance of the proposed approach. Measured reflection coefficient is achieved by vector network analyzer. The measured and simulated reflection coefficients of antenna are compared in Figure 6. From measured results, we can notice that the antenna is operating with a -10dB bandwidth from 3.1 to 7.5 GHz. A good agreement between simulation and measurement results is observed with a little discrepancy which is mainly owing to the fabrication margin. It could also be because to the impact of the feeding cable, feeding connector and the antenna fixation support as the structure is small.

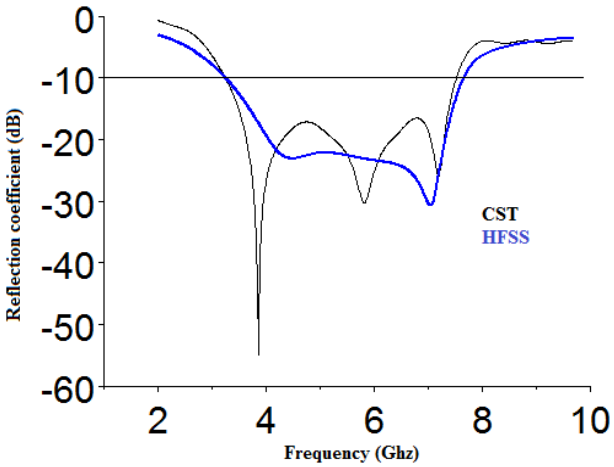


Figure 5 Simulated reflection coefficient of the propounded antenna

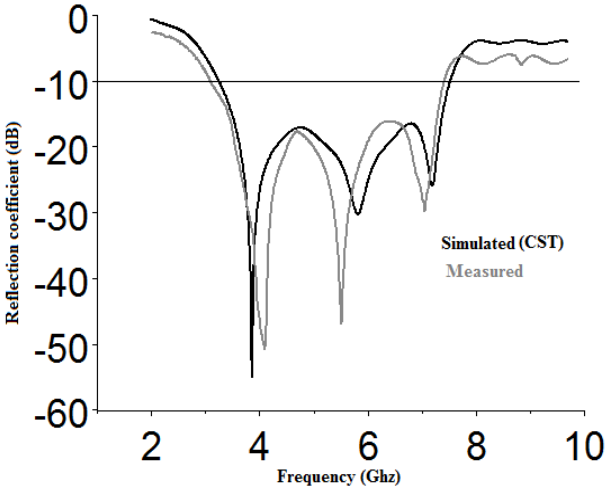


Figure 6 Measured and simulated reflection coefficient of the UWB antenna

The electromagnetic solver, CST Microwave STUDIO, is used to numerically analyze and improve the antenna's configuration. Figure 7 presents the simulated results of the propounded antenna with the rectangular slot combined with semi-circle slots width, from 0.5 to 1.5 mm. It should be noticed that one variable at the time was varied, the others being constant. It is apparent that the bandwidth for UWB band decreases as the width decreases from 1.5 to 0.5 mm. Consequently, it is concluded that b= 1.5 mm is the optimal value for the bandwidth from 3.1 to 7.5 GHz, covering the whole UWB range with three resonant frequencies at 3.9, 5.95 and 7.3 GHz.

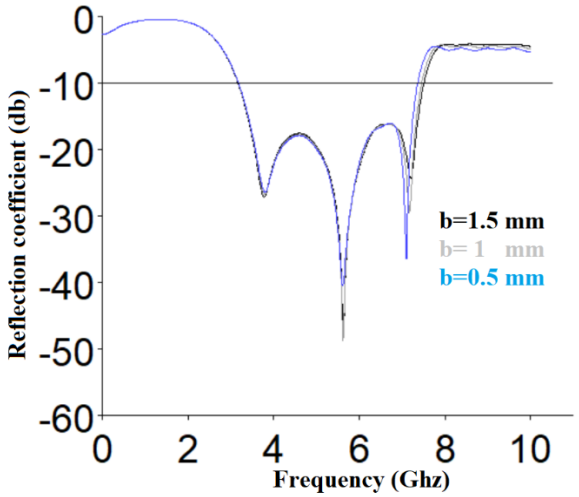
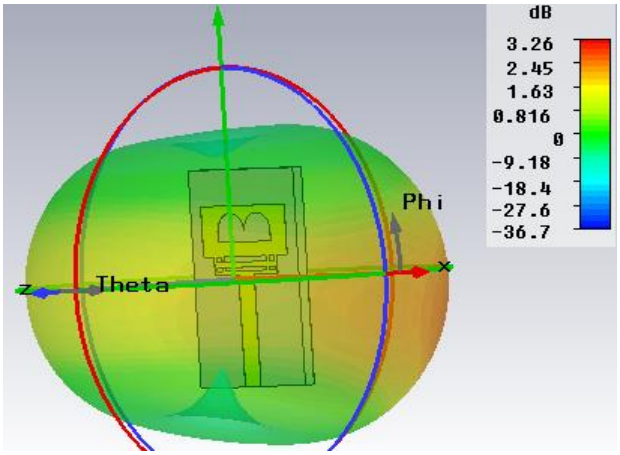
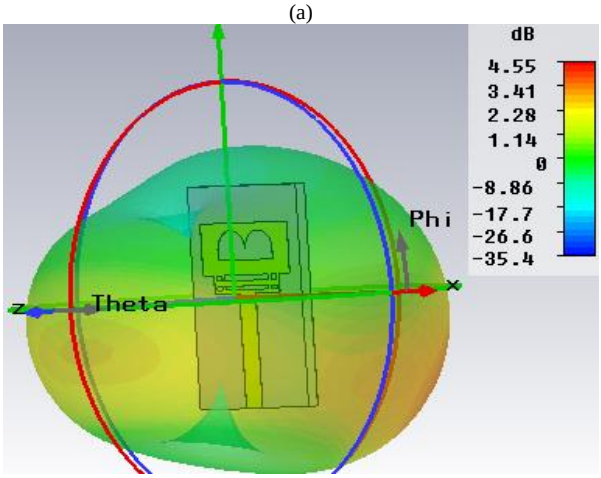


Figure 7 Simulated reflection coefficient response of the antenna as a function of "b". All other parameters are the same as listed in Table 1

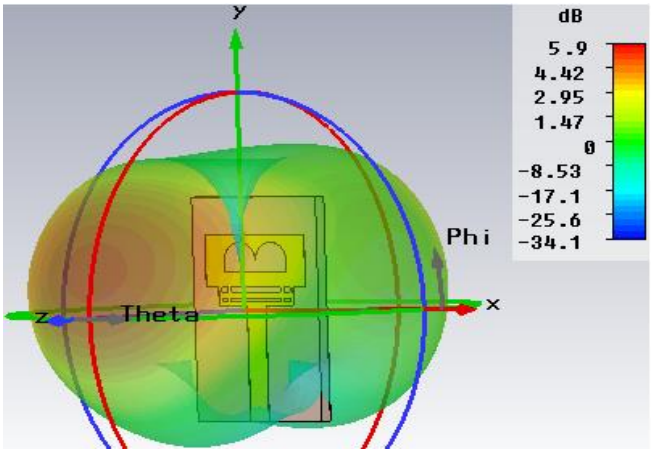
The propounded UWB antenna has an acceptable quasi omnidirectional radiation pattern needed to receive information signals from all directions. Figure 8 presents three-dimensional radiation patterns at three frequencies 3.9, 5.95 and 7.3 GHz. The gain of the antenna attain 3.26 dB at 3.9 GHz, 4.55 dB at 5.95 GHz and 5.90 dB at 7.3 GHz.



(a)

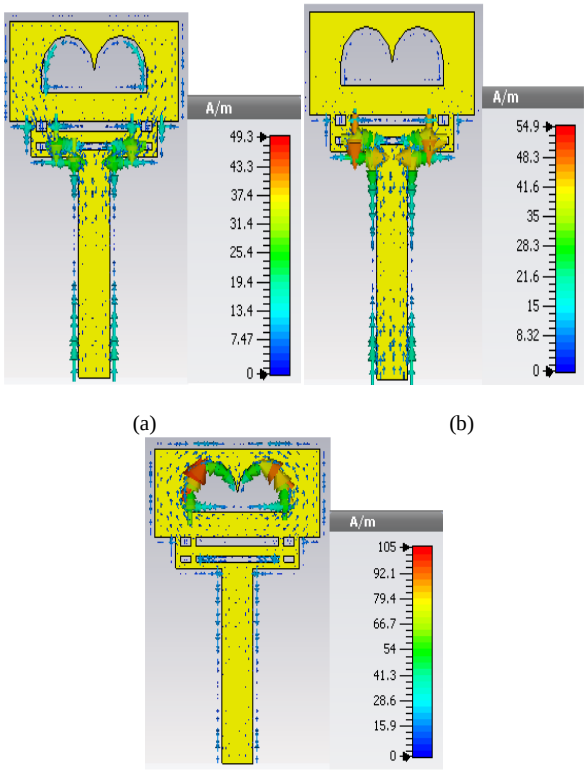


(b)



(c)

Figure 8. 3D radiation patterns of the propounded antenna at (a) 3.9, (b) 5.95 and (c) 7.3 GHz.



(a) (b)

(c)

Figure 9 Surface Current distributions at (a) 3.9 GHz, (b) 5.95 GHz and (c) 7.3 GHz

Surface current distribution has been depicted in Figure 9 and analyzed at discrete frequencies (a) 3.9 GHz, (b) 5.95 GHz and (c) 7.3 GHz, to give a physical perception and have a deeper understanding on the resonant behavior of the antenna. Figure 9 (a) and (b) shows that the surface current distributions are mainly flow and distributed though along the rectangular slots and the feed part. From Figure 9 (c), it can be observed that most of surface currents are concentrated on the edges of the interior and exterior of semi-circle slots. Effective coupling between different slots affords wide-band matched impedance bandwidth, therefore, the surface current is distributed evenly over the radiating patch. The maximum gain of the propounded antenna is presented in Figure 10. As expected, the antenna has a good gain over the whole operating frequency range with a maximum gain up to 5.9 dB.

The simulated radiation efficiency of the antenna is presented in Figure 11. It can be noticed that the antenna reaches a maximum radiation efficiency of 81% and more than 60% over the whole UWB frequency range.

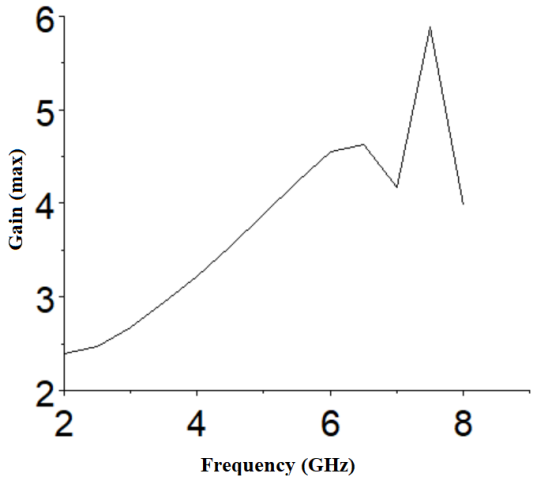


Figure 10 The gain of the propounded antenna

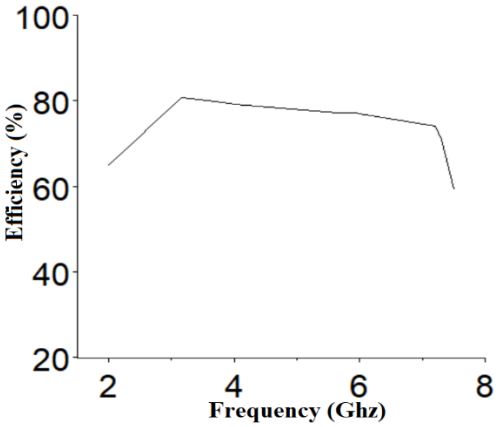


Figure 11 simulated radiation efficiency of the antenna

A comparison between our antenna and other published research is demonstrated in Table 2 in terms of their dimensions, bandwidth, gain and applications, in order to reinforce the concept of design. The almost stable radiation pattern with gain up to 5.9 dB makes our antenna suitable for being used for UWB applications.

Table 2 A comparison between the propounded antenna and other published work

References	Dimensions (in mm)	Bandwidth UWB	Gain at resonant frequency (dB)	Applica-tions
[13]	13x15	3.07-11.64	5.49	UWB applications
[14]	22x24	2.8-11.4	3.6 5.4	UWB applications
[15]	27x32.42	3.1-5.2 5.8-10.6	2.33 5.49	UWB applications
Our work	20x30	3.1- 7.5	3.2 4.5 5.9	UWB applications

4. Conclusion

The micro-strip patch antenna is easy to manufacture, replaceable, low profile and highly efficient. Such antennas are strongly recommended in satellite and wireless communication. In this paper, a new UWB patch antenna is presented with a wide bandwidth from 3.1 to 7.5 GHz with three resonant frequencies at 3.9, 5.95 and 7.3 GHz. The optimized structural parameters have been reached after many optimization and parameter sweeps on antenna performance. The propounded antenna can be used in multiple UWB applications that requires a wide bandwidth and reduced return loss at the operating frequency of the frequency range. The simulated results provide that, the maximum bandwidth attained numerically is 4.4 GHz due to the multiple slots used in the antenna. The results have revealed that our new fabricated antenna is definitely highly convenient for UWB applications such as high-resolution radar, military communication, communications and sensors, position location and tracking.

References

[1] Federal Communications Commission (FCC) "New Public Safety Applications and Broadband Internet Access among Uses Envisioned by FCC Authorization of Ultra-Wideband Technology", February 2002.

[2] Federal Communications Commission (FCC), "Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems", First Report & Order, Washington DC, February 2002.

[3] Chen, D. and C. H. Cheng, "A novel compact ultra-wideband (UWB) wide slot antenna with via holes," Progress In Electromagnetics Research, Vol. 94, pp. 343–349, 2009. doi:10.2528/PIER09062306

[4] A. Ibrahim, M. A. Abdalla, and A. Boutejdar, "A Printed Compact Band-Notched Antenna Using Octagonal Radiating Patch and Meander Slot Technique for UWB Applications", Progress In Electromagnetics Research M, Vol. 54, pp. 153–162, 2017. doi:10.2528/PIERM16122805

[5] Changzhou Hua, Yunlong Lu, and Taijun Liu, "UWB Heart-Shaped Planar Monopole Antenna with a Reconfigurable Notched Band", Progress In Electromagnetics Research Letters, Vol. 65, pp. 123–130, 2017. doi:10.2528/PIERL16120203

[6] A. Pirooj, M. Naser Moghadasi, F. B. Zarrabi, and A Sharifi, "A Dual Band Slot Antenna for Wireless Applications with Circular Polarization", Progress In Electromagnetics Research C, Vol. 71, pp. 69–77, 2017. doi:10.2528/PIERC16111401

[7] C. A. Balanis. Antenna Theory analysis and design. New York, 2007

[8] Kharakhili, F. G., M. Fardis, G. Dadashzadeh, A. Ahmadi, and N. Hojjat, "Circular slot with a novel circular microstrip open ended microstrip feed for UWB applications," Progress in Electromagnetics Research, Vol 68, pp. 161–167, 2007. doi:10.2528/PIER06071901

[9] Marqués, R.; Mesa, F.; Martel J.; Medina, F. "Comparative analysis of edge and broadside coupled split ring resonators for metamaterial design: Theory and experiment". IEEE Transactions on Antennas and Propagation, Vol. 51, No. 10, pp. 2572-258, October 2003. DOI: 10.1109/TAP.2003.817562

[10] I. Latif, L. Shafai and S. K. Sharma, "Bandwidth Enhancement and Size Reduction of Microstrip Slot Antennas". IEEE Transactions on Antennas and Propagation, Vol. 53, No. 3, pp. 994-1003, March 2005. DOI: 10.1109/TAP.2004.842674

[11] M. Hamza QADDI, Marko SONKKI, Sami MYLLYMAKI, Hassan MHARZI, M. Nabil SRIFI, and Heli JANTUNEN, "Novel compact patch antenna for Ultra-wideband (UWB) applications", International Journal Of Microwave And Optical Technology, Vol. 13, pp. 343-350, July 2018.

[12] M. Särestöniemi, T. Tuovinen, M. Hämäläinen, K. Y. Yazdandoost, E. Kaivanto, and J. Iinatti, ''Applicability of Finite Integration Technique for the modelling of UWB channel characterization,'' in Proc. ISMICT, San Diego, CA, USA, 2012, pp. 1–4. DOI: 10.1109/ISMICT.2012.6203045

[13] M. Z. Mahmud et al "A triangular coupled-resonator antenna for ultra-wide applications" Applied physics A, 2017. https://doi.org/10.1007/s00339-016-0639-x

[14] Rezaul Azim, Mohammad Tariqul Islam, and Norbahiah Misran. "Compact Tapered-Shape Slot Antenna for UWB Applications". IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, Vol. 10, pp.1190-1193, October 2011. DOI: 10.1109/LAWP.2011.2172181

[15] Arun Kumar. "A microstrip UWB antenna for next generation communication system". Int. J. Wireless and Mobile Computing, Vol. 15, No. 3, pp.270–278, 2018. doi.org/10.1504/IJWMC.2018.096010

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3,16-24 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Vowel Classification Based on Waveform Shapes

Hakan Tora[*,1], Gursel Karacor[2], Baran Uslu[3]

[1]*Atilim University, Department of Aircraft Electrics and Electronics 06832, Turkiye*

[2]*Grodan (Rockwool B.V.) Industrieweg 15, 6045 JG Roermond, Netherlands*

[3]*Atilim University, Department of Electrical and Electronics Engineering 06832, Turkiye*

A R T I C L E I N F O

Article history:
Received: 06 March, 2019
Accepted: 16 April, 2019
Online: 07 May, 2019

Keywords :
Vowel recognition
Speech waveform features
Image processing
ANN
SVM
XGBoost

A B S T R A C T

Vowel classification is an essential part of speech recognition. In classical studies, this problem is mostly handled by using spectral domain features. In this study, a novel approach is proposed for vowel classification based on the visual features of speech waveforms. In sound vocalizing, the position of certain organs of the human vocal system such as tongue, lips and jaw is very effective on the waveform shapes of the produced sound. The motivation to employ visual features instead of classical frequency domain features is its potential usage in specific applications like language education. Even though this study is confined to Turkish vowels, the developed method can be applied to other languages as well since the shapes of the vowels show similar patterns. Turkish vowels are grouped into five categories. For each vowel group, a time domain speech waveform with an interval of two pitch periods is handled as an image. A series of morphological operations is performed on this speech waveform image to obtain the geometric characteristics representing the shape of each class. The extracted visual features are then fed into three different classifiers. The classification performances of these features are compared with classical methods. It is observed that the proposed visual features achieve promising classification rates.

1. Introduction

Vowel classification has been an attractive research field with growing intensity over the recent years. It is closely related to voice activity detection, speech recognition, and speaker identification. Vowels are the main parts of speech and the basic building units of all languages and an intelligible speech would not be possible without them. They are the high energy parts of speech and also show almost periodic patterns. Therefore, they can be easily identified by time characteristics of their speech waveforms. Each vowel is produced as a result of vocal cord vibrations. The frequency of these vibrations is known as pitch frequency, which is a characteristic feature of the speech and the speaker. Pitch frequency variations occur mainly at voiced parts which are mostly formed by vowels. Consequently, vowels are an important source for features in speech processing.

Detecting the locations of the vowels in an utterance is critical in speech recognition because their order, representing the syllable form of the word, can help in determining the possible candidate words in speech. In addition, voice activity detection can be accomplished by determining the voiced parts of the speech which are mainly constituted from vowels. Speech processing technologies using spectral methods are also dependent on vowels and other voiced parts in speech. These methods are mostly built on the magnitude spectrum representation, which displays peaks and troughs along the frequency axis. Voiced segments of speech cause such peaks in the magnitude spectrum. The frequencies corresponding to the peaks, known as formants, are useful for both classifying the speech signal and identifying the speaker. Therefore, vowels are inevitable in the area of speech processing [1].

There are quite a number of studies on vowel classification in the literature. Most of them are based on frequency domain analysis using features such as formant frequencies [2,3], linear predictive coding coefficients (LPCC), perceptual linear prediction (PLP) coefficients [4], mel frequency cepstral coefficients (MFCC) [5,6,7,8], wavelets [9], spectro-temporal features [10], and spectral decomposition [11]. However, there are fewer studies using time domain analysis [12, 13]. There are also vowel classification studies for the imagined speech [14].

[*]Hakan Tora, Email: hakan.tora@atilim.edu.tr

Although most of the studies on speech recognition make use of the acoustic features, the visual characteristics obtained from speech waveform shapes can also carry meaningful information to represent the speech. Shape characteristics, for example, envelope of the waveform, area under that waveform, and some other geometrical measurements can be utilized for classification purposes. Extracting these properties can be accomplished by basic image processing techniques such as edge detection and morphological processing. In other words, a speech waveform can be treated as an image. The notion of visual features is perceived as the shapes of the mouth and lips in general, and used also in vowel classification [15]. There exist some articles in the literature concerning the speech and sound signals as an image. Many of them utilize visual properties from the spectral domain. Matsui [16] et al. propose a musical feature extraction technique based on scale invariant feature transform (SIFT), which is one of the feature extractors used in image processing. Dennis et al. [17] use visual signatures from spectrogram for sound event classification. Schutte offers a parts-based model, employing graphical model based speech representation, which is applied to spectrogram image of the speech [18]. Dennis et al. [19] propose another method for recognizing overlapping sound events by using visual local features from the spectrogram of sounds and generalized Hough transform. Apart from these time-frequency approaches, Dulas deals with the speech signal in the time domain. He proposes an algorithm for digit recognition in Polish making use of the envelope pattern of the speech signal. A binary matrix is formed by placing a grid on the speech signal of one pitch period. Similarity coefficients are, then, calculated by comparing the previous and next five matrices around the matrix to be analyzed [20, 21]. Dulas also implements the same approach for finding the inter-phoneme transitions [22].

In this paper, we propose the visual features obtained from the shapes of speech waveforms to classify vowels. We are inspired by the fact that one can determine the differences among the vowels by visually inspecting their shapes. The proposed approach, called herein Speech Vision method (SV), henceforth considers the speech waveform as an image. The images corresponding to the respective vowels are formed from two-pitch period speech segments. After applying several image processing techniques to these waveform images, some useful geometrical descriptors are extracted from them. Later, these descriptors are used for training Artificial Neural Network (ANN), Support Vector Machine (SVM), and eXtreme Gradient Boosting (XGB) models to recognize the vowels. Experiments show that comparable recognition rates are obtained. The use of visual features makes a clear distinction between the application areas of classical frequency domain approach and our suggested method. A possibility of application is in the field of language education, especially language learning of a foreign language, where one needs to test learner's pronunciation of vowels, or the learner tries to make the shape of the vowel as he/she sees both his/her own pronunciation and the ideal shape of the corresponding vowel on a screen for example. By the same token, the method could also be used in the speaking education of those with hearing disabilities. Another alternative area of application would be text to speech conversion tasks, in which ideal vowel shapes could be used in order to enhance the quality of the digital speech.

This paper is organized as follows. After this Introduction part, in Section II we discuss vowels and their properties. Section III presents the proposed method in details. Tests and results are given in Section IV, and a comparison with other vowel classification studies in the literature is carried out in Section V. Finally, conclusions and discussion appear in Section VI.

2. Characteristics of Vowels

In the Turkish language, there are 8 vowels and 21 consonants. The vowels are {a,e,ı,i,o,ö,u,ü}. There are 44 phonemes, 15 of which are obtained from vowels and the rest from consonants. The production of vowels basically depends on the position of the tongue, lips and jaw. For instance, for the vowel "a", tongue is moved back, lips are unrounded and the jaw is wide open. Therefore, all the vowels are generated differently depending on the various positions of the parts of the mouth. By considering the shape of the mouth, Turkish vowels are distributed as given in Table 1 [23]. According to this table, there are several categories for the vowels. For example, {a,ı,o,u} are vocalized with the tongue pulled back, while {e,i,ö,ü} are vocalized with the tongue pushed forward. Similarly, {a,e,ı,i} are generated with lips unrounded and {o,ö,u,ü} are generated with lips rounded. We establish vowel groups to be classified in this study according to the position of lips and jaw.

Table 1: Classes of Turkish vowels

	Unrounded (lips)		Rounded (lips)	
	Wide (jaw)	Narrow (jaw)	Wide (jaw)	Narrow (jaw)
Back (tongue)	a	ı	o	u
Front (tongue)	e	i	ö	ü

When a voice plot is stated, it is basically meant to be the graph of voice intensity against time. A sample plot of a recorded vowel is given in Figure 1.

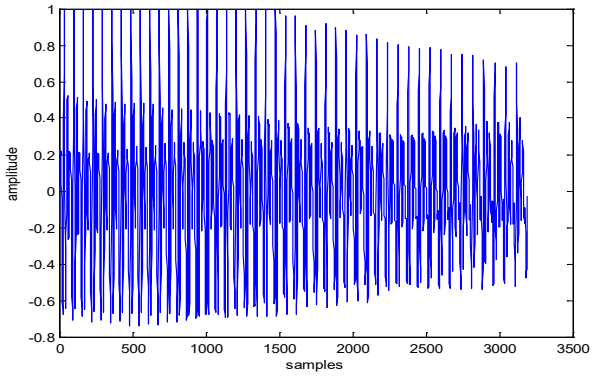


Figure 1: Waveform of Vowel "a"

It is noted that the vowel has a certain waveform. If we take a closer look, we can see that there is a repeating pattern in the waveform. This pattern is illustrated in Figure 2. The duration of each repeating pattern is known as the pitch period. This pattern keeps repeating with slight perturbations until the intensity starts to die off. When we focus on an interval of one pitch period of waveforms of all the vowels, we obtain the shapes illustrated in Figure 3. The vowels used in this paper come from a database [24]. As seen in that figure, each waveform generally differs from others in terms of appearance. The similar pattern can be experienced in certain English vowels, which sound like their corresponding

Turkish counterparts. Figure 4 shows these vowels chosen from the words within parenthesis.

The argument in this study is that the vowels can be identified by examining their waveform shapes as an image. In other words, visual features extracted from the waveform images can make vowel classification possible without the need for spectral features such as MFCC, LPCC and/or PLP coefficients. From this point of view, the proposed technique contributes to the feature selection part in speech processing. Therefore, some of the image processing and machine vision techniques are applied to those waveforms. The main novelty of this work lies in providing visual features for speech waveforms.

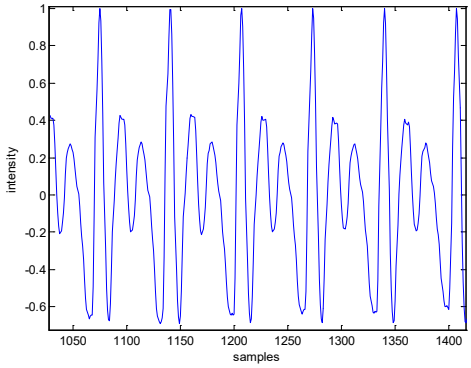


Figure 2: Repeating Patterns in Vowel "a"

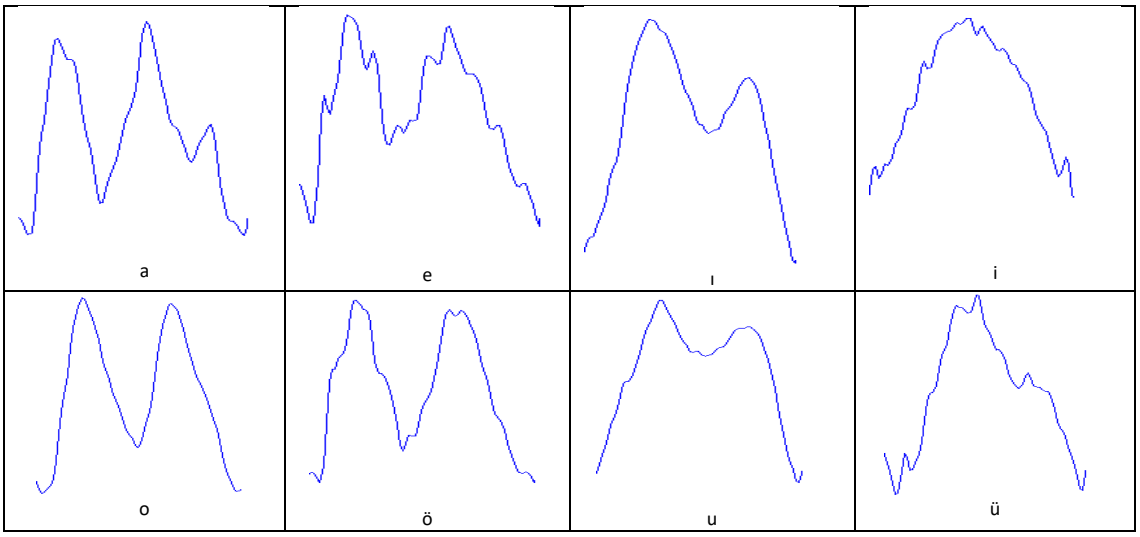


Figure 3: Sample Pitch Period Plots for 8 Turkish Vowels

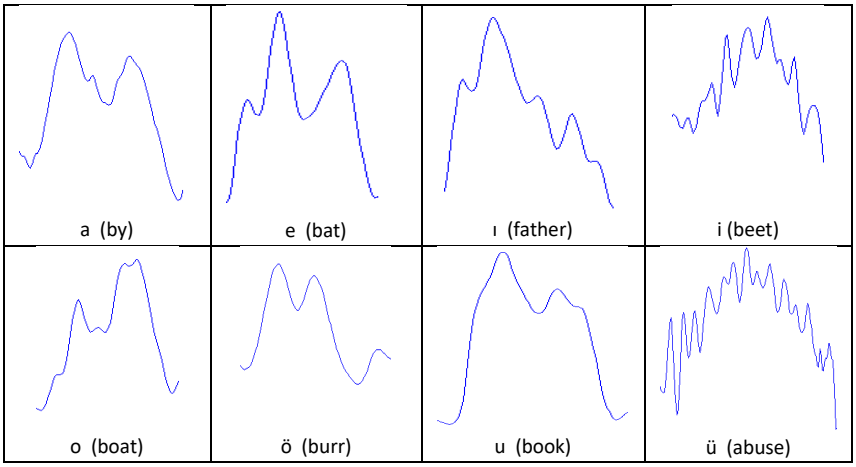


Figure 4: Sample Pitch Period Plots for 8 of the English Vowels

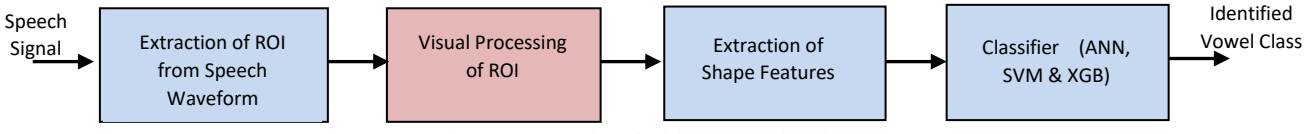


Figure 5: Proposed Method for Vowel Classification

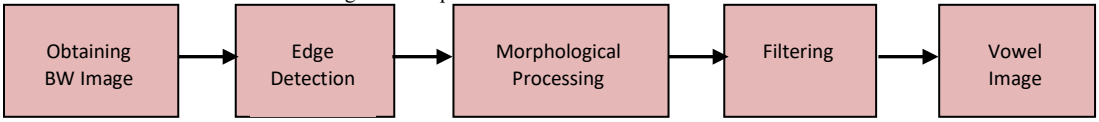


Figure 6: Operations in the "Visual Processing of ROI"

3. Speech Vision Methodology

The overall view of our proposed method is shown in Figure 5. In addition, Figure 6 shows the operations carried out in "Visual Processing of ROI" block. Our method comprises four main parts: the first is the extraction of Region of Interest (ROI), the second is visual processing of ROI, the third is extraction of shape features, and the fourth is the ANN/SVM/XGB part, where inputs are formed from the matrix and fed into the previously trained model to obtain a classification result. The following subsections explain the functions of each block in detail.

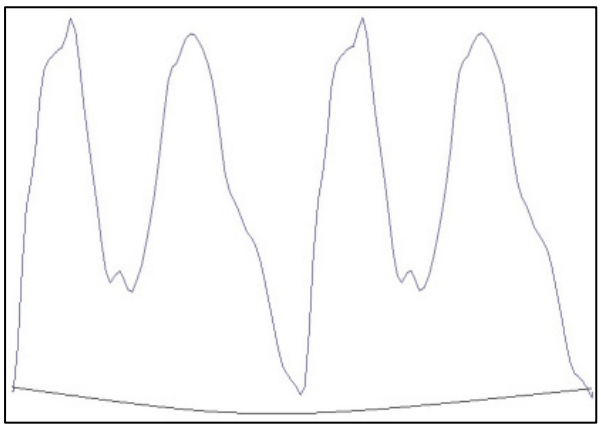


Figure 7: A Sample Closed Shape of Double Pitch Periods

3.1. Processing of region of interest

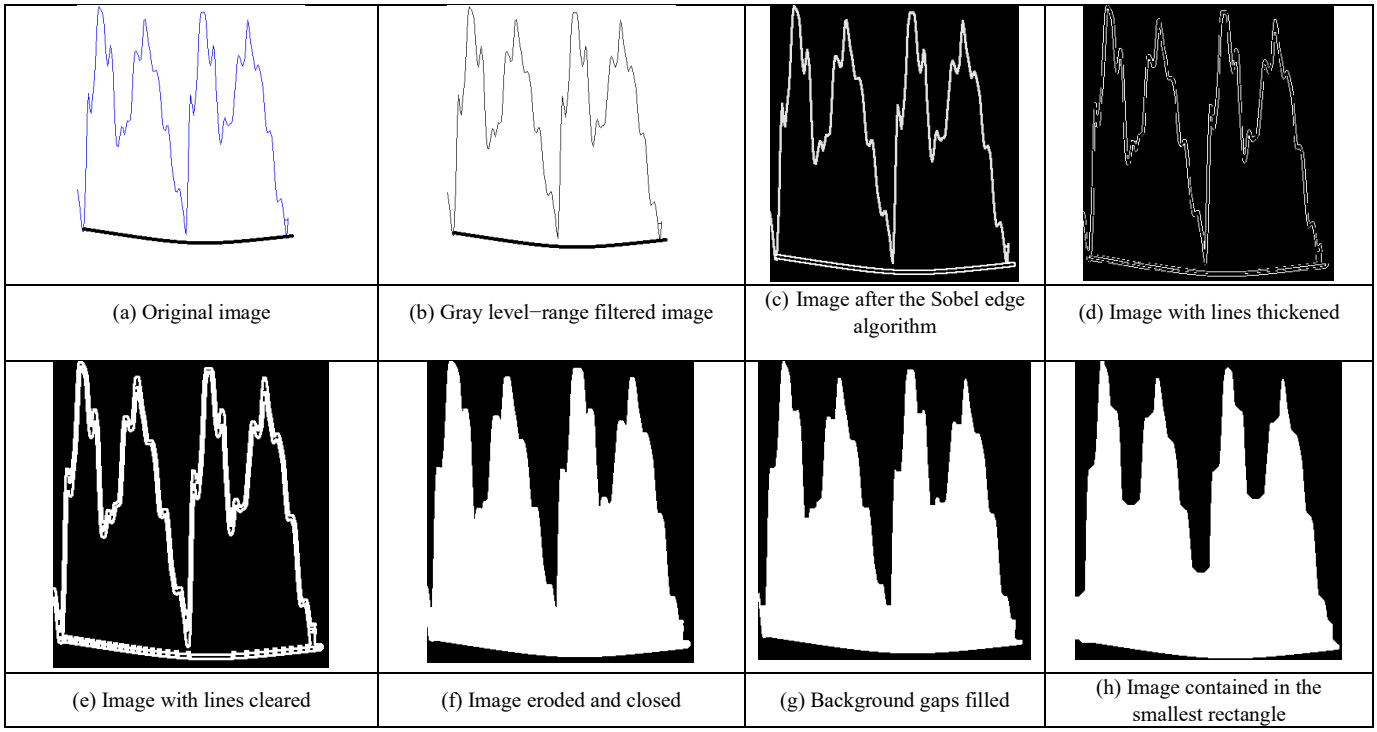
The input speech signals are segmented into two-pitch length waveform images as seen in Figure 7. The reason for choosing double pitch periods is that the shapes of a single, double, and triple pitch periods are compared, and two consecutive pitch periods give the highest scores in classification. As can be seen in Figure 8a, the image contains little jagged edges because of the noise level and the style of the speaker. In order to make vowel recognition speaker independent, one should dispose of those rapid ups and downs. Hence, we apply a sequence of image processing operations to smooth these details and, consequently, obtain a more general appearance of the waveform.

A selected waveform image to be processed is shown in Figure 8a. Then, a range filter which calculates the difference between maximum and minimum gray values in the 3x3 neighborhood of the pixel of interest is applied to the obtained gray-level image. The resulting image can be seen in Figure 8b. After this, we determine the edges of this image using Sobel algorithm with a threshold value of 0.5.

The image obtained is shown in Figure 8c. Following this, we apply a morphological structuring for line thickening, whose result is given in Figure 8d. Then, we clear the edges and borders using 4-connected neighborhood algorithm and obtain the image shown in Figure 8e. Following this operation, we erode the image and close it using a morphological closing method, whose result is shown in Figure 8f. Finally, the gaps on the background are flood-filled while changing connected background pixels (0's) to foreground pixels (1's). The result is seen in Figure 8g. A closing operation is applied to this figure and the resulting image is later contained in the smallest rectangle as depicted in Figure 8h. In the morphological operations performed on the images, we used structuring elements of line with length of 3 and angles of 0 and 90 degrees, as well as diamond with size of 1 and disk with size of 10.

3.2. Extraction of shape features

The geometric features that characterize the waveform image seen in Figure 8h are presented in this section. Since the aim is to analyze the rough shape rather than the detailed one, the features are selected in a way that represents the general



(a) Original image | (b) Gray level−range filtered image | (c) Image after the Sobel edge algorithm | (d) Image with lines thickened

(e) Image with lines cleared | (f) Image eroded and closed | (g) Background gaps filled | (h) Image contained in the smallest rectangle

Figure 8: Steps for Image Operations

silhouette of the waveform. The authors in [25] used ten features describing the general silhouettes of aircraft. In this study, we use these features along with the orientation angle as an additional feature. Table 2 lists these features. They are calculated by using the function of *regionprops* in Matlab [26]. In the features F1, F2, F6, and F11, the white region in Figure 8h, referred to as the image region, is approximated by an ellipse.

The followings are the descriptions of the features;

Table 2: Features Obtained from the Processed Image

Feature	Name of the feature
F1	*Major axis length*
F2	*Minor axis length*
F3	*Horizontal length*
F4	*Vertical length*
F5	*Perimeter*
F6	*Eccentricity*
F7	*Mean*
F8	*Filled area*
F9	*Image area*
F10	*Background area*
F11	*Orientation angle*

F1- *Major axis length*: the length of the longer axis of the image region in pixels. See Figure 9.

F2- *Minor axis length*: the length of the shorter axis of the image region in pixels. See Figure 9.

F3- *Horizontal length*: horizontal length of the image region in pixels. See Figure 10.

F4- *Vertical length*: vertical length of the image region in pixels. See Figure 10.

F5- *Perimeter*: perimeter of the image region in pixels, shown in red. See Figure 10.

F6- *Eccentricity*: a parameter of an ellipse indicating its deviation from the circularity, whose value ranges from 0 (circle) to 1 (line).

F7- *Mean*: the ratio of the total number of 1's in the binary image to the total number of pixels.

F8- *Filled area*: the total number of white pixels in the image.

F9- *Image area*: estimated area of the object in the image region which is correlated with the filled area. The area is calculated by placing and moving a 2x2 mask on an image. Depending on the corresponding pixel values in the mask, the area is computed. For example, if all the pixels in the mask are black, then the area is zero. When all are white, then the area equals one. The other distributions of pixels in the mask result in area values between zero and one.

F10- *Background area*: estimated area of the black region in the image.

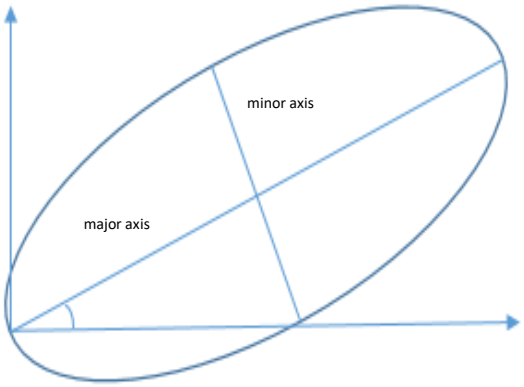


Figure 9: Major Axis, Minor Axis, and Orientation Angle

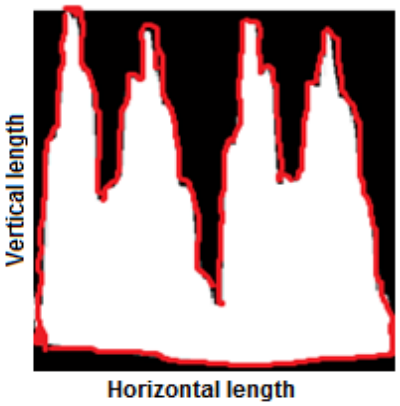


Figure 10: Horizontal and vertical lengths, and perimeter

F11- *Orientation angle*: the angle between the horizontal axis and the major axis of the ellipse approximating the image region. See Figure 10.

All the features describe the spatial domain properties of the underlying image. On the other hand, these images are the time domain representations of the speech signals. Thus, classifying the images corresponds to recognizing the speech sounds. Adopting such simple features in speech recognition leads to promising results, as shown in our experiments.

3.3. Classifiers

A general description of the employed classifiers is given here in order to facilitate a better understanding. We utilized three widely used classifiers in our study; namely ANN, SVM, and XGB method. It is well known that these are among the strongest classification tools for pattern recognition applications. They are all able to classify nonlinearly distributed input patterns into target classes. The classifiers are trained using the features in Table 2.

When sounding a vowel; the position of mouth, tongue, and lips is the key factor. The dotted and non-dotted (front and back) vowels in Turkish are quite similar in the way that only the position of the tongue changes when sounding the dotted and non-dotted vowels. Out of the eight vowels in Turkish, five vowel classes are formed in this study, combining 'dotted'

vowels with non-dotted ones. Those combined vowels were: 'ı' and 'i', 'o' and 'ö', 'u' and 'ü'. Besides, the vowels 'a' and 'e' are treated as separate classes. Therefore, these five vowel classes are considered as the outputs of the classifier.

Following a parameter optimization, an ANN is constructed with a multi-layered feed forward network structure having 11 inputs, 5 outputs, 2 hidden layers with 22 and 13 neurons, respectively. A hyperbolic tangent is chosen as activation function. The network is trained by back propagation algorithm.

As another classifier, SVM is implemented using the kernel Adatron algorithm, which optimally separates data into their respective classes by isolating the inputs, which fall close to the data boundaries. Hence, the kernel Adatron is especially effective in separating sets of data which share complex boundaries. Gaussian kernel functions are preferred in this study.

As a third classifier, a decision tree-based XGB method is used. Again, following a parameter optimization, a multi-class XGB model is employed with 89 booster trees having a maximum depth of three, whilst default values are used for the rest of the parameters.

4. Tests and Results

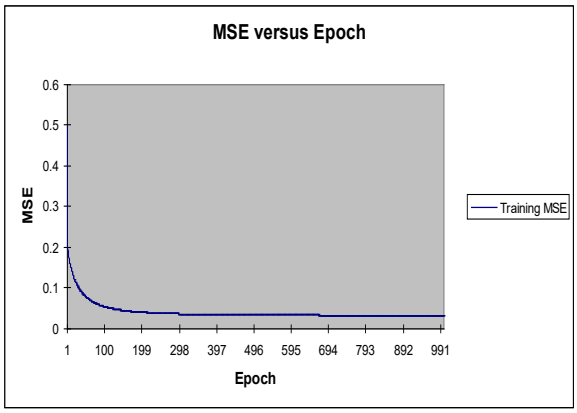
For the design of experiments, 551 samples are used; consisting of 100, 88, 90, 76, 197 samples for Class1 through Class5, respectively, for vowel classification. The vowels are parsed from the diphone database developed in [24]. Noisy conditions are not considered because we aimed to use the classification of the ideal shaped waveforms in different applications as opposed to classical voice recognition techniques. The data are randomized in order to achieve a fair distribution, 80% of which is used for training, 15% for testing, and the remaining 5% for cross validation. The ANN and SVM are trained until the results cannot improve the validation set any further. The Neurosolutions software is used for this process [27]. During the training process how the mean squared training error changes for the SVM and ANN is illustrated in Figure 11 as an example. The Python software is used for XGB modeling and training [28].

A statistical error and R-value analysis is made on the test data in order to compare the produced outputs of the trained models with the actual values that indicate whether estimations succeed or not. The results of this analysis appear in Table 3 and Table 4 for the training and test sets respectively. It can be seen from the tables that ANN and XGB perform better in terms of almost all criteria with XGB having a slightly better performance. ANN performs very well on all vowel classes except Class 4, i.e. 'o' and 'ö' vowels in Turkish; whereas XGB has more than 80% sensitivity on all classes. In Classes 1 and 3, there is a 100% correct classification for all classifiers.

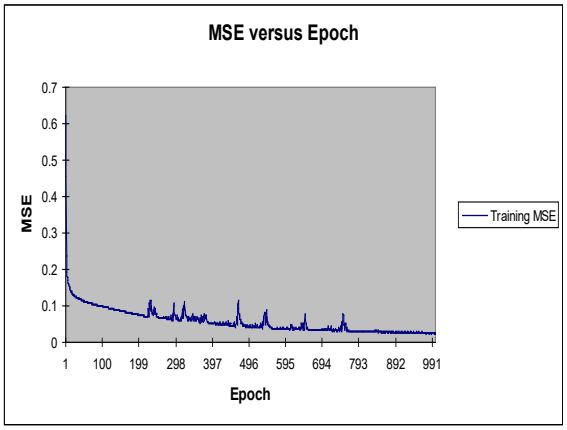
In both tables, MSE is Mean Squared Error, NMSE is Normalized Mean Squared Error, and R is linear correlation coefficient. NMSE is calculated as follows:

$$NMSE = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}(d_{ij}-y_{ij})^2}{\sum_{j=0}^{P}\frac{N\sum_{i=0}^{N}d_{ij}^{\,2}-(\sum_{i=0}^{N}d_{ij})^2}{N}} \quad (1)$$

where P is the number of output processing elements (neurons), N is the number of exemplars in the data set, y_{ij} is the network output for exemplar i at processing element j, and d_{ij} is the desired output for exemplar i at processing element j. Since NMSE is an error term, values closer to zero denote better predictability. MSE is simply the numerator of NMSE.



(a)



(b)

Figure 11: Mean Squared Error as (a) SVM and (b) ANN Training

Another statistically meaningful variable used for predictability performance is the correlation coefficient R. It is used to measure how well one variable fits on another, linear regression wise. In our case, these variables are predicted against the desired outputs. The R value is defined as:

$$R = \frac{\frac{\sum_i (y_i-\bar{y})(d_i-\bar{d})}{N}}{\sqrt{\frac{\sum_i (d_i-\bar{d})^2}{N}}\sqrt{\frac{\sum_i (y_i-\bar{y})^2}{N}}}, i = 0,\ldots,P \quad (2)$$

where y is the network output, and d_i is the desired output.

Table 3: Statistical Parameter Analysis and Comparison over Training Sets

Performance	Class 1(A)			Class 2(E)			Class 3(I-İ)			Class 4(O-Ö)			Class 5(U-Ü)		
	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB
MSE	0.0149	0.00127	0.0013	0.00588	0.0302	0.0045	0.0047	0.00133	0.0027	0.00665	0.01883	0.0049	0.0079	0.0396	0.0052
NMSE	0.0997	0.0085	0.009	0.04328	0.0222	0.034	0.03534	0.01003	0.0021	0.0567	0.16061	0.1343	0.03407	0.017	0.025
R	0.9668	0.99724	0.9876	0.9875	0.9914	0.9921	0.98271	0.99672	0.9855	0.98464	0.91815	0.9528	0.98898	0.9917	0.9901
Sensitivity (%)	100	100	100	100	100	100	100	100	100	100	99.554	100	100	100	100

Table 4: Statistical Parameter Analysis and Comparison over Test Sets

Performance	Class 1(A)			Class 2(E)			Class 3(I-?)			Class 4(O-Ö)			Class 5(U-Ü)		
	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB
MSE	0.0205	0.0006	0.0005	0.0987	0.04529	0.0095	0.01309	0.0004	0.0006	0.07955	0.04946	0.031	0.1074	0.0519	0.063
NMSE	0.1398	0.0041	0.005	0.78047	0.35815	0.089	0.08244	0.00255	0.0034	0.62906	0.39109	0.3824	0.48823	0.2359	0.2628
R	0.94	0.99836	0.9993	0.62122	0.83046	0.9172	0.95904	0.99895	0.9808	0.61131	0.78621	0.8086	0.71667	0.8796	0.8781
Sensitivity (%)	100	100	100	54.17	73.68	93.33	100	100	100	80	84.62	80	86.21	93.55	87.88

The mean squared error (MSE) can be used to determine how well the network output fits the desired output, but it does not necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, we can change the MSE without changing the directionality of the data. The correlation coefficient R solves this problem. By definition, the correlation coefficient between a network output y and a desired output d is defined by Eq. (4). The correlation coefficient is limited to the range [-1 1]. When $R = 1$, there is a perfect positive linear correlation between y and d; i.e., they vary accordingly. When $R = -1$, there is a perfect linear negative correlation between y and d;

i.e., they vary in opposite ways (when y increases, d decreases by the same amount). When $R = 0$, there is no correlation between y and d; i.e., the variables are called uncorrelated. Intermediate values describe partial correlations.

We evaluate the performances of all classifiers in terms of sensitivity, specificity, accuracy, and precision. These parameters are statistical measures for classification. Values close or equal to 100% are desirable. They are related with true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values, as explained below:

TP: Number of cases belonging to a certain class that are correctly classified.

TN: Number of cases not belonging to a certain class that are correctly classified.

FP: Number of cases belonging to a certain class that are incorrectly classified.

FN: Number of cases not belonging to a certain class that are incorrectly classified.

These parameters are calculated by the following equations:

$$Sensitivity = \frac{TP}{TP + FN} \quad (3)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (6)$$

Table 5 shows the performance results of the classifiers adapting the proposed features on the basis of sensitivity, specificity, accuracy and precision. As can be seen, the performance of each classifier justifies that the visual features can be successfully employed in vowel classification. It is noted that the ANN and XGB classifiers perform better than the SVM. Since the XGB method is decision tree-based, and not a black-box, it is possible to see which features are more useful in the model as shown in Figure 12. In fact, this is a score that denotes the goodness of each feature during the building of the boosted decision tree model based on the splits. The more the feature is used in split decisions, the higher the score. The overall score for a feature is calculated as the average of the scores of that feature across all decision trees of the model.

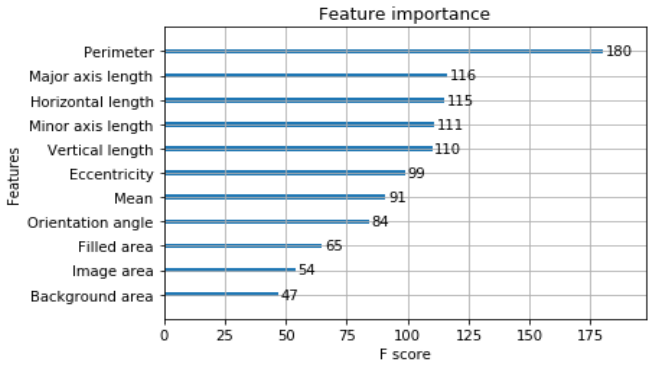


Figure 12: Feature Importance's for XGB Model

In order to show the effectiveness of the offered visual features, the same vowel classes are also classified by utilizing the MFCCs, which are commonly used for speech recognition. Table 6 depicts the sensitivity results obtained from the proposed and MFCC features classified by all three classifiers. This table also includes the classification performance of the study in [7] which classifies the Turkish vowels by MFCCs using ANN.

It is fair to say that the proposed SV method yields better results, on average, on all classes with the exception of Class 2(E). It is observed that unlike the case of visual features, when MFCCs are used SVM performs slightly better than ANN and XGB.

Table 5: Classification Parameter Analysis and Comparison over Test Sets (%)

Performance	Class 1(A)			Class 2(E)			Class 3(I-İ)			Class 4(O-Ö)			Class 5(U-Ü)		
	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB	SVM	ANN	XGB
Sensitivity	100	100	100	54.17	73.68	93.33	100	100	100	80	84.62	80	86.21	93.55	87.88
Specificity	100	100	100	97.4	98.78	94.25	100	100	100	92.31	95.45	98.85	88.89	94.29	97.1
Accuracy	100	100	100	87.13	94.06	94.12	100	100	100	91.09	94.06	96.08	88.12	94.06	94.12
Precision	100	100	100	86.67	93.33	73.68	100	100	100	53.33	73.33	92.31	75.76	87.88	93.55

Table 6: Comparison of Methods Using MFCCs and Proposed Visual Features

Method	Class 1(A)	Class 2(E)	Class 3(I-İ)	Class 4(O-Ö)	Class 5(U-Ü)
ANN (proposed visual features)	100	73.68	100	84.62	93.55
SVM (proposed visual features)	100	54.17	100	80	86.21
XGB (proposed visual features)	100	93.33	100	80	87.88
ANN (MFCC)	83.33	80	66.67	57.14	63.63
SVM (MFCC)	85.71	80	77.78	66.67	60
XGB (MFCC)	50	70	60	60	90
ANN Method in [7] (MFCC)	88	81	76	78	81

Table 7: Comparison of Various Vowel Classification Studies

Ref.No	Language	Input Features	Clasifier	Performance
7	Turkish	MFCC	ANN	80.8
29	Australian English	Frequency Energy Levels	Gaussian &ANN	88.6
30	English	MFCC	SVM	72.34
31	English	Formant frequencies	ANN	70.5
32	English	Formant frequencies	ANN	70.53
33	English	Tongue and lip movements	SVM	85.42
34	Hindi	MFCC	HMM	91.42
35	Hindi	Gammatone Cepstral Coeficients + MFCC + Formants	HMM	91.16
36	Hindi	Power Normalized Cepstral Coefficients	HMM	88.46
Speech Vision (SV)	Turkish	**Time Domain Visual Features**	ANN	**90,37**
			SVM	**84,08**
			XGB	**92,24**

4.1. Comparison with Relevant Studies

In order to evaluate the performance of the SV approach more objectively, a literature search on various vowel classification performances is also carried out. In detail, a brief comparison is given in Table 6 with the results of another study; however, this study also contained Turkish vowels [7] whereas we are keen to look into the success rates of vowel classification in different other languages. On the other hand, it should be pointed out that the indirect comparison here is just to give a rough idea about the performance of the SV approach among other vowel classification results in general.

Harrington and Cassidy conducted a study on vowel classification in Australian English, using frequency energy levels with Gaussian and ANN classifiers [29]. Indeed, there are a number of studies classifying English vowels with SVM and ANN classifiers using MFCC and Formant Frequencies [30, 31, 32]. Another study was also conducted using tongue and lip movements to classify English vowels with SVM [33]. In addition, there are a few studies on vowel classification in the Hindi language using various frequency domain features and employing Hidden Markov Model (HMM) classifiers [34, 35, 36]. A comparison of these various studies with our SV approach, in terms of sensitivity, is given in Table 7.

5. Conclusion and Discussion

This paper describes a novel approach introducing visual features for classifying vowels. The proposed approach makes use of the geometric features obtained from speech waveform shapes. Shape-based features from speech signals have rarely been employed for speech recognition. On the other hand, the features that are widely used are usually in the transform domain, i.e. spectrograms. However, the techniques using spectrograms involve computational costs due to the Fourier transform calculations. In our approach, the recorded two-pitch long speech waveform is first processed to extract the visual features. For this purpose, the waveform is treated as an image. Therefore, several aforementioned image processing techniques are utilized. Then, the features are obtained from the processed waveform image. Finally, ANN, SVM, and XGB classifiers are trained for the vowels to be classified. The test results show that using visual features accomplish quite satisfactory performances.

It is fair to say, in comparison with the success rates of classical speech features; our speech vision approach introduces a promising performance. As it can be clearly seen in Table 7, it has the highest performance with the XGB classifier, which is slightly above 92%, among all the compared studies. Additionally, our neural network and SVM classifiers result in better or comparable scores with the others. Thus, it is clear that the proposed visual features work well for Turkish vowel classification.

These features can be used in applications where the visual part would make a difference such as in teaching hearing disabled individuals to speak. Although we applied the proposed features to Turkish vowels, it could be adapted to other languages easily, since the vowels in all languages share similar characteristics in the time domain. Combining both acoustic and visual features for vowel classification can be considered for future work.

References

[1] M. Benzeghiba, et al. "Automatic speech recognition and speech variability: A review" Speech Communication, 49.10, pp. 763-786, 2007.

[2] B. Prica and S. Ilić, "Recognition of vowels in continuous speech by using formants" Facta universitatis-series: Electronics and Energetics, 23(3), pp. 379-393, 2010.

[3] S. Phitakwinai, S., H. Sawada, S. Auephanwiriyakul and N. Theera-Umpon, "Japanese Vowel Sound Classification Using Fuzzy Inference System", Journal of the Korea Convergence Society, 5(1), pp. 35-41, 2011.

[4] H. Hermansky, H. "Perceptual linear predictive (PLP) analysis of speech" Journal of the Acoustical Society of America, 87(4), pp. 1738-1752, 1990.

[5] N. Theera-Umpon, C. Suppakarn and A. Sansanee, "Phoneme and tonal accent recognition for Thai speech" Expert Systems with Applications, 38.10, pp. 13254-13259, 2011.

[6] H. Huang, et al. "Phone classification via manifold learning based dimensionality reduction algorithms" Speech Communication, 76, pp. 28-41, 2016.

[7] O. Parlaktuna, et al. "Vowel and consonant recognition in Turkish using neural networks toward continuous speech recognition" in Electrotechnical Conference Proceedings., 7th Mediterranean. IEEE, 1994.

[8] C. Vaz, A. Tsiartas and S. Narayanan, S., "Energy-constrained minimum variance response filter for robust vowel spectral estimation. In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on (pp. 6275-6279), 2014.

[9] M. Cutajar, et al. "Discrete wavelet transforms with multiclass SVM for phoneme recognition" in: EUROCON, 2013 IEEE. pp. 1695-1700, 2013.

[10] Sivaram, G. S. and H. Hermansky, H. "Sparse multilayer perceptron for phoneme recognition" Audio, Speech, and Language Processing, IEEE Transactions on, 20(1), pp. 23-29 2012.

[11] P. Thaine, and G. Penn, "Vowel and Consonant Classification through Spectral Decomposition" in Proceedings of the First Workshop on Subword and Character Level Models in NLP, pp. 82-91, 2017.

[12] M. T. Johnson, et al., "Time-domain isolated phoneme classification using reconstructed phase spaces" Speech and Audio Processing, IEEE Transactions on 13.4: pp. 458-466, 2005.

[13] J. Manikandan, B. Venkataramani, P. Preeti, G. Sananda and K. V. Sadhana, "Implementation of a phoneme recognition system using zero-crossing and magnitude sum function" in TENCON 2009 IEEE Region 10 Conference, pp. 1-5, 2009.

[14] T. J. Lee and K. B. Sim, "Vowel classification of imagined speech in an electroencephalogram using the deep belief network" Journal of Institute of Control, Robotics and Systems, 21(1), pp. 59-64, 2015.

[15] K. I. Han, H. J. Park and K. M. Lee, " Speech recognition and lip shape feature extraction for English vowel pronunciation of the hearing-impaired based on SVM technique" in Big Data and Smart Computing (BigComp), 2016 International Conference on pp. 293-296. IEEE, 2016.

[16] T. Matsui, et al. "Gradient-based musical feature extraction based on scale-invariant feature transform" in Signal Processing Conference, 19th European. IEEE, pp. 724-728, 2011.

[17] J. Dennis, H. D. Tran, L. I. Haizhou, "Spectrogram image feature for sound event classification in mismatched conditions", Signal Processing Letters, IEEE, 18.2, pp. 130-133, 2011.

[18] K. T. Schutte, "Parts-based models and local features for automatic speech recognition" PhD Thesis, Massachusetts Institute of Technology, 2009.

[19] J. Dennis, H. D. Tran and E. S. Chng, "Overlapping sound event recognition using local spectrogram features and the generalised hough transform" Pattern Recognition Letters, 34.9, pp. 1085-1093, 2013.

[20] J. Dulas, "Speech Recognition Based on the Grid Method and Image Similarity" INTECH Open Access Publisher, 2011.

[21] J. Dulas, "Automatic word's identification algorithm used for digits classification" Przeglad Elektrotechniczny, 87, pp. 230-233, 2011.

[22] J. Dulas, "The new method of the inter-phonemes transitions finding" Przegląd Elektrotechniczny, 88.10a, pp. 135-138, 2012.

[23] İ. Ergenç, Spoken Language and Dictionary of Turkish Articulation, Multilingual, 2002.

[24] Ö. Salor, B. L. Pellom, T. Ciloglu, K. Hacioglu, and M. Demirekler, "On developing new text and audio corpora and speech recognition tools for the Turkish language" in INTERSPEECH, 2002.

[25] A. G. Karacor, E. Torun E. and R. Abay, "Aircraft Classification Using Image Processing Techniques and Artificial Neural Networks" International Journal of Pattern Recognition and Artificial Intelligence, Vol. 25 No 8, pp. 1321-1335, 2011.

[26] Matlab, The Mathworks Inc. (1984 - 2013), www.mathworks.com.

[27] Neurosolutions,NeuroDimension Inc,(1994-2015), www.neurosolutions.com

[28] G. Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), 1995.

[29] J. Harrington, and S. Cassidy, "Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs" Australian English. Language and Speech, 37(4), pp. 357-373, 1994.

[30] P. Clarkson, and P. J. Moreno, "On the use of support vector machines for phonetic classification" in Acoustics, Speech, and Signal Processing, 1999. Proceedings, 1999 IEEE International Conference on, pp. 585-588, 1999.

[31] R. Carlson, and J. R. Glass, "Vowel classification based on analysis-by-synthesis" in ICSLP, 1992.

[32] P. Schmid, and E. Barnard, "Explicit, n-best formant features for vowel classification" in Acoustics, Speech, and Signal Processing, ICASSP-97., 1997 IEEE International Conference on, pp. 991-994, 1997.

[33] J. Wang,, J. R. Green, and A. Samal, "Individual articulator's contribution to phoneme production" in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on pp. 7785-7789, 2013.

[34] S. Mishra, A. Bhowmick and M. C. Shrotriya, "Hindi vowel classification using QCN-MFCC features" Perspectives in Science, 8, pp. 28-31, 2016

[35] A. S. T. I. K. Biswas, P.K. Sahu, A. N. I. R. B. A. N. Bhowmick, and M. A. H. E. S. H. Chandra, "Hindi vowel classification using GFCC and formant analysis in sensor mismatch condition" WSEAS Trans Syst, Vol. 13, pp. 130-43, 2014.

[36] M. Chandra, "Hindi Vowel Classification using QCN-PNCC Features" Indian Journal of Science and Technology, 9(38), 2016.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 25-30 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

An Immutable Algorithm Approach to Improve the Information Security of a Process for a Public Organization of Ecuador

Segundo Moisés Toapanta Toapanta*[1], Andrés Javier Bravo Jácome [1], Maximo Giovanny Tandazo Espinoza[1], Luis Enrique Mafla Gallegos[2]

[1] *Departament of Engineering Systems, Universidad Politécnica Salesiana(UPS), Guayaquil, Ecuador*

[2] *Faculty of Engineering Systems, Escuela Politécnica Nacional del Ecuador (EPN), Quito, Ecuador*

ARTICLE INFO

Article history:
Received: 28 March, 2019
Accepted: 22 April, 2019
Online: 07 May, 2019

Keywords:
Algorithms
Cryptography
Security

ABSTRACT

Currently, information security is among the main characteristics that must be achieved within the security of private and public organizations worldwide. For this reason, globally recognized algorithms such as the AES, IDEA, RC5, DES, RSA are researched with the aim of identifying the most suitable and obtaining a greater degree of security and speed of encryption in order to mitigate the information vulnerabilities between processes and be applied as a feasible alternative in an electoral process. The deductive method was used to analyze the information obtained in the references. After the study it is possible to conclude that to improve security in the processes of public organizations in Ecuador it is necessary to implement cryptographic mechanisms.

1. Introduction

The arrival of Information Technologies for the digital communications of the organizational processes, security problems have increased in an increasing way, that is why it is important to know about the different existing algorithms in order to reduce security problems of the information[1]. In the last decade there have been incidents where computer systems are vulnerable in the presence of hackers, cybercriminals and hacktivists, for such events governments have analyzed the negative effects that can cause these attacks and have developed different defense strategies to deal with the different attempts of intrusions. Most private and public organizations implement the use, creation or customization of cryptographic algorithms specialized in safeguarding the security of digital communications.

Why is it imperative to implement immutable algorithms in the processes of public organizations in Ecuador?

It is necessary to ensure the security of the information of possible attacks of theft, and to guarantee the security in the communications between the processes of the public organizations of Ecuador.

The general objective of the research is to establish a cryptographic algorithm could use a public institution of the Ecuadorian state that provides greater security in the communication and treatment of information.

The articles analyzed in relation to the subject are:

Performance analysis of encryption algorithms for security[1], A Comparative and Analytical Study on Symmetric[2], Dessign of new security algorithm[3], Comparative Analysis of NPN Algorithm & DES[4], Proposed Symmetric Key Cryptography Algorithm[5], Comprehensive Study of Symmetric Key and Asymmetric Key Encryption Algorithms[6], Performance Evaluation of Cryptographic Algorithms: DES and AES[7], DES and AES Performance Evaluation[8], Differential fault analysis against AES-192 and AES-256 with minimal faults[9] ,Implementing the IDEA Cryptographic Algorithm in Virtex-E and Virtex-II FPGAs[10], User Defined Encryption Procedure for IDEA Algorithm[11], Performance evaluation for CAST and RC5 encryption algorithms[12], Selection of parameter 'r' in RC5 algorithm on the basis of prime number [13] , Design and implementation of algorithm for des cryptanalysis [14], A-RSA: Augmented RSA [15], High speed implementation of RSA algorithm with modified keys exchange[16].

The deductive method was used to analyze the information obtained in the references and identify the characteristics of each of the security algorithms; that allow improving the security of organizations. This method is taken in view of the fact that the

*Segundo Moisés Toapanta Toapanta, Email: stoapanta@ups.edu.ec

information obtained must be analyzed to consider the qualities presented by each of the algorithms.

In this research phase it can be obtained that AES cipher algorithm is the most efficient due to its features, working together with hardware and software for the application of cryptography in digital communications processes.

2. Materials and Methods

2.1.1 Materials

Cryptography is the skill of writing in an enigmatic way, that is, it is a process of transformation of any data readable to an encrypted data. This ensures that the data can't be objective of any attack coming from the organization or outside it. For the interaction of processes applying cryptography it is necessary to apply the encrypted method and decrypted method[2].

This science allows the secure transmission of sensitive information in unsafe processes so that it can't be interpreted by third parties[3].

Cryptography, in addition to providing confidentiality and privacy, within its main features that it provides are: authentication, data integrity, non-repudiation, etc[4,5]. The two main methods are Symmetric Key Cryptography and Asymmetric Key Cryptography[6].

Immutable Algorithm refers to the fact that the data it handles within its encryption process are not going to be modified or altered, the immutable algorithms are algorithms that work in blocks, for that reason not even a single bit of each output of the round.

2.1.2 Definition AES Algorithm

AES is used as a standard algorithm for US federal organizations. AES consists of a key mechanism of 128 bits, 192 bits and 256 bits. Starting from an initial key of 16 bytes (128 bits), which we can show as a block or matrix of 4x4 bytes, 10 keys are generated, these resulting keys plus the first key are called subkeys[7,8].

The algorithm is classified into AES-128, AES-192 and AES-256 which have 10 rounds, 12 rounds and 14 rounds respectively. Each of the rounds is composed of 4 transformations, unlike the last round. These transformations are: SubBytes, ShiftRows, MixColumns and AddRoundKey, as mentioned the last round lacks a transformation which is MixColumns[9].

- SubBytes: It consists of 16 identical boxes (8x8). In this step, a non-linear substitution is performed where each byte is replaced with another, that is, $Sb_{ij} = S(a_{ij})$. The denotative is SB.
- ShiftRows: In this step a transposition is carried out, each row of the box moves cyclically in different displacements. Row 0 does not scroll, row 1 moves by 1 byte, row 2 moves by 2 bytes and row 3 by 3 bytes. The denotative is SR.
- MixColumns: This is a mixing operation that runs in the columns of the box, coupling the 4 bytes in each column using a linear transformation. The denotative is MC.
- AddRoundKey: This is an XOR operation bit by bit with the key of the round.

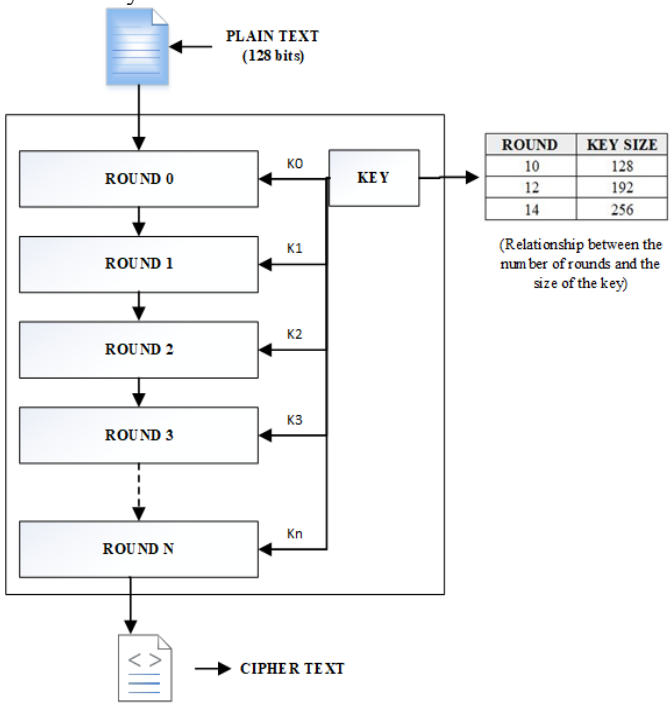


Figure 1: Process AES (Advanced Encryption Standard).

2.1.3 Definition IDEA Algorithm

IDEA is an encryption algorithm that works with a block of 64-bit flat text. It implements a 128-bit input key that it uses to generate 52 subkeys of 16 bits each. The decryption process is the same encryption process but applied in reverse[10].

The encryption stream contains a total of 8 rounds, after the round number 8 performs a transformation in the output. Its operation consists of the first four sub-blocks are 16 bits of key which combines them with four blocks of 16-bit flat text. The exit of each round is the entrance of the next round[11].

IDEA is an encryption algorithm that is based on the concepts of confusion and diffusion, implementing elementary operations, are the following:

- XOR
- Sum of module 2^{16}
- Module product $2^{16} + 1$

In a round of IDEA:

1. Multiply X1 by sub-key Z1.
2. Add X2 with sub-key Z2.
3. Add X3 with sub-key Z3.
4. Multiply X4 by sub-key Z4.
5. XOR between step 1 and step 3.
6. XOR between step 2 and step 4.
7. Multiply step 5 by subkey Z5.

8. Add the step 6 and step 7.
9. Multiply step 8 by subkey Z6.
10. Add the step 7 and step 9.
11. XOR between step 1 and step 9.
12. XOR between step 3 and step 9.
13. XOR between step 2 and step 10.
14. XOR between step 4 and step 10[11].

Transformation of the output:

1. Multiply X1 by sub-key Z1.
2. Add X2 with sub-key Z2.
3. Add X3 with sub-key Z3.
4. Multiply X4 by sub-key Z4[11]

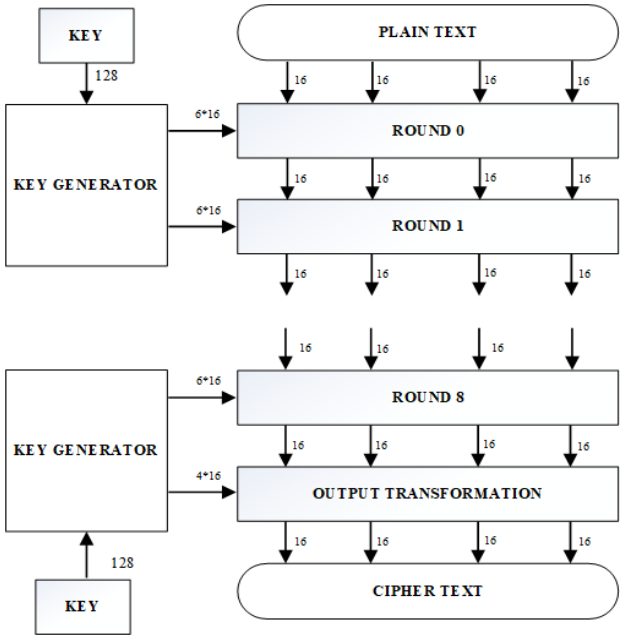


Figure 2: Process IDEA (International Data Encryption Algorithm)

2.1.4 *Definition RC5 Algorithm*

RC5 was created by the cryptographer Ronald Rivest and is a symmetric method block cipher algorithm. Its operation is simple and very fast since it only implements three classes of computational operations (such as XOR, shift, etc.) and a minimum memory consumption[12].

To encrypt the entry of the block is a plain text that is divided into 2 sub-blocks A and B, the output of this block is an encryption text of 2 w-bit length. Equation(1) shows the operations to encrypt[13].

$$
\begin{aligned}
&A = A + S[0];\\
&B = B + S[1];\\
&\text{For } i = 1 \text{ to } r \text{ do}\\
&A = ((A \text{ XOR } B) << B) + S[2*i];\\
&B = ((B \text{ XOR } A) << A) + S[2*i+1];
\end{aligned}
\tag{1}
$$

To decrypt the cipher text, it is treated as a data block and again divided into two sub-blocks. Therefore, one can arrive at the deduction that the decryption method is the inverse of encrypting. Equation(2) shows the operations to encrypt[13].

$$
\begin{aligned}
&\text{For } i = r \text{ downto } 1 \text{ do}\\
&B = ((B - S[2*i+1]) >> A) \text{ XOR } A;\\
&A = ((A - S[2*i]) >> B) \text{ XOR } B;\\
&B = B - S[1];\\
&A = A - S[0];
\end{aligned}
\tag{2}
$$

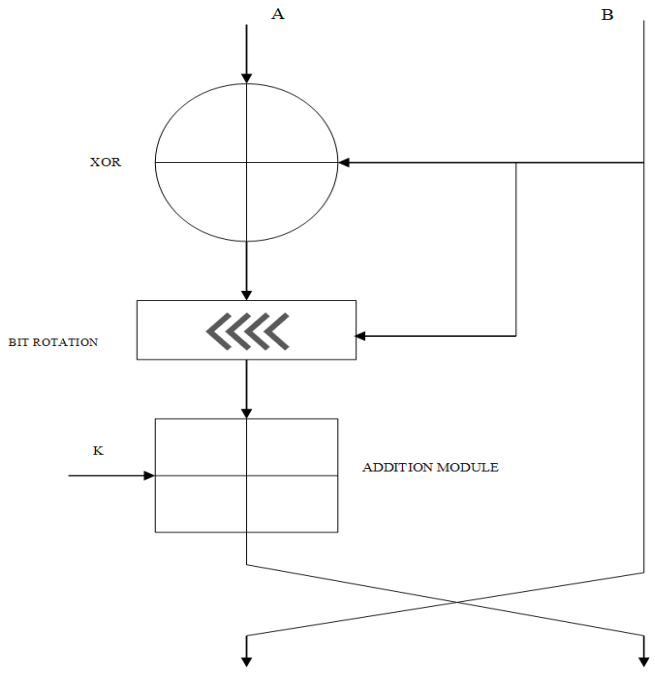


Figure 3: Process RC5 (Rivest Cipher 5)

2.1.5 *Definition DES Algorithm*

DES is a symmetric key encryption algorithm, it is among the first encryption methods that was implemented commercially, it was mainly used as a security standard for the processing of federal information in the US. It is an encryption algorithm that operates on data bocks, 64-bit blocks, with a 56-bit secret key. It has 16 rounds each with two permutations[14].

The encryption process consists of two permutations called P boxes, which correspond to preliminary permutation and last permutation, and sixteen rounds of Feistel. Each of the 16 round uses a different 48-bit key-key generated by the encryption key that implements a predefined algorithm.

The f function of Feisnel is composed of four sections:

1. Box expansion P
2. Mix of box P.
3. Replacement
4. Permutation

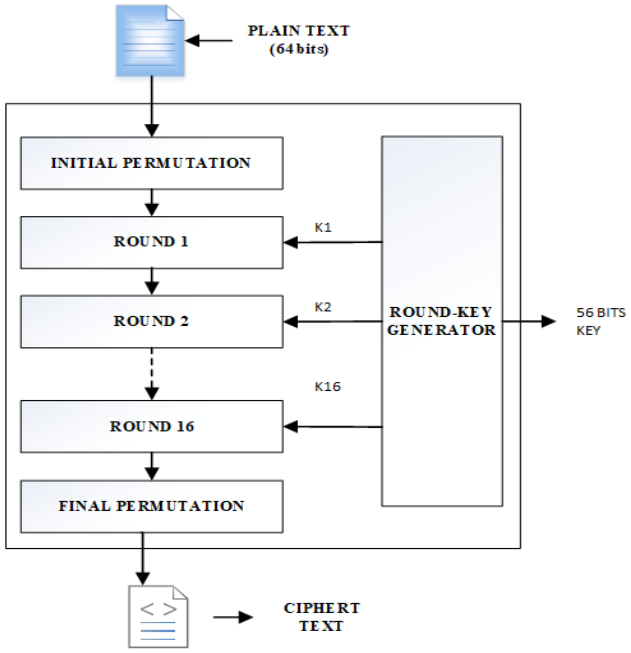


Figure 4: Process DES (Data Encryption Standard)

2.1.6 *Definition RSA Algorithm*

RSA was founded in 1978 and is an asymmetric cryptographic algorithm named after its founders Rivest, Shamir and Adelman[15]. It is one of the most popular and recognized for the exchange of keys, digital signatures, encryption of data blocks. RSA uses an encryption of variable size of blocks and a key of variable size.

It uses 2 prime numbers to generate the public key and private keys, its size is from 128 to 4096 bits. These two different keys are used in order to encrypt and decipher[15]. It should be noted that RSA is absolutely slow in its methods of encrypting and decrypting; therefore, it is not recommended for large data[16].

RSA encryption is simply a modular expression. The module "n" is not more than the product of 2 large prime numbers (between 100 and 300 digits) chosen at random, both the public key and the private key is obtained from the following equation (3).

$$e = d\text{-}1 \bmod Ø(n) \quad (3)$$

The encryption operation is performed by public keys "n" and "e" of the following equation (4).

$$C = M^e \pmod n \quad (4)$$

While to recover the original message from the encrypted message is done with the equation (5).

$$M = C^d \pmod n \quad (5)$$

2.1.7 *Methods*

We analyzed the available data of the different cryptographic algorithms AES, RC5, IDEA, DES and RSA and it was deduced from the analysis which would be the best to apply it within the public organizations of Ecuador and as a possible alternative of security in digital electoral processes.

Secondly, the following characteristics have been considered as the relevant points to proceed to evaluate the cryptographic algorithms AES, IDEA, RC5, DES and RSA.

Table 1: Selected features for the evaluation of the AES, IDEA and RC5 algorithms

Characteristic	AES	IDEA	RC5	DES	RSA
Type	Symmetric Algorithm	Symmetric Algorithm	Symmetric Algorithm	Symmetric Algorithm	Asymmetric Algorithm
Key length (bits)	128 bits 192 bits 256 bits	128 bits	128 bits 192 bits 256 bits	56 bits	128, 256, 1024, 2048 y 4096 bits
Block size	128 bits	64 bits	32 bits 64 bits 128 bits	64 bits	-
Number of Rounds	10, 12 o 14	8	12	16	-

Then algorithms were implemented using Netbeans 8.2 as a development environment. The algorithms are programmed in JAVA language under OS Windows 10. The test platform is a laptop (ASUS Q504U) with Intel Core i5-7200U CPU 2.5GHZ 2.71GHZ and 12GB of RAM.

The speed test consists of the time that an encryption algorithm takes to transform a plain text. The speed test helps us measure the performance in units of time. In this paper we also consider the decryption speed for the algorithms to be evaluated.

3. Results

3.1. Cipher Test

In this test we have used 3 files of different sizes, 1 Mb, 10 Mb, and 100 Mb; which will be the files to be encrypted in order to obtain the times when encrypting the file. We have to keep in mind that each cryptographic algorithm uses different key sizes.

Table 2: Results of the cipher times obtained for the cryptographic algorithms with their different keys.

FILE (MB)	AES 256 BITS	IDEA 128 BITS	RC5 256 BITS	DES 256 BITS	RSA 128 BITS
1 MB	21 ms	40 ms	67 ms	70 ms	-
10MB	59 ms	242 ms	617 ms	418 ms	-
100MB	1489 ms	1920 ms	1874 ms	3804	-

3.2. Decryption Test

In this test we perform the decryption of the files of the previous test in order to obtain the time it takes for each algorithm to decrypt.

Table 3: Results of decryption times obtained for cryptographic algorithms.

FILE (MB)	AES 256 BITS	IDEA 128 BITS	RC5 256 BITS	DES 256 BITS	RSA 128 BITS
1 MB	36 ms	46 ms	67 ms	56 ms	-
10MB	230 ms	390 ms	617 ms	474 ms	-
100MB	2125 ms	2269 ms	2745 ms	3825	-

3.3. Analysis of result

In the following table we show a weighting assigned to each cryptographic algorithm by the characteristics of key strength and encryption speed. The assigned value corresponds to the value obtained by each algorithm (1 to 3, with 3 being the highest value and 1 being the lowest). As a result, we can show that the AES algorithm provides greater security and the IDEA gives us less security compared to the RC5 algorithm, and as the algorithms of less security this DES and finally we have the RSA algorithm.

Table 4: Analysis of the results.

Characteristic	Cryptographic Algorithm				
	AES	IDEA	RC5	DES	RSA
Key strength	3	2	3	1	3
Encryption speed	3	1	2	2	-
Decryption speed	3	2	3	1	-
Total	**9**	**5**	**8**	**4**	**3**

In the following figure we can compare the weights and show that the algorithm AES has better results in comparison with the algorithms IDEA, RC5, DES and RSA. It is worth noting that the RSA algorithm could not be performed because it is only capable of encrypting data of less than 254 bytes and the files under test exceeded its maximum quota.

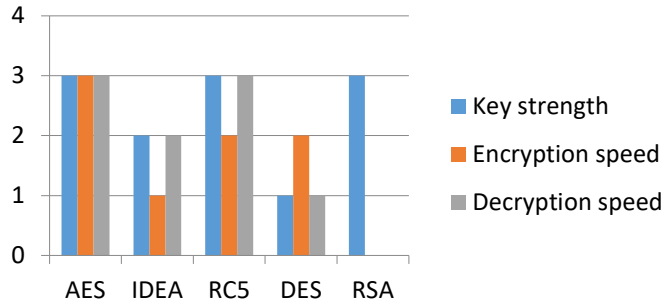


Figure 5: Analysis of results. The evaluation diagram by characteristic of each cryptographic algorithm is shown

a. The AES cryptographic algorithm has a value of 3 in the key strength characteristic, this indicates that it has a greater number of combinations among the other analyzed algorithms IDEA and RC5.
b. AES cryptographic algorithm has a greater number of stages (rounds), this means that its design is more complex than the IDEA and RC5 algorithms.
c. The cryptographic algorithm AES has the highest number in encryption speed however it has the same value in deciphering as algorithm RC5. This means that both are fast, however, the AES cryptographic algorithm is faster at the time of encrypting.

4. Discussion

The majority of developed countries adopt the use of proprietary algorithms since these are designed to their needs and vulnerabilities, there are developing countries which implement the use of immutable algorithms, with the observations made on the cryptographic algorithms it is possible to show that the cryptography has a fundamental role in information security.

The results obtained in this investigation were the criteria of the authors on the importance of the information; it is considered that the AES algorithm can be an alternative to improve the security of the information for a process of a public organization of Ecuador or any other organization.

For this, it was obtained as a result that the AES cryptographic algorithm is the most optimal, according to the score obtained; Within the research, the variables that characterize it were identified: key strength, block size, number of rounds and encryption speed; characteristics mentioned by the authors of the references(1-16).

AES exceeds the algorithms DES, RC5, IDEA, RSA; both in hardware and in software, it is considered as a combination of security, speed, performance and applicability. It also provides greater security by its multiple possible combinations between block sizes, as well as in the use of longer keys and its 3 possible numbers of rounds.

Can speak of immutability on the AES encryption algorithm; its method of encryption of rounds provides us with the security that the exit data of each round can't be altered on the way to the entrance of the next round.

The implementation of the AES cipher can be considered as a reference for the design of the security of the processes of public organizations of Ecuador, as well as possible viable alternative for an electoral process.

It is concluded that with the implementation of the AES cipher algorithm to improve the security of the information and the treatment thereof, guaranteeing the availability, integrity and confidentiality of the communications and services offered by the public organizations of Ecuador.

5. Conclusion and Future Work

The cipher algorithms play a very important role in the security of the processes inside and outside the organizations, our work in the research was to study the algorithms AES, IDEA, RC5, DES and RSA. By implementing the AES Algorithm, it was obtained that it is the fastest algorithm and gives us greater security.

It is advisable to achieve a degree of security in the keys, for this it is necessary to take into consideration the use of extensive keys, alphanumeric keys, keys with the use of uppercase and lowercase, use of special characters; this way we can delay the time in which he performs a brute force attack on a cryptographic algorithm.

The training of IT staff in cryptology, since in the future they could be responsible for technological advancement and implement their own cryptographic algorithms. Reinforcing information security policies in organizations, since the established rules could be violated due to ignorance or negligence.

It is recommended as future works the analysis of algorithms with the use of Hash methods to guarantee the integrity of the data against brute force attacks.

6. Acknowledgment

The authors thank the Salesian Polytechnic University of Ecuador, the research group of the Guayaquil Headquarters "Information Technology, Security and Information for a Globalized World" (CSITGW) created in accordance with resolution 142-06-2017-07-19 and the Secretariat of Higher Education Science, Technology and Innovation (Senescyt).

References

[1] M. Panda, "Performance analysis of encryption algorithms for security," *Int. Conf. Signal Process. Commun. Power Embed. Syst. SCOPES 2016 - Proc.*, pp. 278–284, 2017.

[2] B. Mandal, S. Chandra, S. S. Alam, and S. S. Patra, "A comparative and analytical study on symmetric key cryptography," *2014 Int. Conf. Electron. Commun. Comput. Eng. ICECCE 2014*, pp. 131–136, 2014.

[3] M. Dubai, T. Mahesh, and P. Ghosh, "Design of new security algorithm: Using hybrid Cryptography architecture," *3rd Int. Conf. Electron. Comput. Technol. (ICECT), 2011.*, pp. 99–101, 2011.

[4] M. Sharma, R. B. Garg, and S. Dwivedi, "Comparative analysis of NPN algorithm & des Algorithm," *Proc. - 2014 3rd Int. Conf. Reliab. Infocom Technol. Optim. Trends Futur. Dir. ICRITO 2014*, 2015.

[5] A. Anand, A. Raj, R. Kohli, and V. Bibhu, "Proposed symmetric key cryptography algorithm for data security," *2016 1st Int. Conf. Innov. Challenges Cyber Secur. ICICCS 2016*, no. Iciccs, pp. 159–162, 2016.

[6] M. B. Yassein, S. Aljawarneh, E. Qawasmeh, W. Mardini, and Y. Khamayseh, "Comprehensive study of symmetric key and asymmetric key encryption algorithms," *2017 Int. Conf. Eng. Technol.*, pp. 1–7, 2017.

[7] S. Kansal and M. Mittal, "Performance Evaluation of Various Symmetric Encryption Algorithms," *Int. J. Netw. Secur. Its Appl.*, vol. 6, no, pp. 105–109, 2014.

[8] B. Bhat, A. W. Ali, and A. Gupta, "DES and AES performance evaluation," *Int. Conf. Comput. Commun. Autom. ICCCA 2015*, pp. 887–890, 2015.

[9] C. H. Kim, "Differential fault analysis against AES-192 and AES-256 with minimal faults," *Fault Diagnosis Toler. Cryptogr. - Proc. 7th Int. Work. FDTC 2010*, pp. 3–9, 2010.

[10] J. M. Granado, M. A. Vega, J. M. Sanchez, and J. A. Gomez, "Implementing the IDEA Cryptographic Algorithm in Virtex-E and Virtex-II FPGAs," vol. 03, pp. 109–112, 2006.

[11] V. S. Prajwal and K. V. Prema, "User Defined Encryption Procedure for IDEA Algorithm," *2018 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2018*, pp. 1668–1671, 2018.

[12] T. Nie, Y. Li, and C. Song, "Performance evaluation for CAST and RC5 encryption algorithms," *2010 Int. Conf. Comput. Control Ind. Eng. CCIE 2010*, vol. 1, pp. 106–109, 2010.

[13] H. S. Gill, "Selection of parameter 'r' in RC5 algorithm on the basis of prime number," *2014 Recent Adv. Eng. Comput. Sci. RAECS 2014*, pp. 1–4, 2014.

[14] H. D. Zodpe, P. W. Wani, and R. R. Mehta, "Design and implementation of algorithm for des cryptanalysis," *Proc. 2012 12th Int. Conf. Hybrid Intell. Syst. HIS 2012*, pp. 278–282, 2012.

[15] A. Karakra and A. Alsadeh, "A-RSA: Augmented RSA," *Proc. 2016 SAI Comput. Conf. SAI 2016*, pp. 1016–1023, 2016.

[16] S. A. Nagar and S. Alshamma, "High speed implementation of RSA algorithm with modified keys exchange," *2012 6th Int. Conf. Sci. Electron. Technol. Inf. Telecommun. SETIT 2012*, pp. 639–642, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 31-37 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Resourceful Residual Energy Consumption in TDMA Scheduling for IoT-based Wireless Sensor Network

Amir Rizaan Rahiman*, Md. Ashikul Islam, Md. Noor Derahman

Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400, Malaysia

ARTICLE INFO

Article history:
Received: 04 March, 2019
Accepted: 23 April, 2019
Online: 09 May, 2019

Keywords:
IoT
Wireless Sensor Network (WSN)
Time Division Multiplexing Algorithm (TDMA)
Sensor node
Energy consumption
Broadcasting

ABSTRACT

Recently, wireless sensor network (WSN) gets more concern due to the robustness in the latest communication technology iteration such as big data, IoT and 5G. Such daily usage of these technologies includes smart home, smart farming, smart traffic control etc. Moreover, WSN becomes the best preference for mobile objects in data accumulating in a wild range area. Routing distance, signal interference and routing computational cost give a significant impact to the WSN nodes lifetime. Unsynchronized node time allocation slot and neighbor discovery are the main factors in the energy consumption issue faced by the WSN. Higher energy consumption reduces the network lifetime and WSN nodes performance. This paper discusses the optimization of energy-topology (E-T) factors for distributed time division multiplexing algorithm (TDMA) slot scheduling for high-speed data link capacity. The E-T factor is based on the influence of residual energy and topology on the time slot allocation. Both node residual energy and topology information have shown a respectable impact on the TDMA node slot allocation. Moreover, the numbers of neighbors and the network residual energy have been proved both nodes execution time and energy utilization can be reduced in the algorithm. The algorithm performance has been evaluated based on the previous experiment parameters with new high-speed data link. The experimental results have shown a significant improvement in residual energy consumption for the proposed optimized TDMA slot allocation

1. Introduction

The recent revolution in the current Internet technology is the Internet of Things (IoT). It becomes a new communication technology enabling the development of the connection between different types of platforms like smart home, smart traffic monitoring system, autonomous vehicles, military operation, environment monitoring, sizeable agricultural area monitoring, cave and mining hole etc. [1]. Moreover, it becomes the main reason for the interconnection and co-operation between cloud computing, mobile Internet, distributed network and many more. Today, it slowly overtaking the modern wireless telecommunications, it's a novel paradigm. The Wireless Sensor Network (WSN) is the medium between this interconnection communication. It consists of sensor nodes handling the information transfer between source and sink or base station. Fig. 1 shows the common WSN architecture.

The usage amount of the WSN nodes growth as much as the IoT grows. All real-time systems totally depend on a wireless sensor for various tasks, such as data collection, environmental monitoring or even decision making according to the event occurs. Recently, advanced technology makes communication devices smaller with highly integrated low power consumption. Sometimes those devices are including tiny microprocessors supplied by the tiny scale of energy [2]. The WSNs is a combination of advanced micro devices that are used in collecting the IoT data due to its real-time processing system and easy deployment capability. It is cost effective, low power consumption, easy to implement with a multifunctional wireless system that bridge between physical space and information space. One of the greatest advantages of the WSN is it can distribute over a large geographical area which can be even between two cities or countries or can be into the ocean to monitor events like earthquake, tsunami etc.

However, the main limitation of WSNs is limited energy resource and lifetime where the sensor node is powered by small batteries. In addition, it is hard to change or add a new battery or energy resource due to rugged environments. For example, in the military and dangerous experimental works like checking volcano

* Amir Rizaan Rahiman, Email: amir_r@upm.edu.my

eruption, war zone monitoring, and deep-ocean search and rescue are benefited from the WSN. The period of monitoring would be a few days, months or years. For a large and long period network, only battery power is a suitable and efficient solution to power to the nodes. Generally, in a wireless environment, low energy efficiency and high data rate network become a challenging topic since the sensor nodes usually battery driven. Thus, managing the limited power supply in a proper and efficient way becomes an important task for the assisted IoT WSN [3]. Other factors of WSN nodes breakdown are continuously changing of network and wireless signal interruption. Power breakdown becomes a new challenge of the WSN time slot allocation and optimization.

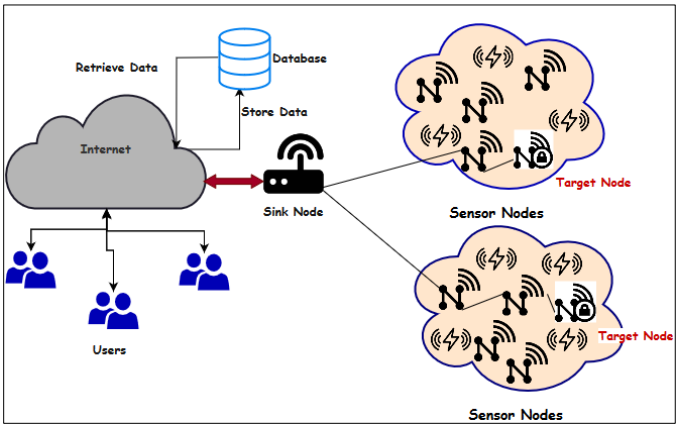


Figure 1: The WSN Architecture

Sensor nodes mobility and unfitting time slot allocation create a complicated network topology [4]. For instance, the dynamic network topology of the IoT networking and unutilized neighbor nodes discovery resulting in transmitting more message than usual. This will cause more energy consumption and reduce network performance. It is known that fast battery draining is among the major limitation of the WSN nodes. Therefore, most of the researchers are working to get better performance of the WSNs focusing on the time slot allocation topological [5 – 18]. However, these algorithms still have some limitations due to the node's distribution factor. Some of the solutions work efficiently for randomly distributed nodes but consumes more accessing time and energy in assigning the neighbor nodes. Another limitation occurs include limited data link capacity and network topology and coverage range. Recently, the network range (e.g., industrial and experimental) becomes broader (e.g., a manufacturing company, rural development monitoring, geographical network etc.). For example, even for smart traffic management system requires long distance network to monitor vehicle movement. On the other hand, every technology moving forward to catch up with high-speed Internets like 3G, 4G and some developed countries already tested the 5G.

The two attentions of this paper are:

- To overcome the range limitation of the mesh topology by analyzing the optimized energy-topology (E-T-DRAND) with star topology that supports long-range network.
- To analyze the optimized algorithm behavior with a minimum of 10 Mbps data link capacity. The reason is, the existing E-T-DRAND algorithm capable to allocate time slots without collision and can implement priority control with minimum message complexity that helps to reduce the energy consumption.

The simulation of this study was implemented on Object TCL (OTCL) environment in NS2.35 platform to evaluate the performance of optimized E-T-DRAND algorithm. The simulation results show improvements in message complexity, running time, number of rounds, and energy consumption. The rest of this paper is as follows. Background of the WSN scheduling algorithms and their related works are being discussed in Section 2. Section 3 explains the analysis mechanism of the proposed optimized algorithm. The performances of the experimental works are being discussed in Section 4. Section 5 concludes this paper.

2. Related Works

2.1. The WSN

The WSN consists of an outsized amount of sensor nodes where the nodes transfer the information from source to sink or base station. The sensor nodes play an important role to get better performance form the network, long service time efficient and low energy consumption. In addition, slot allocation, resources management makes the node energy consumption and performance getting better. The WSN main issue is establishing an energy-efficient routing protocol that gives a significant impact on the sensor lifetime. Several studies were developed to focus on improving the efficiency of the nodes energy utilization.

2.2. DRAND and E-T-DRAND TDMA Slot Scheduling

The TDMA energy-based distribution scheduling topology algorithm in IoT has been proposed by [5]. The mesh topology has been considered where the proposed E-T-DRANT (mainly analyzed the DRAND) being added an extra decision by referring to residual energy among the WSN nodes and the neighbor nodes. Since the time slot for both DRAND and E-T-DRAND is identical, the residual energy information with the priority control algorithm is being combined during allotting the time slot. The difference between DRAND and E-T-DRAND is the neighbor nodes discovery phase.

In E-T-DRAND, when one hop receives the broadcasted message with the energy information, the energy information table for the neighbor is being updated. Therefore, the two-hop neighbor receiving two-hop nodes of energy information. It is helpful for the nodes to manage nodes within two hops. The information of residual energy is being captured and stored in the information table for sending the requested time slot assignment. On the other hand, the time slot allocation control in E-T-DRAND scheduling is implemented by adopting an energy topology factor wherein the DRAND is done randomly. There are four stages for time slot allocation, IDLE, REQUEST, GRANT and RELEASE. The priority control algorithm helps in minimizing the node allocation process by putting some rules (e.g. less residual energy have higher priority). These rules ensure that only one node is applying for the time slot allocation to avoid a collision. The outcomes for both slots scheduling can be present by three complexity analysis, i) TDMA time slots allocation without collision, ii) Maximum and minimum message delays, and iii) Message transmission delay.

The DRAND has been improved in terms of the balancing message complexity, time complexity. Therefore, the previous work improves the efficiency of the TDMA slot scheduling.

2.3. Energy Efficiency in WSN Nodes

Jian Shen *et al.* [6] proposed a protocol named energy-efficient centroid-based routing protocol (EECRP) for WSN-assisted IoT. The work proposed a clustering algorithm that constructs the node residual energy based on node position. The algorithm clusters the nodes to dead nodes and cluster head (CH) nodes and has shown reducing the energy consumption average, especially for long-range communication without affecting the network lifetime. The Distributed TDMA Slot-scheduling (DTSS) algorithm for WSNs has been proposed by Bhatia and Hansdah [7]. The DTSS priority is to perform the scheduling with schedule length restriction requirement to a maximum degree of interference graph and accessing time. In the algorithm, the node needs to know only the intended receiver IDs and its advantage is the neighborhood nodes can take different slots simultaneously. To extend the wireless network lifetime in the IoT network, Hakan *et al.* [8] optimize the energy-aware for the Routing Protocol for Low Power and Lossy Networks (RPL). The work used Contiki OS which enables the emulation of extensive support for IoT protocols and being carried out by the Cooja simulation tool of Contiki OS which enables emulation of a wide variety of embedded hardware.

Nguyen *et al.* [9] work to focus on energy harvesting-aware routing for the WSN with the IoT. The algorithm is able to adapt with a different traffic load from the different application together with the residual energy and the arrival harvesting energy. The authors also proposed a model arrival-harvested energy prediction model at the WSNs nodes that include the stochastic characteristic of the ambient energy sources. In addition, a new parameter called "extra bakeoff" that combined with the energy prediction process to define the route to forward the packet with the cost metric has been introduced. In order to ensure the heuristic WSN service selection and composition with respect to the service classes, Zhou *et al.* [10] coincide sensor node with various sensing functionalities. The cooperative integration and strategy satisfy certain independent concurrent users request that control energy efficiency and minimize the WSN service problem. Compartmental clustering model of an opportunistic signal in WSN for nodes energy efficiency has been proved by [11]. The opportunistic signal of the optimal cluster partition is obtained through the compartmental model rather than the state-of-the-art attenuation models. Thus, the energy efficiency of optimal clustering is utilized from those signals which are carried out from various sources in an indoor environment.

Switching the roles between the sensor nodes to guarantee the WSNs energy consumption has been suggested by [12]. The nodes lifetime and fair evolution of energy level are guaranteed in this work by the positioning of sensor nodes such as a distance between the nodes in the network. Software-defined WSN has been proposed by [13] to reduce the data transmitting overhead between the sensor nodes. The proposed centralized routing algorithm increases the network lifetime in terms of the WSNs nodes consumption. Enabling IoT for underwater WSNs with balanced energy consumption by adaptive routing protocol for the sensor node been proposed by [14]. The underwater WSNs becomes difficult to work due to multifold limitations (e.g., power consumption, path loss, and inadequate bandwidth). To prolong the lifetime of underwater WSNs, a new routing protocol called balanced energy adaptive routing (BEAR) has been proposed.

Zhang *et al.* [15] proposed the energy-harvest wireless sensor network (EH-WSN) algorithm (environment-based harvesting energy) to prolong the network lifetime. However, energy harvesting poses lots of limitation which is difficult to achieve in the real-life. The algorithm adopts a double-stage capacitor structure to ensure node synchronization in situations without energy harvesting and uses an integrator to get ultra-low power measurement. Adjusting the sensor timeslot according to the environment changing to minimize the energy consumption in the TDMA channel accessing method has been proposed by [16]. In this work, all the sensor nodes used a fixed timetable slot for transmitting the data. The work is suitable for long time monitoring since the TDMA cannot be implemented in a scenario where the environment remains constant. If the environment is unchanged, the timeslot size remains the same, which is not the best solution for the WSN network.

A stochastic model in capturing the node predicted energy consumption by a schedulable sensor per cycle of operation has been proposed by [17]. The sensor node minimum and maximum energy consumption per cycle are being computed by the Semi-Markov theory can determine it expected life cycle. The particle-swarm optimization solution being proposed by [18] to emphasize the optimal energy efficiency (EE) problem. The work optimizes the harvesting time and transmits power for non-orthogonal multiple access (NOMA)-based wireless powered WSN (WPSNs) with EE maximization for the wireless energy TDMA. A novel mode for clustering the cluster head (CH) and the cluster formation in the WSN has proposed by [19]. The sensor network divides the nodes into zones (based on residual energy and distance) as per geographic locations. The clustering will be initiated when the CH residual energy become below the threshold value.

The above algorithms discuss and propose the distributed TDMA scheduling algorithms that improve the time slot allocation performance to some intensity by proposing several reliable solutions. Though, the difficulties, high energy consumption, message complexity and collision problem still ongoing which need more attention and further study.

3. Optimized E-T-DRAND Time Slot Scheduling

In this paper, the existing E-T-DRAND algorithm is being re-evaluated and optimized by comparing with existing performance records from the aspect of message complexity, time complexity, number of experiment rounds, and energy utilization. The algorithm is being re-implemented in a randomly distributed network mode, where there is no central node and the algorithm being executed in more than one node at the same time. The neighbor nodes amount becomes the important factor and the node energy being represented as residual energy. The residual energy definition refers to the following "Energy Topology factor":

$$E \text{—} T \rightarrow F\,(Ei + \alpha\, x\, Ti + Ni) \quad (1)$$

Parameter Ei refers to the residual energy of node i while parameter F is the sorting priority algorithm. The coefficient

parameter α refers to the number not being allocated neighbor nodes and $\alpha \in [0, 1]$. *Ti* is the number of neighbor nodes and *Ni* is the neighbor nodes energy information. The optimized algorithm improves the E-T-DRAND algorithm by adding neighbor node energy information among the WSN nodes where the characteristics are similar to the DRAND.

3.1. Neighbor Node Discovery

Fig. 2 shows the pseudo code for neighbor node discovery. The optimized algorithm uses the "HELLO" message containing the node residual energy information. In addition, the message also being used to manage the priority of the nodes time slot allocation. The information was used for updating energy information for the single-hop neighbor. For the two-hop neighbor, it is being discovered from the energy information table of the single-hop neighbor. Thus, when a start receiving the broadcast message, the node can get the energy information inward two hops. The node updates the released information with real-time energy information. During sending a request for time slot assignment, the nodes capture and store the residual energy information, thus, it helps the node to manage and optimize all the nodes within two hops from its own.

```
1. oneHopNeighbor[i]: Array of one hop neighbor nodes
2. twoHopNeighbor[j]: Array of two hop neighbor nodes
3. do {getOneHopNeighbor(i): collect data for one-hop
   neighbor
4. while {message.NdType == FirstHop))
5. do{getTwoHopNeighbor(j): collect data for two-hop
   neighbor}
6. while{ getOneHopNeighbor == true || message.NdType ==
   SecondHop}}}
```

Figure 2: Neighbor discovery pseudo code

3.2. Implementation

The most important entity in the WSN is proper TDMA slot allocation when the network dynamically changes. The improved E-T-DRAN is being implemented as the following:

- The energy topology factor always being scheduled together with the time slot allocation. During the time slot allocation request, at a first time, only one node can send a message. The time slot allocation process has four states namely, i) IDLE, if no request, ii) REQUEST, need to do time slot allocation or discover a new neighbor, iii) GRANT, the neighbor node time slot allocation successful and iv) RELEASE, no acknowledgement from the node.
- The priority control is being implemented together with time slot allocation. If any node has less residual energy, more neighbor node will get high priority. The first distributed priority order is based on the number of neighbor nodes and the "HELLO" message priority broadcast, see Fig. 3. More neighbor means node consume extra energy when the time slot allocation fails. On the contrary, when the node residual energy is lower, all next one-hop nodes neighbor will get higher priority, the additional overhead of these nodes increase energy consumption when the failure of the time slot allocation. If any node decides to allocate a time slot but the nodes in its two hops occupy less residual energy, the node will stop and declining its priority.

```
1. Int Node A;
2. Int Node B; #one-hop neighbor
3. Int Node C; #two-hop neighbor
4. Ticket_number[]: Priority array
5. Value initialization: ticket_number[A] +=neibourCount;
6. do {ticket_number[B]++}
7. while { has_unslotted_one-hop neibor_node B &&
   has_smaller_residual_energy(unslotted_one-hop
   neibor_node B}
8. do {ticket_number[C]++}
9. while { (has_unslotted_two-hop neibor_node C &&
   has_smaller_residual_energy(unslotted_two-hop
   neibor_node C}
10. do {Send a requested time slot packet, and transit
    current state to REQUEST;}
11. while {max(ticket_number[])==ticket_number[A])
```

Figure 3: Priority control pseudo code

4. Experimental Setup and Evaluation

To evaluate the algorithm described in Section 3 is being implemented on the network simulator NS2.35 which is one of the best network simulators for multiple network models. The simulator version 2.35 has been chosen due to its scalable operating environment based on the Object TCL (OTCL) environment. To draw the nodes and the interface measurement NSG2.1.jar file is being used to draw the nodes and the interface measurement. The topology consists of 250 nodes, randomly being distributed on 300 X 300 m planes, see Fig. 4. The network nodes have 50 m broadcast communication range with a link capacity of 10 Mbps. The topology irregularity is extended to the maximum reflects by the randomness of the nodes. The experiment has been carried out 10 times similar to the configuration experiment in [5].

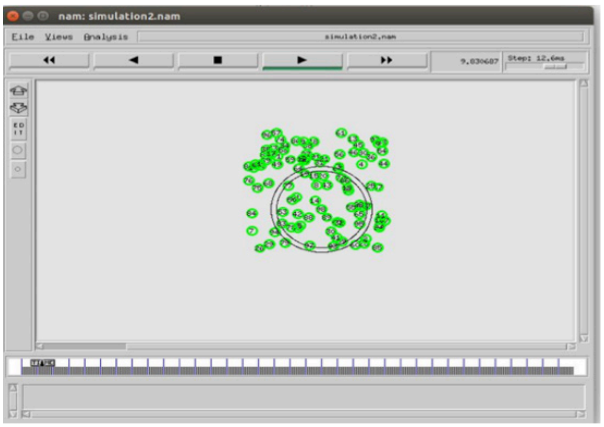


Figure 4: The topology structure of NS2.35 simulation.

The performance schemes of the experiment are evaluated and presented from different aspects that are message complexity, time complexity and energy consumption. The number of neighbor nodes varies from 5 to 55 in the two hops range and the simulation parameters are shown in Table 1. We evaluate the proposed environment and algorithm by comparing with original ET-DRAND algorithms from the aspects of message complexity, time

complexity, number of rounds, energy consumption. As we know the distributed system does not have any central node so different nodes are executing the program and algorithms at a time. The performance evaluation of the experiment presents here through statistical results in figure 4.

Table 1: Performance Metrics

Parameters	Values
Number of nodes	250
Link capacity	10 Mbps
Broadcasting range	50 m
Initial energy	1900 mA
Receiving energy	15 mA
Transmitting energy	20 mA
Idle energy	5 mA
Average delay	1s
Message Selection	Randomly

4.2. Message Complexity

For the experiment, the message selection is randomly and message complexity metric refers to a number of messages being broadcasted over any nodes. The number of messages broadcasted depending on the node neighbor size for the time slot allocation.

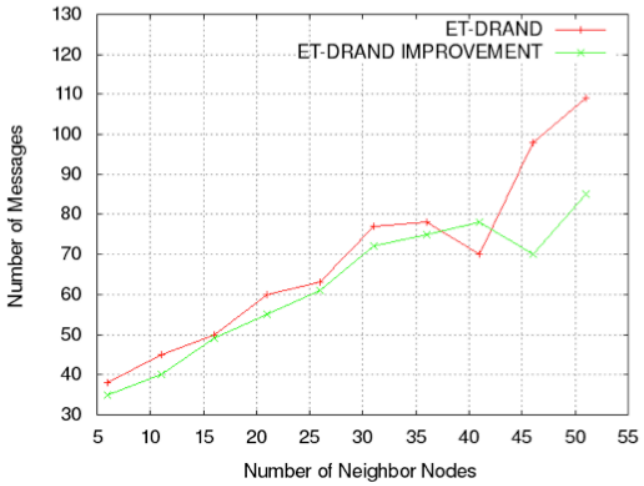


Figure 5: Number of messages in time slot allocation.

The average number of messages being broadcasted is shown in Fig. 5. The figure shows the average number of messages for a node to acquire a time slot in different size of neighbor nodes, where the performance between the existing result and the optimized algorithm is similar when the neighbor modes less than 40. When the number of the neighbor node increases more than 40, then the existing experiment result takes more messages due to a low datal ink rate. The experiment result takes 20.7% less than the previous experiment because when the data link capacity has been increased. This result shows that with high data link rate, the optimized E-T-DRAND algorithm reduces message complexity.

4.3. Running Time Complexity

Fig. 6 presents the average running time of existing ET-DRAND algorithm and the improvement result in the range of the different neighbour nodes successfully allocating the time. When the number of nodes reaches more than 30, the running time keeps increase. As we know, more nodes mean many messages and many collisions. Therefore, when the number of nodes more than 30, the collisions happened regularly due to collision slot request occurs continuously. However, the ET-DRAND algorithm performs better than DRAND algorithm when the number of nodes increases and the amount of message increase. The result confirms the optimized work takes almost the equal running time for time slot allocation and resource allocation with the previous algorithm.

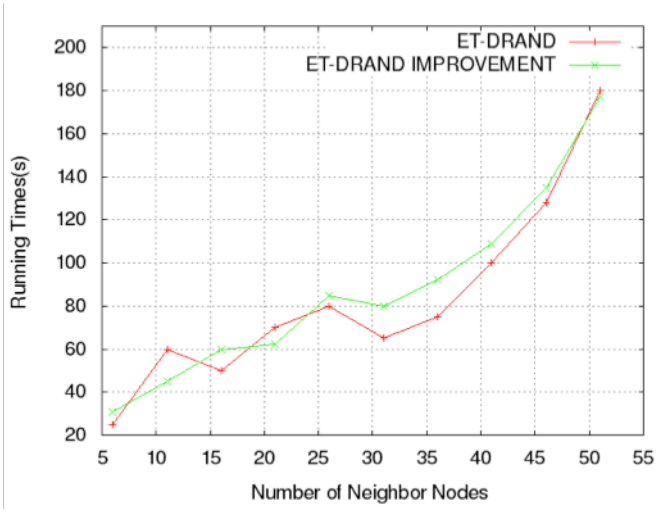


Figure 6: Average running time in time slot allocation.

4.4. Number of Rounds

Note that, this experiment is being carried out using a high-speed data link Internet. Fig. 7 above shows the average number of rounds successfully took to complete the time slot allocation. The running time shows the number of rounds/circle the experiment took to discover all the neighbor nodes. With the number of neighbour nodes increases, it is necessary to allocate time slots after multiple rounds.

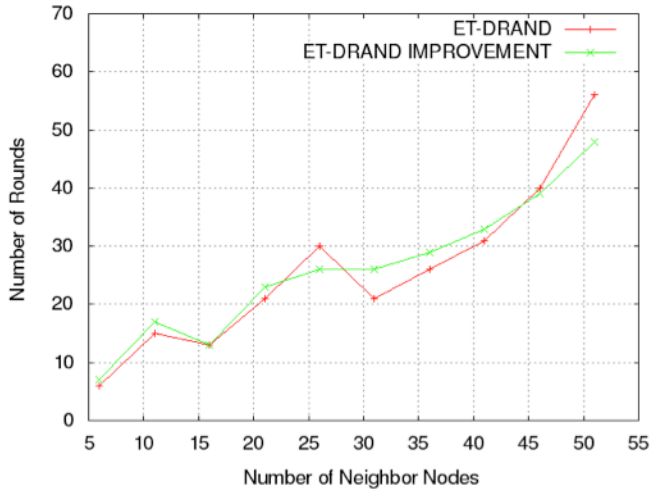


Figure 7: Number of rounds required in time slot allocation.

On the other hand, the running time, the collision increases with the increase of nodes. As mentioned above, the high data link shows an improvement when the neighbor nodes are above 45.

The experiment takes less round in allocating successful time slot and increase the resources allocation.

4.5. Average Energy Consumption

Fig. 8 shows the node average energy consumption for different size of neighbor. The experiment result shows less energy consumption when the neighbor nodes less than 20. On the contrary, the result shows the similarity with the previous experiment. However, when there are more than 40 nodes the nodes consume 5-6% more energy. This is due to the high collision. Since the experiment was done using high data link Internet, so more messages are passed, and more collisions occur. The experiment result proves that E-T-DRAND stills the best solutions in terms of time slot allocation and energy consumption.

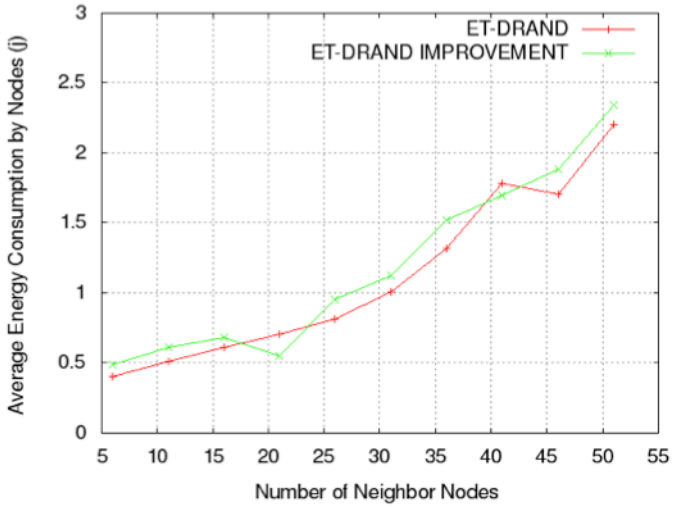


Figure 8: Neighbor nodes average energy consumption.

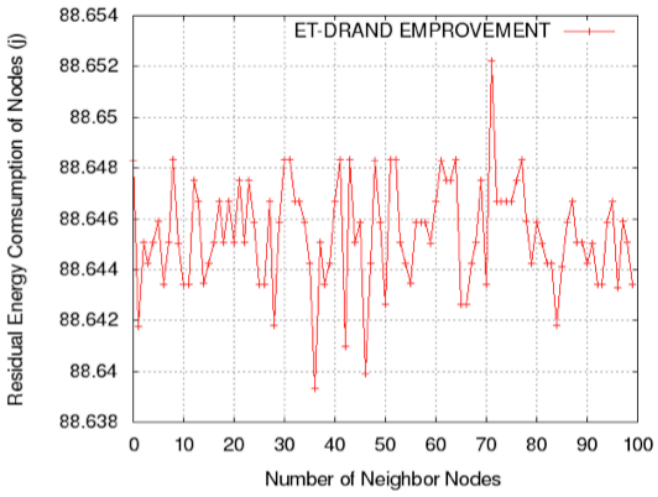


Figure 9: Neighbor nodes residual energy consumption.

4.6. Neighbor nodes residual energy consumption

Fig.9 summarizes the average neighbor nodes energy consumption after the experiment being repeated 10 times. It shows that the nodes group between 30 and 40 consumes less energy than others while the highest energy consumption hold by the neighbor nodes 70. This residual energy is to observe the node lifetime and nodes performance.

5. Conclusion

With the IoT, anything can be controlled and monitored, but for the WSN the power limitation and resources utilization issues make a boundary limit for the network performance improvement. In the distributed network environment, power consumption issues become the main challenges where lacking concentration on time and energy consumption of the WSN nodes makes IoT improvement slowly. In this study, the proper way for time slot allocation by combining with energy topology factor has been investigated. The allocation of the time slot for the MAC protocol in WSN have been analyzed and present the behavior of the algorithm in the mesh topology. The randomly distributed nodes causing the collision would increase the execution time. By updating the node time allocation with energy topology, the energy consumption of neighbor node can be reduced due to the randomness. The energy topology factor of the distributed TDMA slot-scheduling algorithm improves the percentage of nodes performance and reduces message complexity. The world moving for fast Internet like 4G and 5G where the 5G can be one of the best ways to improve communication. In future, implementing E-T-DRAND algorithm with the 5G networks that is our next goal in reducing energy consumption and improves the WSN for the IoT to collect huge data. Since the big data made a huge impact on information technology and business, both the WSN and the IoT indirectly play their rules.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The authors wish to thank the reviewers for their constructive comments and suggestions helped in improving this research paper and Universiti Putra Malaysia for providing the funding in carrying out this study.

References

[1] Y. Yang, M. Zhong, H. Yao, F. Yu, X. Fu, O. Postolache, "Internet of things for smart ports: technologies and challenges" IEEE Instru. Meas. Mag., 21(1), 34 – 43, 2018. https://doi.org/10.1109/MIM.2018.8278808

[2] M. Gautam, V. Sejwar, "A Brief Review of WSN Energy and Routing Protocols" in International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 849 – 855, 2017. https://doi.org/10.1109/I-SMAC.2017.8058299

[3] Z. Ning, Q. Song, Y. Yu, Y. Lv, X. Wang, X. Kong, "Energy-aware cooperative and distributed channel estimation schemes for wireless sensor networks" Int. J. Commun. Syst., 30(5), e3074, 2017. /doi/full/10.1002/dac.3074

[4] V. Gabale, B. Raman, P. Dutta, and S. Kalyanraman, "A classification framework for scheduling algorithms in wireless mesh networks" IEEE Commun. Surveys Tuts., 15(1), 199 – 222, 2013. https://doi.org/10.1109/SURV.2012.022412.00068

[5] Y. Li, X. Zhang, J. Zeng, Y. Wan, F. Ma, "A distributed TDMA scheduling algorithm based on energy-topology factor in internet of things" IEEE Access, 5, 10757 – 10768, 2017. https://doi.org/10.1109/ACCESS.2017.2710304

[6] J. Shen, A. Wang, C. Wang, P. C. Hung, C. Lai, "An efficient centroid-based routing protocol for energy management in WSN-assisted IoT" IEEE Access, 5, 18469 – 18479, 2017. https://doi.org/10.1109/ACCESS.2017.2749606

[7] A. Bhatia, R. C. Hansdah, "A Distributed TDMA Slot Scheduling Algorithm for Spatially Correlated Contention in WSNs" in 27th International Conference on Advanced Information Networking and Applications Workshops (AINA), Barcelona, Spain, 377 – 384, 2013. https://doi.org/10.1109/WAINA.2013.23

[8] H. Erdol, S. Gormus, M. C. Aydogdu, "A Novel Energy Aware Routing Function for Internet of Things Networks" in 10th International Conference on Electrical and Electronics Engineering (ELECO), Bursa, Turkey, 1314 – 1318, 2017.
[9] T. D. Nguyen, J. Y. Khan, D. T. Ngo, "An Effective Energy-Harvesting-Aware Routing Algorithm for WSN-based IoT Applications" in IEEE International Conference on Communications (ICC2017), Paris, France, 2017. https://doi.org/10.1109/ICC.2017.7996888
[10] Z. Zhou, J. Xu, Z. Zhang, F. Lei, W. Fang, "Energy-efficient optimization for concurrent compositions of WSN Services" IEEE Access, 5, 19994 – 20008, 2017. https://doi.org/10.1109/ACCESS.2017.2752756
[11] S. Kumar, "Compartmental modeling of opportunistic signals for energy efficient optimal clustering in WSN", IEEE Commun. Lett., 22(1), 173 – 176, 2018. https://doi.org/10.1109/LCOMM.2017.2763948
[12] S. Hamrioui, P. Lorenz, "ES-WSN: Energy Efficient by Switching between Roles of Nodes in WSNs" in 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 2015. https://doi.org/10.1109/GLOCOM.2015.7417047
[13] F. Junli, W. Yawen, S. Haibin, "An Improved Energy-Efficient Routing Algorithm in Software Define Wireless Sensor Network" in 2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Xiamen, China, 2017. https://doi.org/10.1109/ICSPCC.2017.8242610
[14] N. Javaid, S. Cheema, M. Akbar, N. Alrajeh, M. S. Alabed, N. Guizani, "Balanced energy consumption based adaptive routing for IoT enabling underwater WSNs" IEEE Access, 5, 10040 – 10051, 2017. https://doi.org/10.1109/ACCESS.2017.2706741
[15] Y. Zhang, H. Gao, S. Cheng, J. Li, "An efficient EH-WSN energy management mechanism" Tsinghua Science and Technology, 23(4), 406 – 418, 2018. https://doi.org/10.26599/TST.2018.9010034
[16] A. Biazi, C. Marcon, F. Shubeita, L. Poehls, T. Webber, F. Vargas, "A Dynamic TDMA-based Sleep Scheduling to Minimize WSN Energy Consumption" in 2016 IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC), Mexico City, Mexico, 2016. https://doi.org/10.1109/ICNSC.2016.7478994
[17] V. Agarwal, R. A. Decarlo, L. H. Tsoukalas, "Modeling energy consumption and life time of a wireless sensor node operating on a contention-based MAC protocol" IEEE Sensor, 17(16), 5153 – 5168, 2017. https://doi.org/10.1109/JSEN.2017.2722462
[18] M. Song, M. Zheng, "Energy efficiency optimization for wireless powered sensor networks with nonorthogonal multiple access" IEEE Sensors Letters, 2(1), 1 – 4, 2018. https://doi.org/10.1109/LSENS.2018.2792454
[19] A. Chunawale, S. Sirsikar, "Minimization of Average Energy Consumption to Prolong Life Time of Wireless Sensor Network" in 2014 IEEE Global Conference on Wireless Computing & Networking (GCWCN), Lonavala, India, 2014. https://doi.org/10.1109/GCWCN.2014.7030887

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 38-46 (2019)

www.astesj.com

Special Issue on Recent Advances in Engineering Systems

ASTESJ

ISSN: 2415-6698

Stress Level Classification Using Heart Rate Variability

Tlija Amira[1,2,*], Istrate Dan[1], Badii Atta [3], Gattoufi Said[2,4], Bennani Az-eddine[5,6], Wegrzyn-Wolska Katarzyna[7]

[1]*BMBI UMR 7338, UTC, 60200, Compiegne, France*

[2]*SMART Lab, ISG TUNIS, 2000, Bardo, Tunisia*

[3]*Chair of Secure Pervasive Technologies, Computer Science Department, University of Reading, RG6 6AH, United Kingdom*

[4]*CEMIS-University of Nizwa-Sultanate of Oman, Oman*

[5]*TSH COSTECH, UTC, 60200, Compiegne, France*

[6]*Neoma Business School, 5100, Reims, France*

[7]*Allianstic, EFREI PARIS, 94800, Villejuif, France*

ARTICLE INFO

Article history:
Received: 20 December 2018
Accepted: 08 April, 2019
Online: 09 May, 2019

Keywords:
Heart Rate Variability (HRV),
RR interval,
Wearables sensors,
IoT,
Support vectors machines,
Classification

ABSTRACT

The research programme reported in this paper is set within the framework of our research under the theme of ICT support for Active Healthy Ageing (AHA). This longitudinal empirical research is focused on the study of the impact on the management of cardio-vascular disease if supported by sustained health monitoring using wearable connected devices. One of the key objectives is stress monitoring and its classification during the daily routine of life thus enabling psycho-physiological monitoring to study the correlation between emotional states including variable stress levels and the evolution and prognosis of cardiovascular disease. In this paper, the calibration phase will be studied in order to distinguish between two emotional states: i) meditation and ii) stress condition. For this, the Heart Rate Variability (HRV) features are used as extracted from the RR interval and a support vectors machine (SVM) classifier deployed which resulted in 74% and 87% recognition accuracy based on HRV data for the recognition of the two emotional states, namely meditative, and, stressed, respectively.

The main objective is to prevent Cardio-Vascular Disease (CVD) in healthy people and to treat those who already suffer from it. Creating a reference database was our first step in this research project. The sensor choice was made based on doctors' recommendations. The work methodology was as follows: first validate the « objective data » issuing from the calibration state. Second, set up the automatic algorithm and detect automatically the patient's emotional states during the experimentation period (subjective data). Third analyse the physical activities correlated to the blood pressure and emotions. This study has involved the challenge of distinguishing the influence of stress versus relaxation on the Cardio-Vascular function and in particular on the risk of exacerbation of pre-existing Cardio-Vascular Disease.

* Tlija Amira, Email: tlija.amira@gmail.com

1. Introduction

The Internet of Things (IoT) represents a vision of ambient intelligent devices to support smart environments, lifestyles and hyper-connectivity based on the semantic interoperability of network-centric sensor devices as connected objects [1]. It represents different types of connection (Bluetooth, wireless, etc.) between several sensors within pre-established protocols. The connected objects reach several areas and fields. They can be found in the thermal energy sector, automated manufacturing, smart agriculture, intelligent buildings and healthcare sector. They are linked and connected in order to record, communicate and exchange data in real-time (Figure 1).

Chao Li [2] mentioned that by 2020 the number of connected sensors worldwide would reach 50 billion. It means that in less than 5 years the sensors number have almost doubled (in 2015 it was 25 billion objects). Combining these sensors with computer science tools like big data would increase the French gross domestic product (GDP) by 7% in 2025 [3]

Fig. 1. Data exchange and connected devices (Pixabay)

Hata [4] maintains that our world is mainly composed of three major fields: the environment and its influence on human beings, the technology that enables the sensing and perception of the environment and the physiological and emotional data of the human beings themselves, which in this study in multi-modal HealthCare Monitoring are connected through sensor devices by a complex system-of-systems that includes many operators and exploits the IoT architecture. When operating a healthcare system, the patient's environment must be managed ensuring their security and safety. The number of connected sensors and the necessity to record data in continuous and anonymous way, accurately using a self-synchronised data acquisition method constitute the paradigmatic challenges in this domain. The sensors measure various parameters and provide real-time data streams that need to be processed for data fusion and model building. Hata's health management framework [4] is based on a causality model as is this case study. Sensors measure the psycho-physiological states of the user as well as the relevant states of their living environment and the interaction between them. Emotions and human responses are closely related to hormone balance and the environment (lifestyle, work-style). In order to analyse emotion, the fact that feelings are personally interpreted must be taken into consideration. Some people cannot or will not identify their feelings and even if they do, it is still a subjective assessment. For example, a person might state: "I feel stressed" but physiologically it may that the person is just "tired", or may state "I feel relaxed" while the data indicates that they are stressed. The intercommunications between several areas in addition to the complication of administrating data and decision-making have to be taken into consideration.

2. Cardiovascular diseases

As mentioned in 2017, cardiovascular disease was one of the main causes of death in the world (according to the World Health Organisation). This affects 17 million people all over the world. CVD mainly concerns heart stroke and blood vessels. These diseases are not infections and cannot be passed from person to person. There are different types of CVD, the most common being coronary disease which causes one death in three in the United States every year (source centres for diseases control and prevention). In France, the main cause of death is due to CVD by 33% compared to 5% from digestive system diseases (source INSERM- INSEE). This kind of chronic disease is controlled by risk factors. Some of them cannot be changed like age, gender, family history. However, others can be moderated such as physical activities, obesity, high blood pressure, etc. by reducing those risks with healthier and better life style choices. This is the main objective of this study: to ensure better health care system using information technology. Chest pressure, heart pain, shortness of breath, nausea and feeling heartburn are the main symptoms of heart attack. These aspects are warnings that our body sends us of a dysfunction in our body.

Kernay [5] emphasised the importance of preventing, controlling and treating hypertension diseases, with high priority worldwide especially for developed countries. As figure 2 shows, being in developed or developing countries does not make a significant difference in hypertension risks. Being able to avoid hypertension, coronary disease, and peripheral arterial disease among other CVD is a worldwide healthcare system challenge. The European Cardiovascular Disease report for 2017 [7], shows that despite the age, sex and ethnicity difference between European citizens, CVD still the leading cause of death in Europe.

The aim is to prevent the disease if possible and to manage its treatment using data-analytic tools and connected sensors. Three main factors for CVD have been chosen with doctor's rrecommendations: Physical activities, regulation blood pressure and controlling emotions.

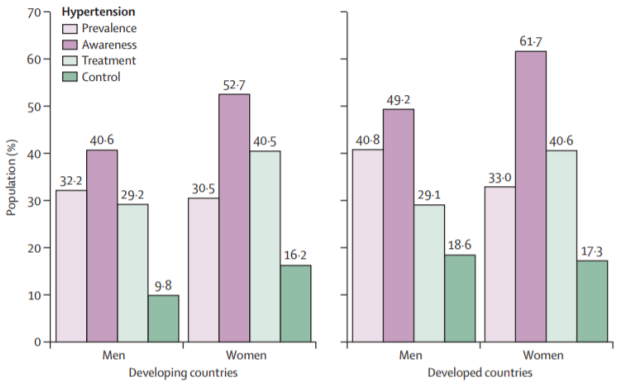


Fig. 2. Developed and developing countries: hypertension statistics (Source [6])

In this paper, a methodology will be presented to describe the main idea and objectives, define the experimentation protocol, explain the sensor choice and show the main results segmentation. However, the most challenging part in this research is controlling emotions for stress. For that, how emotion can be detected is explained based on a complex system of heart rate variability.

3. Heart rate variability

The heart rate variability measures the inter-beat interval. Analysing heart rate variability based on RR interval was the focus of the research. Much research has been conducted to show that there is a link between the emotional state, physiological and physic activities.

The HRV analysis can be divided into three areas: the long term (within 24hrs), the short term (about 5 min) and the ultra-short term (less than 5 min). Some research proposes the ultra-sort term from 1 min to 4 min [8]. Between the time domain, the frequency domain and the nonlinear metrics there are about 50 indicators for HRV. Monitoring people during daily routine life was the focus of this research. The calibration phase duration is about 10 minutes of meditation (calm respiration in order to calculate the heart rate frequency in a resting state) and 15 minutes of stress (stressful games like the Stroop colour test, mental arithmetic, etc.). The RR interval is being recorded with a frequency of one per second. In the literature, the HRV analysis has been analysed for a short-term duration and 24hr measurements. There is a total absent of studies conducted on ultra-short term. The purpose of this research is to bring out several indicators relevant to identifying each emotional state during ultra-short-term testing.

The Autonomic Nervous System (ANS) is the part of the nervous system in our body that regulates the function of human organs and the cardiac muscles and glands. It is the vital system for controlling many phenomes in our body: such as temperature changes sending extra blood to some organs, causing a slower heartbeat, etc. This system is divided in two main areas: the parasympathetic system and the sympathetic system. They serve the same organs but have opposite roles. While one excites our body and supports fight or flight reactions to the environment, the other calms it down. When feelings of anxiety, panic, fear, happiness, etc. are experienced, these are initiated by our sympathetic system thus preparing the body for physical activities and emotional responses. In other words, "stress and emotions" are evoked mainly by the sympathetic system while "relaxation and calm responses" are evoked by the parasympathetic system. The analysis of emotions is the focus of this work on stress based on HRV that was calculated from the RR interval extracted from the heart rate monitor belt. The data recording is 1s for each observation. There is not a unique indicator that can categorise the HRV. This complex system takes into consideration the paradoxical behaviour between the parasympathetic system and the sympathetic system as well as the respiratory mechanism. In much research, the LF/HF ratio (Low Frequency/High Frequency) has been considered reliable in presenting sympathovagal balance. Wilhelm [9] confirms that using this ratio for a two-dimensional representation framework can help in stress recognition. However, using only this indicator is not enough for this research thus analysing daily data including different kinds of emotion, such as fear, panic, happiness, stress, satisfaction, etc. This paper presents the potential benefit of using a Support Vector Machine (SVM) to identify mental stress based on heart rate variability. The complexity of heart rate indicates the flexibility of the Autonomic Nervous System (ANS) and sympathovagal modulation effects on HR variability. Several HRV indicators were taken into consideration from the time and frequency domains to Poincaré and statistical analysis. The objective was to distinguish between the effects of sympathovagal fluctuations and emotion on HR variability.

Emotions and the way they are expressed are different from one person to another. Many studies were conducted into this area, and most of them show that having a positive attitude, surrounded by positive feelings, helps people to live longer compared to those who are always feeling angry, afraid and stressed. In 1977, the American psychologist Caroll Ellis Izard classified emotions into ten categories; joy, surprise, sadness, anger, disgust, contempt, shame, fear, guilt, interest or excitement. As cited, stress is not technically an emotion. It is a feeling that based on a reaction or a stimulus. A recent study in 2018 [10], based on the state-of-the-art of 37 publications showed that heart rate variability is a reliable indicator for stress recognition.

4. The proposed work and other related works

This is not the first work undertaken on stress and HRV as mentioned in the related work in the paper. But the novel contribution is that this work will be the first to separate different kinds of emotion during daily life. No related work has been conducted that mentioned how to distinguish between happiness, being angry, stressed or feeling relaxed. No related research undertook daily monitoring for patient/participants.

There have been no previous studies deploying a similar methodology of longitudinal real-life study of the psycho-physiological correlates of evolution and exacerbation of cardio-vascular conditions. This study was initiated responsive to earlier mainly clinically controlled studies [11], [12], [7] which involved data capture in the context of a few specifically detectable activities (such as preparing food, climbing stairs, using a smartphone, etc.). As of 2018 some work has been mainly concerned with feature extraction [9]; others have used various approaches for stress detection e.g. a survey approach [10], saliva and cortisol level [11]. Others have used controlled studies to evaluate the impact of fast food consumption [12] lifestyle in terms of cardio-vascular disease risk and some have involved data capture in an IOT-enabled controlled smart home context [13]. The approach followed here is distinguished by being based on an ambulant environment (rather than being clinically controlled), connected-devices-enabled continuous data measurement including contemporaneous self-expression recordings of the participant's perceived feelings and mood. The project objective is to extract information and correlations from different types of data (corporal and sentimental) and to bring up analysis that will present the impact of these variables and features on cardiovascular diseases in terms of prevention and treatment. The experiments started with data acquisition involving healthy individuals paving the way for addressing sensor integration, data acquisition, storage and integration issues and evaluating the experimental set-up for the future. This paper explains the project methodology, experiments

already conducted and those planned to be carried out within INSEAD-Sorbonne Université. Firstly, the sensor choice was explained, secondly, the concept and objectives were described and demonstrated and the questionnaires were tested and validated. Finally, the experimentation and analysis tools were defined. sensors currently available on the market can support the measurement of three categories of data regarding the cardio-vascular function; namely: i) the physiological parameters such as activity level, ii) arterial pressure, iii) stress level. The research purpose is to build an automatic system detection capable of recognising different types of emotion and to categorising feelings and mood according to heart beat, arterial pressure among several other variables. Moreover, the work's contribution is to ensure monitoring participants while they are wearing sensors during their daily life. Despite the fact that not all the stress-creating factors are manageable, the conducted experience gave rise to reliable data that illustrate participants' spontaneous feelings and mood. As the experimentation has been conducted over a long period, the part that concerns data storage and management had to be handled securely.

5. Materials and methods

5.1. Wearables sensors

Before launching this research, a literature review was conducted . The idea was to study the types of existing sensors on the market and to the research carried out already [8]. The sensor choice was based on three hypotheses that are explained as follows:

1. Blood Pressure or the blood circulation is highly correlated with cardio vascular diseases[14, 15] and thus identifying the flow of blood circulation with a connected sensor remains an important part to maintain a stable blood pressure. Hastie and Gluszek [16, 17] mentioned that it is essential to measure heart rate variability in order to manage and control the risk of having a cardiovascular disease.
2. Nutritional attitude and physical activities are two important factors that people should take into consideration in order to prevent the apparition of cardiovascular diseases [18].
3. Feeling overwhelmed, anxious, stressed among several other feelings, are factors that can be the main cause for the appearance of chronic diseases. These cognitive and emotional factors, have not yet been measured in an ambulatory approach [13, 19, 20].

To satisfy the hypothesis mentioned above and to set up this experiment, 3 connected objects were selected. They will measure the physical activity, the blood pressure and heart rate variability. So each participant was given a connected tensiometer, a heart rate monitor belt and a smart watch.

Tensiometer

Following a study of user acceptance, reliability and wearability criteria, the tensiometer «Rest» AD761f was selected as one of the sensors to be integrated within the body area sensor network (see Fig.3). Tensiometers of this brand are among the first to have passed the European Society of Hypertension (ESH) protocol [21], [22]. This device enables the detection of a person's rest condition using the Hemodynamic Stability Determination technique (HSD).

Fig. 3. Rossmax tensiometer

The smartwatch ActiGraph

Dinseh [23,24], states that the actigraph is one of the most reliable widely used accelometers in scientific research. This smart watch is an activity monitor (Figure 4) tested, tried and approved by many researchers and practitioners [25, 26] as a reliable sensor giving accurate data. As the trial was over a long period and the objective was to monitor people during their daily life, many limitations were imposed on the sensor choice. It had to be small, discreet, user friendly, giving reliable data. As the sensor placement affects the degree of data accuracy, all participants were asked to carry this connected device in the same position on their waist (Figure 4)

Both raw and calibrated data were extracted at a range of sampling frequencies (ranging from 30 Hz to 100 Hz).

Fig. 4. Actigraph sensor worn

The heart rate monitoring belt: Polar H7

For emotional state monitoring and measurement specifically stress, a state-of-the-art analysis of the emotional states sensing technology was performed as well as a study of the readymade solutions for this which could be mass-scalable. To quantify emotions and feeling many tools and techniques are available specially to measure stress among other feelings. This mental pressure and feeling is one of the main factors that accentuates the appearance of CVD among chronic diseases [19] [20] [55].

Marek and Hamilton [33, 34] specified measuring HRV as a reliable technique for categorisation of several kinds of emotions. For this reason participants were equipped with the heart rate monitor belt to measure HRV based on the RR interval (figure 5)

Fig. 5. The Heart Rate Monitor Belt

5.2. Methodology

The methodology was based on the aim of this work, to treat people with CVD and to prevent the apparition of chronic diseases especially cardiovascular ones. This project methodology was centred on studies that explain the association between the emotional states and heart rate variability [35, 36].

One of the important emotion types to be monitored is stress which can lead to CVD, specifically chronic stress increasing blood pressure and exacerbating CVD; the physiological and psychological reactions to stress can manifest as both anxiety and depression. The stress monitored was real life stress and not induced or simulated for experimental purposes [37]. Vaping and smoking, being overweight or obese, the absence of physical activities, being anxious and stressed are factors that will put healthy participants at risk. The trial duration was 15 days in which participants wore the three sensors according to the user manual. No restriction was imposed. Participants were not asked to change their daily routine nor to be more active. This study is distinguished by the fact that the key psycho-physiological data relevant to cardio-vascular function was captured continuously over a long period to support offline analysis, rather than on a one–off diagnostic basis. This trial was validated in three steps. First, the experts decided the inclusion and the exclusion criteria to select different participants. Second, three online surveys were tested on the UTC team. According to their feedback more questions were added. Third, the procedure was validated and all the experimentation protocol was validated by the INSEAD ethical committee.

5.3. The experimental protocol

This project was established in collaboration with the Institut Europeen d'Administration des Affaires (INSEAD). The experimentation protocol took the welfare of the participants into consideration. The main research goal is to measure some variables via the sensors and devices and to extract features from this data. The main work needs to be done on patients with CVD. Nevertheless, the study with the INSEAD lab was performed on healthy people. A database reference was necessary and so the trial was launched with participants with no chronic diseases who were closest to the sociodemographic criteria of France.

Before launching this experiment, two main exclusion criteria were imposed: having a chronic disease and not reaching the age of majority (18-19 years old). People would wear the sensors every day during the trial. Physiological and emotional parameters will be measured continuously.

Candidates were assigned coded identities, which were transmitted once the trial started. The protocol ID is ID February2018/1. It was approved by the Ethical Committee in INSEAD. All data are anonymously collected and treated.

A quantitative study was performed before starting the recordings. First doctors from different specialities (neurologist, cardiologist and generalist) were questioned. This enabled the categorisation of the three main areas that will affect chronic diseases generally and CVD specifically. Next a focus group was set up with the UTC lab in order make sure that all the questions are clear and easy to understand (as few questions have medical terms). Age, antecedent diseases, physical activity, eating habits and emotion were the main features that doctors proposed to focus on. An emotional eating a validated questionnaire concerning emotion and eating was conducted inspired by Garaulet et al. [38].

The three main questionnaires were implemented on Qualtrics (within INSEAD Sorbonne université). Each participant had to enter his/her own code to express how the study was conducted. The main three surveys are as follow:

1. The pre-selection questionnaire: Selecting people according to the inclusion criteria
2. Before experiment questionnaire: Analysing participants' physical activity, eating habits, the general assessment of the candidate's health (smoking, alcohol, obesity ...).
3. After experimentation questionnaire:the main purpose is to study the candidate's satisfaction with the experience, focus on participant's recommendation, find out if the experiment helped candidate to change his/her dietary habits and becoming an active person.

a) Procedure

The trial duration is 2 weeks; during which candidates must wear three connected objects. The heart rate monitor belt should be worn at least 2-3 hours per day. The tensiometer for morning and evening measurements and the activity tracker all day long.

After selection of each participant, a meeting took place with them in order to explain experimentation's protocol. Participants were informed that we created a forum free place to ask questions and exchanging with each other and giving their feedback about the experience. They were given a user manual re the tools, how and when to wear sensors. They were also given a notepad in which to write down their feelings, physical activities and emotions while they are wearing the sensors.

They were also told that the experiment was not to be regarded as an alternative to any medical or other treatment and informed about the FORUM, a free and anonymous point of exchange of ideas to be used to analyse participants' responses [39], enabling participants to write their feelings anonymously, their issues as well as their remarks and questions. Before starting the experience, candidates had to read and sign the protocol. Once the study is over, they have to bring back the sensors to INSEAD and so getting a financial compensation for their effort during this trial (Figure 6)

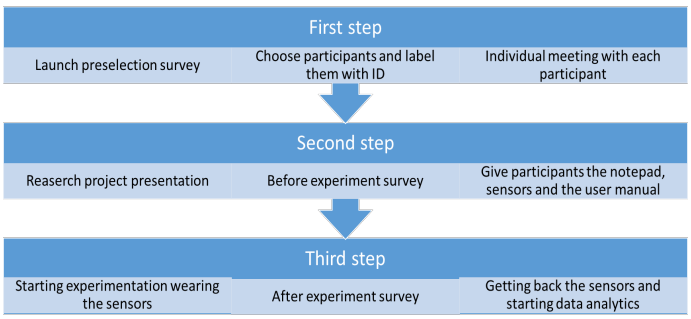


Fig. 6. The description of the experimentation protocol

b) Subjects

Before recruitment began, the intention was to have a population similar to the French population described by the INSEE; 23 men and 22 women. However, recruiting people depends on their flexibility, motivation, availability and eligibility. The trial duration was 4 to 5 months during which 53 participants took part wearing the devices for 15 days each. Candidates were from different socio-professional categories, different levels of study and age category. In this section, some statistical results of the population are presented and participant feedback.

As the graph shows (Figure 7) almost 40% of our population are healthy people with no cardiovascular disease who are aged over 40 years. Recruiting men for this study was not a simple task.

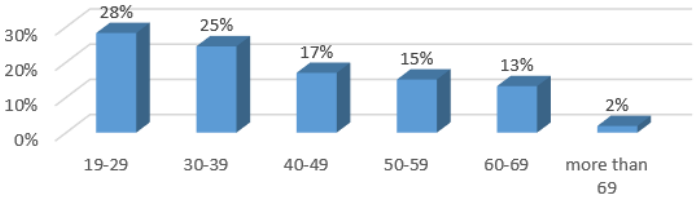


Fig. 7. Age Categories

The final population is composed of ¾ women and ¼ men. Obtaining certain variables during sensor initialisation enabled the comparison of the participants' measurements. The Body Mass Index (BMI) is calculated as follows: BMI = mass / height^2. The interpretation is as follows:

- If BMI <= 18.5 : the participant is in underweight area
- 18.5 <BMI <= 25 : the participant is in a normal weight area
- 25 <BMI <= 30 : the participant is overweight
- 30 <BMI <= 35 : the participant is in a moderate obesity area
- 35 <BMI <= 40 : the participant is in a severe obesity area
- 40 <IMC : the participant is in an obesity morbid area

57% of participants have a normal weight with only 4% having morbid obesity (see Figure 8). 31% is the wear time validation for less than 40% of participants. This shows that participants were motivated to carry out the experiment

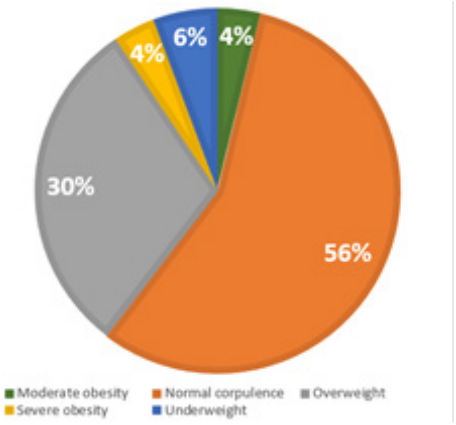


Fig. 8. Obesity Level

6. HRV and calibration phase

Heart rate variability is highly used for emotions analysis like anxiety, fear and specially stress. In fact, the HRV is described by several parameters and biomarkers. There are features that are calculated in the statistical domain, time serial domain and others are in the frequency domain. One powerful biomarkers for stress detection based on HRV analysis is the LF/HF ratio (Low frequency/ High frequency).

The interpretation is based on two hypotheses: the ratio LF/HF must be interpreted. If its value is less than one it shows that the person is feeling relaxed, sleeping or just resting. Otherwise it means that the participant is carrying out a physical activity, running or feeling stressed

The first HRV analysis of recordings on Kubios, showed the confirmation or the two hypotheses mentioned above (Figure 9 and 10)

Fig. 9. Kubios HRV analysis that shows a participant doing a physical activity

Fig. 10. Participant in rest condition (yoga class)

Before starting this experiment with the collaboration with INSEAD Sorbonne Université, a « blind test » was set up with the UTC member lab. The idea was to make sure that the equipment and sensors are connected and that they were friendly user.

While analysing the HRV based on participants, notes (on the notepad) it was deduced that sometimes there is no correlation between what is written on the notepad and the HRV analysis. In fact, emotions are our personal interpretation. For example,

participant A can write on the notepad 'feeling stressed' and while analysing data of participant A we found that he was not stressed. In addition, the fact that people sometimes have difficulty identifying their feelings has to be taken into consideration. It was therefore decided that 'reliable data' was needed which meant recording data while observing what is actually happening. This was called that the 'calibration state'. Games were developed to stress people. Before starting the experimentation, participant had to spend 20 min of recording 'reliable data' starting with 10 min of deep respiration and meditation. After that 10 minutes of playing stressful games that are reported as follows: Starting with 3 minutes of a mathematical mental exercise, next playing the ball game for 3 minutes and finishing with the Stroop Colour Word Test, the cognitive stressful test. These games were developed using Visual Studio and Scratch (Fig. 11). Despite the fact that the game's idea is validated in advanced research [40], it was decided to develop them so they can be user friendly in an interactive way and adjust the speed games, another helpful factor for boosting the participant's stress. Before launching these games, they were trialled on 13 users from the UTC Lab and the graphical interface was adjusted according to their feedback and remarks.

Fig. 11. Interactives games

7. Support Vector machine

To analyse the HRV there are distinct techniques that enabled the extraction of several features. The support vector machine is often used in statistical machine learning theory to solve classification issues in several areas in the medical field. Lanata [41] used this algorithm to identify horses' response to human fear and happiness. Mirhoseini [42] used it to detect early cardiac death while this pattern recognition classifier was also used for health care applications based on Heart Rate Variability [45-50]

The main use of this supervised algorithm is to distinguish between two classes (or more). Having a set of points in a feature space, each point is associated to a label. The dispersion of a group of points with the description is made up of a straight line or a hyperplane (in the context of two dimensions).

Nevertheless, the problem of machine learning only functions if the data has a linear structure, which is not the case here where the SVM with RBF Kernel (Radial Basis Function Kernel) was used to overcome this.

8. Results

Fred Shaffer and Roman Baevsky [43] [44] mentioned all the state of the art of research that has used the ECG signal and RR interval to interpret the HRV. The results shown in this section, concern the calibration phase. An Ultra short-term period (less than 5 min of analysis). Corresponding features were extracted and some statistical variables were added. A feature was created [24] for arrhythmia detection (if the HTI <= 20.42 and RMSSD<= 0.068 then the heart rhythm is normal, and if the HTI> 20.42 then there is a presence of arrhythmia). The 'HRV Tool' implemented by Marcus Vollmer [53] on Matlab was used; in order to extract some features from the RR signal. During this calibration phase 34 participants were chosen and the data divided into two categories: stress label and relaxation label. Work was chosen with a signal processing 'windows of 90s with recovering of 50% [45]. The 24 selected features are described in the table 1.

Table 1. Different features

Domain	Features
Time Domain	SDSD, PNN50, NN50, RMSSD, HRV triangular index (HTI) Mean RR, Standard deviation, coefficient of variation, rrHRV and TINN[44] , [45]
Frequency Domain	VLF = Very Low Frequency (0.0033-0.04 Hz), LF= Low Frequency (0.04-0.15 Hz), HF = High Frequency (0.15-0.4 Hz), pLF, pHF and LF/HF ratio[46],[47]
Statistics analysis	KURTO, Skeww, Interquartile, Arrythmia (that detect the arrhythmia corresponding to [44])
Poincare plot	SD1, SD2 and SD1SD2ratio, approximate entropy (ApEn)[49] ,[50]

The data was divided into two datasets: 80% for training data and 20% for test data. We made a k fold (k= 5) so the final obtained value is an average of the 5 iteration inside the SVM model. The table below shows the main results with 74.44% of good recognition rate:

Table 2. Confusion Matrix

	Class 1	Class 2	% good recognition rate
Class 1	48	38	55.81 %
Class 2	17	115	87.12 %

9. Discussion and conclusion

Our study mainly concerned monitoring patients with chronic diseases at home and controlling heart rate variability with wearable sensors. The experimental protocol was conducted with 50 participants who were equipped with 3 connected objects which. they had to wear for 15 days and to record, on the notepad provided, what they were doing (activity level) and how they were feeling throughout the day. Before launching this experiment and in order to ensure the accuracy of the acquired data, each participant had to undergo a « calibration phase ». In this paper the

methodology and the results obtained from analysing the calibrated data are presented. The SVM on the RR measurements were applied to interpret the heart rate variability based on the activity and stress levels during any period of observation.

34 RR interval signals, including two different beat types (stress and meditation) were selected. 24 features were selected and the support vector machine algorithm was applied. The HRV signal extraction from each RR interval showed over 74% of accurate recognition. The optimisation of the use of features and thus the reduction of the error rate will be the aim of future work. The algorithm has a recognition accuracy of over 87% for class 2 that indicates the stress induced from playing games. The population was divided by age and gender in order to obtain a more efficient classification [54].

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgments

This work was funded within the framework of E-BIOMED Chair IUIS (Institut Universitaire d'Ingénierie en Santé). Experiment will be conducted in collaboration with the Centre Multidisciplinaire des Sciences Comportementales Sorbonne Universités–Institut Européen d'Administration des Affaires (INSEAD).

References

[1] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," Ad Hoc Networks, vol. 10, no. 7, pp. 1497–1516, 2012.

[2] C. Li, X. Hu, and L. Zhang, "The IoT-based heart disease monitoring system for pervasive healthcare service," Procedia Comput. Sci., vol. 112, pp. 2328–2334, 2017.

[3] I. Montaigne, "Big data et objets connectés Faire de la France un champion," Rapport, p. +200, 2015.

[4] Y. Hata, S. Kobashi, and H. Nakajima, "Human Health Care System of Systems," IEEE Syst. J., vol. 3, no. 2, pp. 231–238, 2009.

[5] Y. G. Smeyers and M. Villa, "A study of the internal dynamics of inverting trimethylamine by means of the non-rigid group theory," Mol. Phys., vol. 101, no. 12, pp. 1795–1803, 2003.

[6] K. B. Tibazarwa and A. A. Damasceno, "Hypertension in developing countries," Can. J. Cardiol., vol. 30, no. 5, pp. 527–533, 2014.

[7] Wilkins E. et al., "European Cardiovascular Disease Statistics 2017 edition," Eur. Hear. Network, Brussels, p. 192, 2017.

[8] W. von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, "Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV," *Front. Physiol.*, vol. 8, no. JUN, pp. 1–12, 2017.

[9] H. G. Kim, E. J. Cheon, D. S. Bai, Y. H. Lee, and B. H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," Psychiatry Investig., vol. 15, no. 3, pp. 235–245, 2018.

[10] C. Zhu et W. Sheng, « Multi-sensor fusion for human daily activity recognition in robot-assisted living », in Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, 2009, p. 303–304.

[11] L. Bao et S. S. Intille, « Activity recognition from user-annotated acceleration data », in Pervasive computing, Springer, 2004, p. 1–17.

[12] A. J. Brush, A. K. Karlson, J. Scott, R. Sarin, A. Jacobs, B. Bond, O. Murillo, G. Hunt, M. Sinclair, K. Hammil, et others, « User experiences with activity-based navigation on mobile devices », in Proceedings of the 12th international conference on Human computer interaction with mobile devices and services, 2010, p. 73–82.

[13] D. Huysmans et al., "Unsupervised Learning for Mental Stress Detection - Exploration of Self-organizing Maps," Proc. 11th Int. Jt. Conf. Biomed. Eng. Syst. Technol., no. January, pp. 26–35, 2018.

[14] Z. Jinbiao, L. Mao, and S. Oceanic, "- 8. 6," vol. 10, no. 4, pp. 609–611, 1991.

[15] C. Höcht, "Blood Pressure Variability: Prognostic Value and Therapeutic Implications," ISRN Hypertens., vol. 2013, p. 16, 2013.

[16] C. E. Hastie, P. Jeemon, H. Coleman, L. McCallum, R. Patel, J. Dawson, W. Sloan, P. Meredith, G. C. Jones, S. Muir, M. Walters, A. F. Dominiczak, D. Morrison, G. T. McInnes, and S. Padmanabhan, "Long-term and ultra long-term blood pressure variability during follow-up and mortality in 14,522 patients with hypertension.," Hypertension, vol. 62, no. 4, pp. 698–705, 2013.

[17] J. Głuszek and T. Kosicka, "Visit-to-visit blood pressure variability and cardiovascular and kidney diseases," Nadcisnienie Tetnicze, vol. 20, no. 1, pp. 26–31, 2016.

[18] M. et al. Poelman, "Relations between the residential fast-food environment and the individual risk of cardiovascular diseases in the Netherlands: a nationwide follow-up study," Eur. J. Prev. Cardiol., no. April, 2018.

[19] P. Krolak-Salmon, "La reconnaissance des émotions dans les maladies neurodégénératives," Rev. Med. Interne, vol. 32, no. 12, pp. 721–723, 2011.

[20] S. Montel and C. Bungener, "Les troubles de l'humeur et des émotions dans la sclérose en plaques : une revue de la littérature," Rev. Neurol. (Paris)., vol. 163, no. 1, pp. 27–37, 2007.

[21] Tasker F, de Greff A, Liu B, Shennan AH "Validation of a non-mercury ascultatory device according to the European Society of Hypertension protocol: Rosssmax Mandaus II. Blood Press. Monit".14, pp.121–4, 2009.

[22] Zhang, L., Kang, Y. Y., Zeng, W. F., Li, Y., & Wang, J. G, "Validation of the Rossmax CF175 upper-arm blood pressure monitor for home blood pressure monitoring according to the European Society of Hypertension International Protocol revision 2010", Blood pressure monitoring. Rev. 20(2), pp. 104-107, 2015.

[23] P. Under and T. H. E. Hood, "NIH Public Access," vol. 44, pp. 1–6, 2013.

[24] A. Manuscript, A. Vertical, and A. Counts, "NIH Public Access," vol. 33, no. 2, pp. 219–230, 2013.

[25] Coker-Bolt, P., Downey, R. J., Connolly, J., Hoover, R., Shelton, D., & Seo, N. J. "Exploring the feasibility and use of accelerometers before, during, and after a camp-based CIMT program for children with cerebral palsy". Journal of pediatric rehabilitation medicine, 10(1), pp. 27-36, 2017.

[26] S. R. Dandu, M. M. Engelhard, A. Qureshi, J. Gong, J. C. Lach, M. Brandt-Pearce, and M. D. Goldman, "Understanding the Physiological Significance of Four Inertial Gait Features in Multiple Sclerosis," IEEE J. Biomed. Heal. Informatics, vol. 22, no. 1, pp. 40–46, 2018.

[27] Williams, R. A., Hagerty, B. M., & Brooks, G "Trier Social Stress Test: A method for use in nursing research. Nursing research", 53(4), 277-280, 2004.

[28] R. Schollmeier and T. Universitat, "A Definition of," vol. 15, no. 1, pp. 101–102, 2002.

[29] M. Gröschi, "Données actuelle sur l'analyse hormonale salivaire. Current status of salivary hormone analysis," Ann Biol Clin, vol. 67, no. 5, pp. 493–504, 2009.

[30] D. Nishimura, "Graphically speaking," Science (80-.)., vol. 283, no. 5405, p. 1134, 1999.

[31] M. R. Ebben, V. Kurbatov, and C. P. Pollak, "Moderating laboratory adaptation with the use of a heart-rate variability biofeedback device (StressEraser®)," Appl. Psychophysiol. Biofeedback, vol. 34, no. 4, pp. 245–249, 2009.

[32] M. E. G. H. O'Donnell C., "Detection and management," EmbaseTransfusion Med., no. October 2017, pp. 64–69.

[33] M. Marek, "Guidelines Heart rate variability," Eur. Heart J., vol. 17, no. March, pp. 354–381, 1996.

[34] J. L. Hamilton and L. B. Alloy, "Atypical reactivity of heart rate variability to stress and depression across development: Systematic Review of the literature and directions for future research," Clin. Psychol. Rev., vol. 50, pp. 67–79, 2016.

[35] M. R. Gunnar, D. Ph, A. Herrera, C. E. Hostinar, and B. Sc, "Stress et développement précoce du cerveau," pp. 1–8, 2009.

[36] J. Ham, D. Cho, J. Oh, and B. Lee, "Discrimination of multiple stress levels in virtual reality environments using heart rate variability," 2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., pp. 3989–3992, 2017.

[37] M. Buckert, J. Oechssler, and C. Schwieren, "Imitation under stress," J. Econ. Behav. Organ., vol. 139, pp. 252–266, 2017.

[38] M. Garaulet, M. Canteras, E. Morales, G. López-Guimera, D. Sánchez-Carracedo, and M. D. Corbalán-Tutau, "Validación de un cuestionario de comedores emocionales para uso en casos de obesidad; cuestionario de comedor emocional (CCE)," Nutr. Hosp., vol. 27, no. 2, pp. 645–651, 2012.

[39] B. Liu, "Sentiment Analysis and Opinion Mining," no. May, pp. 1–108, 2012.

[40] D. Huysmans, E. Smets, W. De Raedt, C. Van Hoof, K. Bogaerts, I. Van Diest, and D. Helic, "Unsupervised Learning for Mental Stress Detection - Exploration of Self-organizing Maps," Proc. 11th Int. Jt. Conf. Biomed. Eng. Syst. Technol., no. January, pp. 26–35, 2018.
[41] Lanata, A., Nardelli, M., Valenza, G., Baragli, P., D'Aniello, B., Alterisio, A., ... & Scilingo, E. P. (2018, July). A Case for the Interspecies Transfer of Emotions: A Preliminary Investigation on How Humans Odors Modify Reactions of the Autonomic Nervous System in Horses. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 522-525). IEEE.
[42] Mirhoseini, S. R., JahedMotlagh, M. R., & Pooyan, M. (2016). Improve accuracy of early detection sudden cardiac deaths (SCD) using decision forest and SVM. In International Conference on Robotics and Artificial Intelligence (ICRAI), USA Google Scholar.
[43] Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Frontiers in public health, 5, 258(over view of heart rate)
[44] Jovic, A., & Bogunovic, N. (2011). Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. Artificial intelligence in medicine, 51(3), 175-186.
[45] Guo, H. W., Huang, Y. S., Lin, C. H., Chien, J. C., Haraikawa, K., & Shieh, J. S. (2016, October). Heart Rate Variability Signal Features for Emotion Recognition by Using Principal Component Analysis and Support Vectors Machine. In 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 274-277). IEEE
[46] Salahuddin, L., Cho, J., Jeong, M. G., & Kim, D. (2007, August). Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE (pp. 4656-4659). IEEE.
[47] Park, D., Lee, M., Park, S., Seong, J. K., & Youn, I. (2018). Determination of Optimal Heart Rate Variability Features Based on SVM-Recursive Feature Elimination for Cumulative Stress Monitoring Using ECG Sensor. Sensors, 18(7), 2387.
[48] Venkatesan, C., Karthigaikumar, P., Paul, A., Satheeskumaran, S., & Kumar, R. (2018). ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications. IEEE Access, 6, 9767-9773.
[49] Ashtiyani, M., Lavasani, S. N., Alvar, A. A., & Deevband, M. R. (2018). Heart Rate Variability Classification using Support Vector Machine and Genetic Algorithm. Journal of Biomedical Physics and Engineering, 8(4 Dec).
[50] Lyle, J. V., Charlton, P. H., Bonet-Luz, E., Chaffey, G., Christie, M., Nandi, M., & Aston, P. J. (2017). Beyond HRV: analysis of ECG signals using attractor reconstruction. Computing, 44, 1.
[51] Lee, H. G., Noh, K. Y., & Ryu, K. H. (2008, May). A data mining approach for coronary heart disease prediction using HRV features and carotid arterial wall thickness. In BioMedical Engineering and Informatics, 2008. BMEI 2008. International Conference on (Vol. 1, pp. 200-206). IEEE.
[52] Baevsky, R. M., & Chernikova, A. G. (2017). Heart rate variability analysis: physiological foundations and main methods. Cardiometry, (10).
[53] Vollmer, M. (2015, September). A robust, simple and reliable measure of heart rate variability using relative RR intervals. In Computing in Cardiology Conference (CinC), 2015 (pp. 609-612). IEEE.
[54] BLK Antelmi, I., De Paula, R. S., Shinzato, A. R., Peres, C. A., Mansur, A. J., & Grupi, C. J. (2004). Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. The American journal of cardiology, 93(3), 381-385.
[55] Amira, T., Dan, I., Az-eddine, B., Ngo, H. H., Said, G., & Katarzyna, W. W. (2018, June). Monitoring chronic disease at home using connected devices. In *2018 13th Annual Conference on System of Systems Engineering (SoSE)* (pp. 400-407). IEEE.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 47-52 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Walker Ergonomic Design for Visually Impaired Elderly

Fang-Lin Chao*

Chaoyang University of Technology, Department of Industrial Design, Taichung, 436, Taiwan R.O.C.

ARTICLE INFO

Article history:
Received: 30 March 2019
Accepted: 28 April, 2019
Online: 13 May, 2019

Keywords:
Design
Walker assistants
Pressure
Ergonomic

ABSTRACT

The study proposed a self-driven guidance walker through user-oriented integration design. The common two-wheel drive performs left and right turn. By reducing the weight of the walker body, the durability of the battery can be increased. Secondly, we use a foldable structure so that the walker can be conveniently stored in a small space. A small number of sensors are required to reduce the complexity of signal processing. The front sensors of the walker are infra-red and ultra-sonic detectors, which determine and evade obstacles. The walker handle has three to five force sensors. The combination of the pressure data from these sensors can be used to determine the intent of the user, like turning or stopping.

1. Introduction

With the significant improvements in health science; the population of the senior citizens over 65 years old will be 15% by the year 2025 in Taiwan. Visually impaired persons often encounter difficulties in living an independent life. They need assistive devices to reduce visual obstacles.

Most of the seniors prefer to stay at home. When visually impaired people are unable to complete activities in their lives, they cannot participate in social activities. The assistance is vital for them to participate in ordinary life. Postmenopausal changes decrease the ability for visually impaired women to maintain postural stability [1]. Elderly blind persons had a slower walking speed, shorter stride length and longer time of gait [2].

In [3], the author addresses the motivation of elders to initiate exercise. The elderly are less active than the general population. Fear of injury is a common concern. Therefore, an exercise environment and regimen must be safe. Gradual exercise progression, increasing duration and frequency is significant. People should have the ability to exercise independently [4]. They should be able to maintain a sense of safety throughout the exercise.

1.1. Smart environment

Wireless technology evolved in recent years, designers can bring it to people with accessibility and ergonomic thinking. Many products found are based on a technology perspective and are less conceived from a user perspective. Article [5] suggests a way of deploying ubiquitous computing to cope with these difficulties. A friendly environment reveals itself by talking to the user with a ubiquitous module prototype.

Smart cane [6] offers voice and vibration alert. There is an antenna at the front of the guide stick to sense the RFID tag in the room. The guide wand transmits the signal to the blind person. Sometimes the contents of the tag cannot be read for a fast moving stick.

1.2. Active assistants

Active assistive walker usually has navigation ability. CMU developed robotic assistant in 2002 which lead the users to the destination. An active assistant gather environmental information, achieve obstacle avoidance with sensors. More sensors could increase the processing load and system cost.

An extended Kalman filter is applied to mobile robot navigation in a public environment which relies on geometric beacon [7]. Leonard et al. [8] used the geometric beacons for mapping and localization. The depth camera's data provides indoor localization. Received signal strength (RSS) constructs the radio maps through different data sets and the environment's coordinates [9].

Another study of Tsukiyama [10] proposed a monitoring system of solitary elderly based on urination, kitchen work, and physical cleanliness. The system monitors health status by using water flow sensors attached to faucets. Semi-autonomous control provided simple access through a handheld device. In case of problems, the robot contacts informal caregivers using teleoperation [11]. Another study [12] predicts the possible

* Fang Lin CHAO, Email: flin@cyut.edu.tw

movement according to the output of the force sensor, the direction of motion conveyed by varying the center of mass elevation. A semi-passive walking robot designed using antagonistic actuation [13] with two pairs of legs, parallel links, and one torso. The robot could walk on horizontal surfaces while hung on the bar to prevent falling. The toe length is sensitive with lateral distance, and gait pattern of robot locomotion [14]. Gait function can be used as reference data in the control process. A passive exoskeleton developed with a clutch-spring mechanism and calf muscles. The engaging and disengaging timings were optimized through simulation for extending knee joint application [15].

The active assistants face the challenge of a trade-off between computing power and weight. The pursuit of the function increases the power consumption. Excessive functionality also makes the user unable to learn or unable to afford it. Heavy parts and large battery are also challenges for this to be implemented in living situations.

2. Concept design and prototyping

We proposed the concept design of an active walker with less weight [16, 17] to fulfill the need of the elderly. Smartphones have remote access and communication ability through the USB or Bluetooth connection.

2.1. Design process

Many of the previous designs focused on special features. From the user interviews, we found that such a model is large in size and complicated to construct, which makes it difficult to use. Complex functions also result in increased power consumption and battery drain. Here, conceptual design proposed from the perspective of human factors with an appropriate function, size, and weight.
Figure 1 shows the design process, and user reflection was collect with ergonomic consideration. The design process set a priority of design goal and tradeoffs of product concepts to confirm feasibly.

The digital model creates a prototype 3D printing which utilized in user evaluation. Electrical parts and interface module are placed with mechanical elements. An existing smartphone app has a built-in position sensor and voice notice. The measured data and user's feedback results are collected to adjust the pressure sensing parameters to acquire better control.

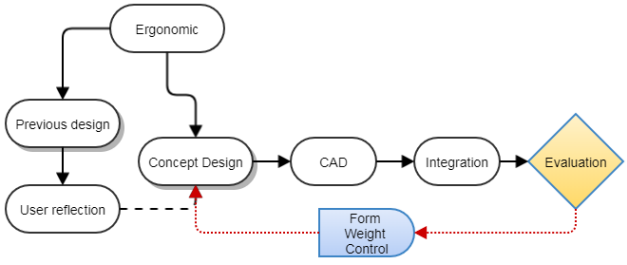


Figure 1: The design process of the people-oriented electromechanical integration design

As shown in figure 1, the ergonomic consideration ensures user's requirements and functional integration. Evaluation test confirms the product features through a series of measurements.

2.2. Electrical components and sensors

The electronic parts and mechanical components are combined to achieve appropriate functions through system integration. The electrical components primarily follow sensing and control circuits.

The walker's body has sensor units placed in front to determine and evade obstacles. The block diagram (Figure 2) shows the existing technologies such as motor controllers and sensors, which have widely been used in different fields. We proposed a link between pressure sensing and user intent.

2.3. Mechanical parts and chassis

The structural and mechanism components are configured with CAD system. The walker's handle (Figure 2) has pressure monitor sensors. The handle sensors located corresponding to palm and finger contact areas. The pressure gathered from the distribution of charges with thin-film sensors. The pressure distribution helps to verify the current status of the user and adjust the walker's speed (Figure 3).
The relationship between the visually impaired and the walker is like the way people lead them.

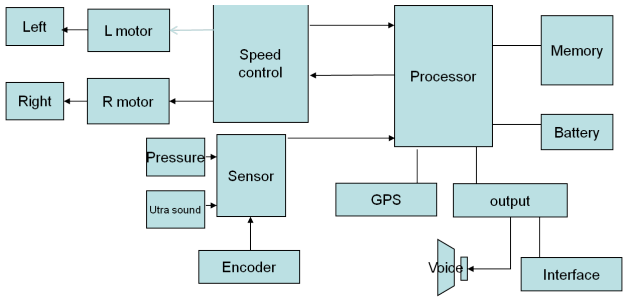


Figure 2: Functional block diagram using a smartphone to control, some software builds in the mobile app. via the Bluetooth interface.

The user configures the parameters of the walker to fit an individual's need. Typically, the handle has three to five sensors. The combination of the data from these sensors can determine the intent of the user (turning or stopping).

Figure 3: Perspective view of walker leads the visually impaired in the moving target.

3. Ergonomic design

The ergonomic design follows some rules; detail concerns: 1) the adjustable handle accommodates users of different heights within weight limit; 2) easy control and reduces the applied force; 3) fashionable style that is apart from handicap stereotypes.

3.1. Overall image

Primary purpose of the assisted walker is to take a walk with the visually impaired. Through the pre-set path and speed, the assisted walker allows the elderly to be involved in minor exercises in daily life, hence promoting a healthier lifestyle. The image of handicap is subjective to the acceptance of users. They are sensitive to negatively perceived in a social context. The form of the assisted walker is more aesthetically pleasing to customers, which may remove some of their mental pressures in purchasing and using the walker.

3.2. Weight and dimension

In [16] participants performed different lifting tasks, with three lifting heights. The working heart rate of the participants showed a significant difference during lifts. One of the design targets is to make it less heavy so that an individual can lift it comfortably. Put the user's needs to the forefront, the weight of the walker was reduced to avoid injury and easy to carry. So we need to choose to remove unnecessary elements and function.

The weight of the car is 23Kg referring to the requirements of airline baggage which is a reasonable load for user. The parts are selected and configured under the premise of this weight limit.

3.3. Way of control

Safety is an important consideration for the elderly. To avoid the sudden rotation caused by the interference of walking, we set the stop button near the thumb. When the elderly consciously can't follow or is too fast, they can manually stop the helper to prevent potential danger.
A user may choose the assistive mode and listen to the related information. Seniors face usability issues related to physical, mental, and cognitive impairments. Most of the elderly have limited experience with technology.

Observing the guidance for blind people is helpful for walker design. Usually, people follow the below steps to activate by a sighted person.

Offer an elbow and back up to the person's outstretched hand. Walk at a comfortable pace by remaining a slightly behind.

Keep arms bent at a 90-degree angle. The visually impaired person will remain a half step behind as you begin walking.

Stop when you need to turn around and tell your follower. He or she will release your arm and wait until you offer it again.

Imagine that when we assist the elderly in crossing the road, we also support him or her on one side; so this is a more natural form of assistance. We choose one-handed operation; the prerequisite is that the elderly can still walk normally. It also avoids interference with the machine while holding both hands.

4. Evaluations

The measured pressure determines the user's intention. The implement procedures are described in Figure 4.

4.1. Moldable clay

Moldable clay was used to demonstrate deformation during initial test. The handle was wrapped with clay evenly. The deformation of clay is related to the pressure distribution during usage.

User's operations would be recorded in the change of shape to determine possible positions of sensors.

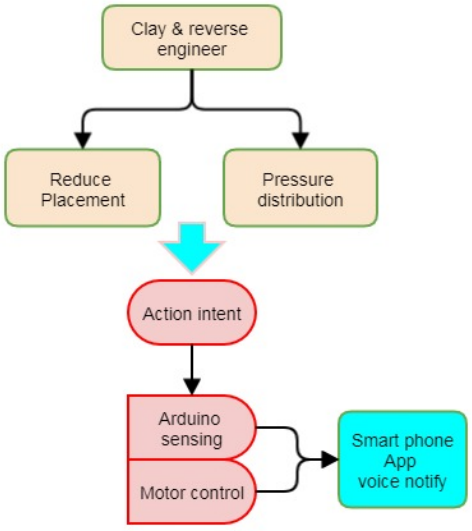


Figure 4: The ergonomic study and implement procedures

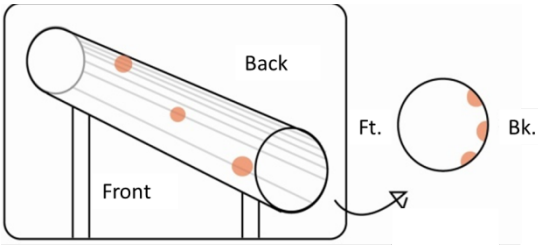


Figure 5: The side view of the pressure detector located on the handle (red dots are high-pressure area). The leading edge (Front or Ft.) is the direction of movement, and the trailing edge (Bk.) is the side of user.

There is less variation among the elderly in terms of palm-size. We choose the positions of sensors according to the magnitude of geometrical change (Figure 5).
When the old man is led forward, he may be towed and fall if the walker runs fast. If the walker turns quickly, the users also have challenges to follow. The sensors acquired the intent and conditions of the elder; so that the robot can adjust its speed.

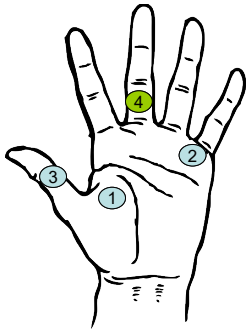


Figure 6: Sensor contact location

4.2. Pressure distribution

As shown in Figure 6, a user holds the handle during a walk. We adjusted the speed and direction of movement based on the pressure distribution. The purpose is to understand the appropriate location to detect the action intently. In this experiment, the driver

did not have a built-in processor and measurement device, so an external sensor and wire is employed to transmit directly to the host PC for subsequent design.

Figure 7: Measurement of pressure distribution.

We will start with five sensors and then re-evaluate that of fewer sensors (3 or 4) in a suitable position based on the pressure distribution observed from the clay.

Five sensors

We collect walking data of five sensors with five participants (Figure 7) to identify the pressure distribution. The measurements is based on the following situations.

- Condition 1： Right-hand control, go forward and intend to turn left,
- Condition 2： Right-hand control, go ahead and expect to turn right,
- Condition 3： Right-hand control, go forward and plan to stop.

The measured results are listed in Table 1-3.

Table 1: Pressure of left turn

User	Thumb	Index	Middle	Ring	Little
A	Low	High	Medium	Low	Low
B	Medium	High	Low	Low	Low
C	Low	High	Low	Low	Low
D	Low	Low	Low	Low	Low
E	Low	Low	Low	Low	Low

We find some variation in Table 1-3, such as:

- Table 1: the index finger of participant A.B.C is different from that of participant D and E.
- Table 2: the little finger pressure of participant E is different from that of others.
- Table 3: the index finger pressure of participant B is different from that of others.

Table 2: Pressure of right turn

User	Thumb	Index	Middle	Ring	Little
A	High	Low	Low	Medium	High
B	Medium	Medium	Low	Low	High
C	High	Low	Low	Low	High
D	High	Low	Low	Low	High
E	High	Low	Low	Low	Low

Table 3: Pressure of sudden stop

User	Thumb	Index	Middle	Ring	Little
A	Low	Medium	Medium	Medium	Medium
B	High	Low	Low	Low	Low
C	Low	High	High	High	High
D	Low	High	High	High	High
E	Low	High	High	High	High

Discussion

Depending on the size of the participant's hand, some users apply pressures by the palm, which made the finger pressure to be less apparent. This variation causes a limitation of using pressure data to determine the user's intent.

The way to overcome this difference is to reduce the unnecessary sensors and to tailor the handle's shape according to the size of the user's hand. Just as people buy a shoe with a specific size, the handle's shape needs to fit the user's hand.

The personal handle can be printed in a 3D facility at a reasonable cost — the specific handle shape made according to the clay model, and adjusted 3D model by reverse engineering. Afterwards, the sensors are then placed in the correct position of the handle.

Three / four sensors

The price of the sensor is high, so we hope to know the possible placement position with a limited number of sensors. Three and four sets of sensors placed to record the user's main motion intent sequentially. The user holds the handle and records the pressure at each position. Based on observations, we infer the position of the sensor that is suitable for the user.

The measurement of three/four sets of sensor placement summarized in (Figure 8, 9). The position for the three sets of sensors for different users is not exactly the same, which is due to the difference in the size of the palm and personal holding habits.

According to the three intentions we want to distinguish (scenario: left turn, right turn, and stop), the test results of the three sensors (Figure 8). We found that the three values at sensor 1 are identical; there is a significant difference in sensor 2; the right turn is similar to the stop at sensor 3. Although the sensor 2 is able to

resolve the three intents, in order to improve the reliability of the results, we continue to try four sensors.

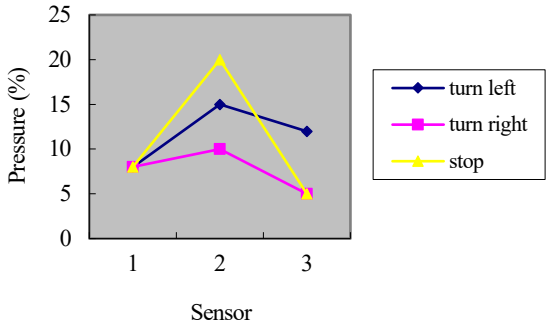


Figure 8: Pressure test results when using three sensors. The three lines represent the values of the pressures presented at each sensor when three actions are performed (left turn, right turn, and stop). The pressure sensing distribution is different when performing three actions.

The fourth pressure sensor was added to the first section of the finger joint (the leading edge of the walking handle). As shown in Figure 9, the test results of the four sensors are shown in the figure. We found that the three values of intent at sensor 1 are identical; there is a notable difference at sensor 2; the right turn is still similar to the stop at sensor 3; there is a significant difference at sensor 4. The sensor 2 and 4 are able to verify the three intents reliably.

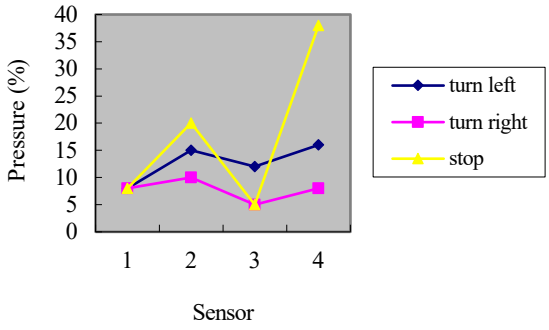


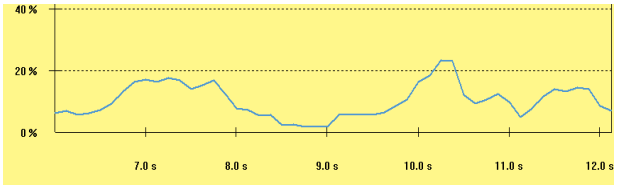
Figure 9: Pressure test results when using four sensors. The three lines represent the values of the pressures presented at the four sensors when performing three actions (left turn, right turn, and stop). The difference in pressure sensing distribution is more obvious when performing three kinds of actions.

4.3. Pressure variation with time

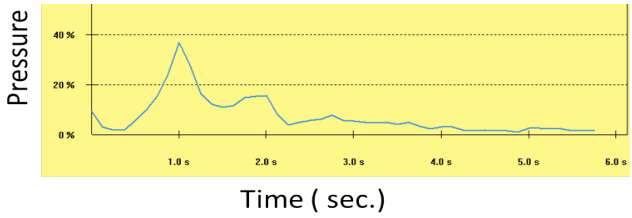
Thin-film sensor applied to acquisition the pressure variation with time — f sensors placed on the different contact area. The sensor is 0.208mm thick and 14mm in width and sensing diameter is 9.53 mm (Figure 10). The sampling rate is 20 Hz and the data is fetched through the serial port.

The measured pressure is plotted in Figure 11; the maximum force occurs in two situations: while turning, or when the walker hits an object. The maximum value does not last long (0.3-0.4 sec.). The result of the fourth sensor near the middle finger is shown in Figure 11(a), it can be found that in the temporary stop operation situation, the value of the fourth sensing zones becomes significantly more extensive, and scenarios can be separated. The measured data show that the 'stop' scenario, the sensor provide good variation to distinguish it at 10.6 sec. Figure 11(b) has a gently varying waveform showing a continuous force-applying state, and a glitch shows a collision.

Figure 10: Thin-film sensor and configurations



(a)



(b)

Figure 11: Sensor's responses: (a) the second sensor during turns left, (b) the fourth sensor during the collision and stop.

We divide the value of the sensor into three different levels; low-level (< 5%), high-level (>14%), the median level in between. If the signal maintained at the same level for a certain period, the phenomenon presented a user's action. The maximum pressure value maintained for 0.3 sec during turning moment (Figure 11.a).

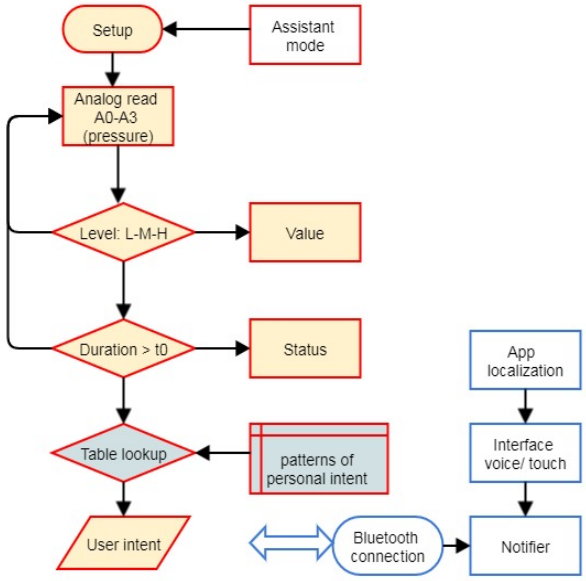


Figure 12: Flow chart of the Arduino sensing module

4.4. Implementation

The sensing module of the walker can implement with Arduino, which connects the four pressure sensors to the analog input pins of A0 to A3 (digital value 0-1023). The percentage of the Y-axis on the left side of Figure 11(b) represents the maximum measured value of the pressure sensor. We set a high pressure zone is above 14% (digital value 143), and a low pressure zone is below 5% (digital value 51), and a medium zone between the two (range 5%-14%). For example, in 2sec time (sampling rate is 20 Hz), the high pressure value is in the range of 0.75 seconds to 1.3 seconds, 1.75 seconds to 2.0 seconds; the low pressure time is between 0.1 seconds and 0.45 seconds, and the rest is medium pressure value. When maintaining the same level for a period of time, we can compare the pressure distribution of the four pressure sensors to predict the user's intention. The sensing module determines the rotation or deceleration through a logic judgment and controls the rotation speed of the motor through the voltage value of the output pin (Figure 12). There are many applications on the phone that handle positioning and navigation. The sensing module can also be connected to the mobile phone via Bluetooth, and the action to be performed is notified to the elderly by the voice of the mobile phone without being overwhelmed.

5. Conclusions

The walker is suitable for elderly in public environments, such as hospitals or retirement homes. The concept design emphasized on ergonomic features; such as weight reduction, folding, sensing the user's intent, and the like. Physical prototype and hardware circuit implementation is ready-to-expand Arduino module to perform force sensing. The smart phone app helps the elderly to move in a safe environment.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research was supported by the National Science Council, Taiwan, ROC, under grant NSC 99-2221-E-324-026-MY2. We gratefully thank Wenjie Xiao and Yusheng Huang for model making and the field test.

References

[1] Z. Maćkowiak, W. Osiński, and A. Salamon, "The effect of sensorimotor training on the postural stability of visually impaired women over 50 years of age," Journal of Women & Aging, 27(1), 68-80, 2015. https://doi: 10.1080/08952841.2014.928140

[2] T. Nakamura, "Quantitative analysis of gait in the visually impaired, disability and rehabilitation" (19)5, 194-197, 1977. https://doi: 10.3109/09638289709166526

[3] E. M. Phillips, J. C. Schneider, and G. R. Mercer, "Motivating elders to initiate and maintain ex-ercise1" Archives of Physical Medicine and Rehabilitation, (85)1, 52-57, 2004.

[4] K. Rector, "Exploring the opportunities and challenges with exercise technologies for people who are blind or low-vision" in 2015 17th International ACM SIGACCESS Conference on Computers & Accessibility, Lisbon, Portugal, 2015.

[5] C. Vlad, F. Röthenbacher, "The Chatty Environment–providing everyday independence to the visually impaired" in Workshop on Ubiquitous Computing for Pervasive Healthcare Applications, UbiComp, Seattle, Washington, USA, 2003.

[6] M. H. A. Wahab, A. A. Talib, H. A. Kadir, A. Noraziah, A. A. Mutalib, "Smart cane: Assistive cane for visually-impaired people" arXiv preprint, 1110-1115, 2011.

[7] Sorour S, Lostanlen Y, Valaee S, Majeed K., "Joint indoor localization and radio map construction with limited deployment load," IEEE Transactions on Mobile Computing, (14) 5, 1031-43, 2015.

[8] J. J. Leonard, H. F. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons" IEEE Transactions on Robotics and Automation, (7)3, 376-382, 1991.

[9] J. Biswas, M. Veloso, "Depth camera based indoor mobile robot localization and navigation" in 2012 IEEE International Conference Robotics and Automation (ICRA), St. Paul, MN, USA, 2012.

[10] T. Tsukiyama, "In-home health monitoring system for solitary elderly" Procedia Computer Science, (63)1, 229-35, 2015.

[11] Mast, M., Burmester, M., Graf, B., Weisshardt, F., Arbeiter, G., Španěl, M., & Kronreif, G., "Design of the human-robot interaction for a semi-autonomous service robot to assist elderly people," in Ambient Assisted Living, 15-29, Springer, Cham, 2015.

[12] Granados, Diego Felipe Paez, et al., "Guiding human motions in physical human-robot interaction through COM motion control of a dance teaching robot," Humanoid Robots (Humanoids), presented at the 2016 IEEE-RAS 16th International Conference, Cancun, Mexico, 2016.

[13] T. Yano, J. H. Lee, and S. Okamoto, "Development of a biped walking robot with antagonistic actuation" International Journal of Mechanical Engineering and Robotics Research, (5)3, 196-201, 2016. https://doi: 10.18178/ijmerr.5.3.196-201

[14] V.-T. Nguyen and H. Hasegawa, "Effect of toe length on biped walking behavior," International Journal of Mechanical Engineering and Robotics Research, (7)6, 599-603, 2018. https://doi: 10.18178/ijmerr.7.6.599-603

[15] Z. Shen, S. Sam, G. Allison, and L. Cui, "A simulation-based study on a clutch-spring mechanism reducing human walking metabolic cost" International Journal of Mechanical Engineering and Robotics Research, (7) 1, 55-60, 2018. https://doi: 10.18178/ijmerr.7.1.55-60

[16] A.S. Abadi, A. Mazlomi, A.H. Jafari, "Effects of box size, frequency of lifting, and height of lift on maximum acceptable weight of lift and heart rate for male university students in Iran" Elvectronic Physician, (7)6, 1365-1370, 2015.

[17] S. Azenkot, C. Feng, and M. Cakmak, "Enabling building service robots to guide blind people: A participatory design approach," in the Eleventh ACM/IEEE International Conference on Human-Robot Interaction, Christchurch, New Zealand, 2016. https://doi: 10.1109/HRI.2016.7451727

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 53-57 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Flame Distribution and Attenuation in Narrow Channels Using Mathematical Software

Bayan Sapargaliyeva[*,1], Aigul Naukenova[1], Bakhyt Alipova[2], Javier Rodrigo Ilarri[3]

[1]*Life Safety and Environment Protection Department, M. Auezov South-Kazakhstan State University, 160012, Kazakhstan*

[2]*Department of Mathematical and Computer Modeling, International IT University, 050040, Kazakhstan*

[3]*Hydraulic Engineering and Environment Department, Polytechnic University of Valencia, 46022, Spain*

ARTICLE INFO

Article history:
Received: 01 March 2019
Accepted: 03 May, 2019
Online: 13 May, 2019

Keywords:
Explosion suppression composition
Fire extinguishing powders
Burning reagents
Efficiency of powder
Heat and mass properties
Mathematical and computer modeling
Flame attenuation, narrow channels

ABSTRACT

The article of given research presents the mathematical and computer modeling of flame attenuation in narrow channels for different fire extinguishing and explosion suppression compositions. Such kind of flame distribution and attenuation could be considered for different media and channels but in this case it was considered for narrow channels. The effect on the combustion reaction is possible with gas mixture components concentration reducing, cooling the combustion zone and slowing down of chain reactions with the help of a phlegmatizing or inhibiting substances, of which the most universal and perspective are powder materials. Therefore, a necessary condition for solving the problems of developing effective flame arresters is to find common indicators and properties of substances that can become criteria for their phlegmatizing ability. The computer model in MatLab was created for methane allowing the choice of diameter and length of channel and velocity of explosion suppression composition, showing the dependence of temperature on length of channel.

1. Introduction

This paper is an extension of work originally presented in 2018 IEEE International Conference" Management of Municipal Waste as an Important Factor of Sustainable Urban Development" (WASTE). The influence of flame propagation as a result of heat losses from the reaction zone is the main factor of ensuring explosion safety for combustible gas systems. Thermal losses make it impossible to spread the flame beyond certain boundaries of composition and instrumental conditions. The mechanism of thermal losses plays a leading role in solving the emerging problems of explosion safety. The objective is to provide a general introduction to the Mechanics of heat transfer, together with its elementary applications to Solid and Fluid Mechanics. As result it would assimilated the principal concepts and laws governing the kinematics and dynamics of deformable media. In addition, this research should understand the application of this theory to the cases (i) of infinitesimal Thermo-Elasticity, and (ii) of Newtonian and perfect Fluid Mechanics [1-4].

[*]Bayan Sapargaliyeva, Kazakhstan, +77017326344, bonya_sh@mail.ru

2. Description of Problem

Mathematical and computer modeling provided a medium of exchange for the diverse disciplines (Mechanics and IT) utilizing mathematical or computer modeling as either a theoretical or working tool. Equal attention was given to the mechanics, methodology and theory of modeling with an attempt to advocate either mathematical or computer modeling, or a combination of the two, in an integrative form [5-7]. When burning in a narrow channel, the key to the mechanism of thermal losses from the reaction zone is the conductive heat conductivity to the walls of the channel. In this case, the reaction zone of obviously explosive gas systems is cooled. In sufficiently narrow channels, heat loss is possible, leading to quenching even of the most rapid-blowing known explosive systems. This feature combustion in narrow channels used in flame arrestors with narrow channels separating apparatus in which possible initiation foci combustion fire-retardant protected by the surrounding space, filled with a potentially explosive environment. The gas and flame have no other way out of the dangerous apparatus.

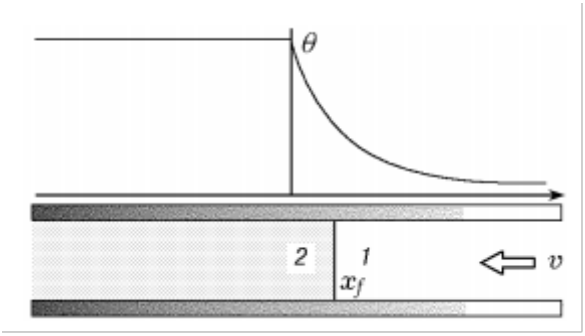


Figure 1: Flame attenuation in narrow channels

The Figure 1 shows a diagram of a flame in a flat channel. As in the classical formulation, approximations are used about the constant density, heat capacity, and thermal conductivity of the gas.

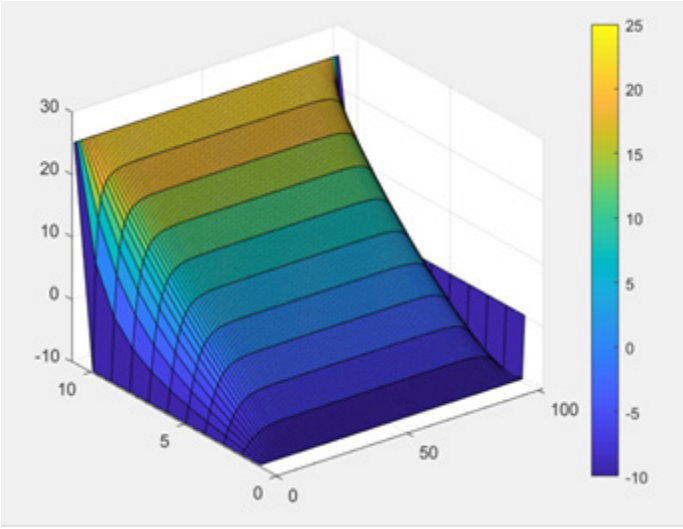


Figure. 2: Temperature distribution from wall to center of channel

This Figure 2 represents the temperature distribution from wall to center of channel. Identifying the relevant heat transfer processes on a heating when the foamed insulation sheet of channel is placed on inner surface and on the outer surface of channel. To avoid singular solution, the round configuration of channel is preferred.

3. Flame attenuation, mathematical modeling

Processes with heat regeneration include filtration combustion of gases, where the flame can stabilize in microchannels between the solid phases, even if the dimensions of the microchannels are substantially smaller than the critical diameter determined from the initial temperature. Combustion in this case is possible provided that the gas phase is surrounded by walls with a temperature exceeding the ambient temperature. This work can be considered as a continuation of the classical works on flame propagation taking into account the transfer of heat from the combustion products to the fresh mixture by an inert medium due to its high thermal conductivity.

Combustion of gas in a channel with heat exchange is the simplest example of a system with a convective heat recovery method. A characteristic feature of combustion in such a system is the thermal interaction between the combustion products and the fresh mixture through the heat-conducting wall. In this linear configuration, unlike devices with special combustion chambers, the flame stabilizes inside the inlet tube. This mechanism of flame stabilization is typical for small-sized systems with recirculation of heat. Studies of the dynamic behavior of a flame in a system with counter-current heat transfer make it possible to study the characteristics of combustion in systems with heat recovery, and also covers a whole series of new phenomena.

It is necessary to solve the Cauchy problem for the one-dimensional heat equation (a second-order partial differential equation that describes the temperature distribution in a given region of space and its variation over time) using parallel programming tools. The problem is solved for a homogeneous heat equation (the system is insulated) on the interval [0, L]: One-dimensional dimensionless equations for temperature have the form

$$\frac{\partial T_{1,2}}{\partial t} = \frac{\partial^2 T_{1,2}}{\partial x^2} + V\frac{\partial T_{1,2}}{\partial x} - \Omega(T_{1,2} - \theta(x))$$

Here the dimensionless time t is measured in units of correlation between thermal properties of gas in channel and adiabatic velocity of gas, i.e. æ/Vb², the dimensionless spatial coordinate x is in units of æ/Vb, where æ is the thermal diffusivity of the gas, and Vb is the adiabatic velocity of the flat flame. The channel wall temperature θ and the gas temperature T are measured in terms of the adiabatic flame temperature Tb [8-10].

The subscripts 1 and 2 refer to the fresh mixture and combustion products, respectively. The gas velocity V is measured in units of Vb. The dimensionless heat transfer parameter $\Omega = 4Nu/Pe^2$ is related to the Nusselt number $Nu=\alpha d_o/\lambda g$ and the Peclet number $Pe=Vbd_o/æ$, which are determined for the channel of height d_o (here α is the heat transfer coefficient; λg is the thermal conductivity of the gas). The boundary conditions for equations (1) far from the flame have the following form:

$$x \to +\infty:\ T_1 \to \sigma,\ C_1 \to 1,\ \theta \to \sigma$$
$$x \to -\infty:\ T_2 \to \Theta,\ \theta \to \Theta > \sigma$$
$$\sigma = \frac{T_0}{T_b},\quad \Theta = \frac{T_w}{T_b}$$

where T0 is the initial temperature of the fresh mixture, Tw is the wall temperature at the hot end of the channel.

Thus, Equation (1) together with boundary conditions (2) at the ends of a narrow channel determines the spread of the flame in the channel. In this case, the temperature distribution in the channel is calculated (the flame temperature of the combustible gas and its mixture before and after mixing, the temperature of the channel walls and their combination are taken into account) as a function of time. Also, the differential equation of the second order includes the velocities and accelerations of the temperature change. The gas characteristics were calculated from some physical parameters: thermal diffusivity of the gas, dimensionless heat transfer parameter (as correlation of numbers of Nusselt and Peclet), heat transfer coefficient; the thermal conductivity of the gas. Thus, taking into account the boundary conditions (2), the boundary value problem was modeled and coded in Matlab.

4. Numerical calculations in MatLab

The numerical realization of heat transfer process was implemented in MatLab, and Adomian Decomposition Method was used for calculation of components of displacement of parts

of fire extinguishing powders and burning reagents. This calculation needs to use some special functions for discretization of components of temperature.

```
% variable definition
Nx = 11;
Ny = 101;
X = 1.0;
Y = 10.0;
dx = X / (Nx - 1);
dy = Y / (Ny - 1);
dt = 1e-4;

% matrix definition
% initializing temperature in our  fire extinguishing and explosion
suppression compositions
U1 = ones(Nx, Ny) * (-10);
U2 = ones(Nx, Ny) * (-10);
U3 = ones(Nx, Ny) * (-10);

% initializing temperature on right border in our fire extinguishing and
explosion suppression compositions
for j=1:Ny + 1
  U1(Nx, j) = 25;
  U2(Nx, j) = 25;
  U3(Nx, j) = 25;
end

% loop till it converges
while 1
%    Using Adomian Decomposition Method to solve this eq
   A = 1 / (dx * dx);
   B = - (1 / dt + 2 / (dx * dx));
   C = 1 / (dx * dx);
   gamma = zeros(Nx);
   beta = ones(Nx) * (-10);

%    First loop using the definition temperature by X axis
   for j=2:Ny - 1
      for i=2:Nx
         F = -(discretize(U1, "y", i, j) + U1(i, j) / dt);
         temp_div = B + C * gamma(i);
         beta(i + 1) = (F - C * beta(i)) / temp_div;
         gamma(i + 1) = -A / temp_div;
      end
%       U2(Nx, j) = - beta(j) / (1 - gamma(j));
      for i=Nx:-1:2
         U2(i - 1, j) = gamma(i) * U2(i, j) + beta(i);
      end
   end

   gamma = zeros(Nx);
   beta = ones(Nx) * (-10);

   A = 1 / (dy * dy);
   B = - (1 / dt + 2 / (dy * dy));
   C = 1 / (dy * dy);
%    Second loop using of temperature by Y axis
   for i=2:Nx - 1
      for j=2:Ny
         F = -(discretize(U2, "x", i, j) + U2(i, j) / dt);
         temp_div = B + C * gamma(j);
         beta(j + 1) = (F - C * beta(j)) / temp_div;
         gamma(j + 1) = -A / temp_div;
      end
      for j=Ny:-1:2
         U3(i, j - 1) = gamma(j) * U3(i, j) + beta(j);
      end
   end
%    for j=Ny:-1:2
%       U3(Nx - 1, j) = U3(Nx - 2, j);
%    end
% And after every iteration check for convergence
   if converge(U1, U3)
      break;
   end
% In the end of every iteration - definition of new layer as old, and
continuation till difference between them become less than epsilon
   U1 = U3;
   surf(U3)
end
% Plot surface to image easier the result
surf(U3)
colorbar
U3((Nx - 1) / 2, (Ny - 1) / 2)

% This function checks for convergence
function result = converge (U1, U2)
% Initial max difference is -1, because of it less than 0, and as we know
% absolute value of difference never be less than 0
max_diff = -1;
% My error rate (Epsilon)
eps = 1e-4;
Nx = 11;
Ny = 101;
% This loop gets max difference Between old and new layers
for i=1:Nx
   for j=1:Ny + 1
      if abs(U1(i, j) - U2(i, j)) > max_diff
         max_diff = abs(U1(i, j) - U2(i, j));
      end
   end
end
% As the result we get Boolean value
result = eps > max_diff;
end

% This function helps me make code easier to understand
function result = discretize (U, axis, i, j)
Nx = 11;
Ny = 101;
X = 1.0;
Y = 10.0;
dx = X / (Nx - 1);
dy = Y / (Ny - 1);
if axis == "x"
   result = (U(i + 1, j) - 2 * U(i, j) + U(i - 1, j)) / (dx * dx);
else
   result = (U(i, j + 1) - 2 * U(i, j) + U(i, j - 1)) / (dy * dy);
end
end
```

The basic mechanisms of combustion of gases in systems with heat recirculation can be clarified from the study of combustion of gases in a narrow channel with a controlled wall temperature. Therefore,

studies of gas combustion in micro-channels with controlled wall temperature are necessary for understanding processes in both microsystems with heat regeneration, where the heating of the combustible mixture of gases occurs due to the heat of the exhaust combustion products, and for the development of new technologies that use combustion of gases in various media. The small dimensions of the channel in which combustion occurs allow the creation of an almost uniform temperature distribution in the cross section. Experiments show that the temperature gradient in the walls of the channel, which was created by an external heat source, remains practically constant, both with the combustion of gas inside the channel and without burning. These observations make it promising to use a microchannel with a temperature gradient to study the structure of the flame and to test existing models with detailed kinetics of chemical reactions.

Burning studies usually consider gas flows in which transport phenomena and chemical reactions play an important role. In order to analyze the combustion process, in addition to the chemical transformations of the substance, it is necessary to take into account the phenomena caused by the release of heat. If we take into account the interconnectedness of the above processes, then the complexity of the theoretical analysis of combustion becomes obvious.

Advances in the theory of combustion of gases were associated, first of all, with the use of simple and at the same time realistic mathematical models.

Software Package of given formulas are special calculations that can contain class folders, function, and class definition files, and other packages. The names of classes and functions in Matlab are scoped to the package folder. A package with name" TED in channels" (Thermoelastodynamics in channels) is unique. Function and class names must be unique only within the package. Using a package provides a means to organize classes and functions. Packages also enable you to reuse the names of classes and functions in different packages.

The proposed code of the software product allows you to transparently change the approaches to calculation, change the parameters to be set. Explicit schemes for solving the differential equation are proposed. Also in Matlab, visualization of the computational process was carried out.

The Graphical User Interface GNUI requires the following parameters:

- Diameter of channel (in m),
- Velocity of motion gas or fluid (in m/sec),
- Length of channel (in m).

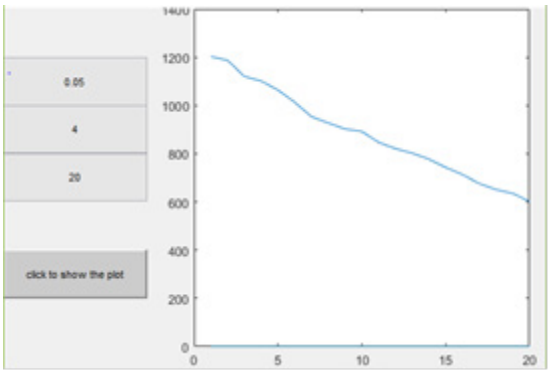


Figure 3: Visualization of flame attenuation in narrow channels in TED in channels

Figure 3 represents the Interface of package and input/output parameters. The graph of the dependence of the change in the combustion temperature of methane on the length of a narrow channel clearly indicates a decrease in temperature during motion away from the flame front, which confirms the physical nature of the problem and the correctness of the chosen mathematical model.

Thus, a mathematical model of flame burning in a narrow channel has been developed, the properties of burning and quenching of gas or liquid in a narrow channel have been studied, depending on the combustion temperature and the flame propagation velocity in the channel. In this case, the inhibitory properties of the substance, flame propagation in thin channels are investigated [11-13]. Based on the 2-D model, a computer model was created in MatLab14. This model requires the specification of some input parameters and provides a graphical representation of the temperature dependence of the length of the channel. The computer model allows you to transform and supplement the subsequent calculations.

As a result of computer modeling of combustion processes in a system with convective heat recirculation, confirmed by some experiments, it was shown that

- in this system

a) it is possible to sustain combustion in a channel with a diameter less than critical,

b) combustion with velocities greater than the normal velocity of the free flame and gas temperature may exceed the adiabatic temperature of the free flame,

c) the flame is resistant to external influences caused by a change in the ambient temperature, changes in the gas flow rate and other factors.

The possibility of flame stabilization in a straight channel with a temperature gradient in the walls, which is created by the combustion products, is shown. These results allow us to conclude that this system is promising for creating, for example, micro-burners.

- A new type of instability has been discovered in the combustion of gas in micro-channels with a longitudinal temperature gradient in the walls of the channel. A feature of this type of flame instability is the presence of an inhomogeneous velocity profile of the gas in the channel and heat exchange with hot walls. The range of values of the gas flow rate at which this type of instability manifests itself is found. It is shown that in a radial channel this type of instability leads to the formation of rotating flame structures, in particular, a spiral flame detected in experiments.

5. Conclusion

The mathematical model of flame attenuation in narrow channels for different fire extinguishing and explosion suppression compositions is given. Such kind of flame distribution and attenuation could be considered for different media and channels but in given research it was considered for narrow channels. The effect on the combustion reaction is possible with gas mixture components concentration reducing, cooling the combustion zone and slowing down of chain reactions with the help of a phlegmatizing or inhibiting substances, of which the most universal and perspective are powder materials. As in the classical formulation, approximations are used about the constant density, heat capacity, and thermal conductivity of the gas. Caushy task

for a second-order partial differential equation that describes the temperature distribution in a given region of space and its variation over time was solved. The numerical realization of heat transfer process was implemented in MatLab, and Adomian Decomposition Method was used for calculation of components of displacement of parts of fire extinguishing powders and burning reagents. The interface with input/output parameters was launched and tested. A package with name" TED in channels" (Thermoelastodynamics in channels) enables to reuse the names of classes and functions in different packages.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

Authors are grateful for scientific and methodological support to M. Auezov South-Kazakhstan State University, Kazakhstan.

References

[1] Alipova B.N., Alexeyeva L.A., Dadaeva A.N. Shock waves as generalized solutions of thermoelastodynamics equations. On the uniqueness of boundary value problems solutions // AIP Conference proceedings, 2017, V. 1798, I. 020003, pp. 1-8 DOI: 10.1063/1.4972595

[2] Fennandez-Pello A. C. Micro-power generation using combustion: issues and approaches. Twenty-ninth intern. symp. on combustion, Pittsburg, 2002, V. 40, no. 5, pp. 883– 899.

[3] Sitzki, L., Borer, K., Wussow, S. Schuster, E., Maruta, K., Ronney, P. and Cohen, A., "Combustion in Microscale Heat-Recirculating Burners", 38th AIAA Space Sciences & Exhibit, Reno, NV, 2001, V. 22, no. 3, p. 1087.

[4] Vican, J. Gajdeczko, B.F., Dryer, F.L., Milius, D.L., Aksay, I.A. and Yetter, R.A."Development of a Microreactor as a Thermal Source for MEMS Power Generation" Proceedings of the Twenty-Ninth International Symposium on Combustion, Sapporo, Japan, 2002, V. 40, no. 5, pp. 906-916.

[5] Lloyd S. A. and Weinberg F. J. A burner for mixtures of very low heat content NATURE, 1974, V. 251, pp. 47–49.

[6] Jones A. R., Lloyed S. A., and Weinberg F. J. Combustion in heat exchangers // Proc. Roy. Soc. Lond., 1978, V. A369, pp. 97–115.

[7] Buckmaster J. D. The effects of radiation on streched flames // Combust. Theory Modelling, 1997, no. 1 pp. 1–11.

[8] Zamaschikov V.V., Minaev S.S. Limits of flame propagation in a narrow channel during gas filtration // Physics of Combustion and Explosion, 2001, V. 37, no. 1, pp. 25–31.

[9] Zeldovich Ya. B. The theory of propagation of a quiet flame // Journal of Energy and Solids, 1941, V. 11, I. 1, pp. 159–168.

[10] Minaev S.S., Babkin V.S. Propagation of a flame in a channel of variable cross section during gas filtration // Physics of Combustion and Explosion, 2001, V. 37, no. 1, pp. 16–24.

[11] Sapargalieva B., Naukenova A., Alipova B., Javier Rodrigo Illari, Shapalov Sh. The analysis of heat and mass properties of the fire extinguishing powder in effectiveness criteria // News of the National Academy of Science of the Republic of Kazakhstan. Series of Geology and Technical sciences, 2018,. ISSN 2518-170X (Online), ISSN 2224-5278 (Print), V. 3, no. 4 (430), pp. 51-61.

[12] M. Raoofian Naeeni, M. Eskandari-Ghadi, A. A. Ardalan, S. Sture, and M. Rahimian, "Transient response of a thermoelastic half-space to mechanical and thermal buried sources," ZAMM **95** (4), 354–376 (2015).

[13] H. Mahmoodi Kordkhieli, G. Ghodrati Amiri, and M. Hosseini, "Axisymmetric analysis of a thermoelastic isotropic half-space under buried sources in displacement and temperature potentials," J. Thermal Stresses **40** (2), 237–254 (2017).

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 58-62 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

An Expanded Derivation of Suppressing Anisotropic Birefringence in a Rectangular Waveguide

Gregory Mitchell*

U.S. Army Research Laboratory, 20783, USA

ARTICLE INFO

Article history:
Received: 06 February 2019
Accepted: 03 May, 2019
Online: 13 May, 2019

Keywords:
Anisotropic media
Birefringence
Anisotropic dispersion
Rectangular waveguide
Biaxial anisotropy

ABSTRACT

Birefringence is a known property of anisotropic media where incidence of a single propagating wave on a biaxial boundary gives rise to what scientists refer to as an ordinary wave and extraordinary wave. This article explores the anisotropic electromagnetic theory that gives rise to the phenomenon of birefringence focusing on both ordinary and extraordinary plane wave propagation. The dispersion equation in an unbounded anisotropic medium leads to a fourth order polynomial solution for the wave number which is shown to be the root cause of birefringence. Finally, a derivation of the unique properties for a partially filled biaxial rectangular waveguide that serves to suppress wave propagation in the vertical direction will negate the effect of birefringence in the case of the first propagating mode.

1. Introduction

At the invitation of the associate editors of the Advances in Science, Technology and Engineering Systems Journal (ASTESJ), the author presents the following manuscript on an in-depth theoretical expansion of a conference paper previously submitted in the Applied Computational Electromagnetics Society (ACES) Symposium proceedings in 2018 [1]. Originally, the author only discussed the explanation of the suppression effect of anisotropic birefringence inside a rectangular waveguide in brief, and the author feels that scientific interest in the subject warrants a more in depth theoretical examination of the phenomenon. The author has also presented other portions of the theory relating to anisotropic wave propagation in a second publication [2]. Therefore, the author is happy to contribute an expansion on this aspect of the anisotropic birefringence theory to a special issue of ASTESJ at the invite of the associate editors. The article will focus on how the derivation of the anisotropic wave equation in a biaxial anisotropic medium leads to birefringence through a fourth order dispersion equation. The final section will explain how constraints on the boundary conditions of modes in rectangular waveguides can negate the effect of birefringence.

Scientists have understood electromagnetic wave propagation in homogeneous anisotropic optical media since the 1950s. At optical frequencies, one has to rely on naturally occurring crystalline media with anisotropic properties. However, as early as 1958 Collin showed that at microwave frequencies, where the wavelength is much greater, it is possible to fabricate artificial dielectric media having anisotropic properties [3].

Recently engineered materials have come to play a dominant role in the design and implementation of electromagnetic devices and especially antennas. Metamaterials, MetaFerrites, and magneto-dielectrics have all come to play a crucial role in advances made in both the functionality and characterization of such devices. In fact, a movement towards utilizing customized material properties to replace the functionality of traditional radio frequency (RF) components such as broadband matching circuitry, ground planes, and directive elements is apparent in the literature and not just replacement of traditional substrates and superstrates with engineered structures. A firm theoretical understanding of the electromagnetic properties of these materials is necessary for both design and simulation of new and improved RF devices.

Inherently, some of these engineered materials have anisotropic properties. Previously, the study of anisotropy had been limited mostly to the realm of optical frequencies where the phenomenon occurs naturally in substances such as liquid crystals and plasmas. However, the recent development of a separate class engineered materials at microwave frequencies has encouraged the study of electromagnetic anisotropy for RF applications. Low loss anisotropic magneto-dielectrics greatly expands the current antenna design space. Anisotropic media provide unique properties not available using isotropic media. These include a high effective refractive index reducing the electrical size of wavelengths in a medium, relatively low magnetic loss at MHz

*Gregory Mitchell, U.S. Army Research Laboratory,
Email: gregory.a.mitchell1.civ@mail.mil

frequencies, and lower density and weight than traditional isotropic ferrite materials. Anisotropic media can be up to five times less dense than traditional ferrites while exhibiting magnetic properties with lower loss tangents. Relevant work in this area includes plane wave solutions to propagation in anisotropic media [4-6]. Researchers have also performed modal decomposition of rectangular waveguides filled with anisotropic magneto-dielectrics along with in depth analysis of cut-off wave numbers and field distributions [7, 8].

An anisotropic electromagnetic medium defines permittivity ($\underline{\underline{\varepsilon_r}}$) and permeability ($\underline{\underline{\mu_r}}$) as separate tensors where the values differ in all three Cartesian directions ($\varepsilon_x \neq \varepsilon_y \neq \varepsilon_z$ and $\mu_x \neq \mu_y \neq \mu_z$). This is known as the biaxial definition of anisotropic material which is more encompassing than the uniaxial definition which makes the simplifying assumption that $\varepsilon_x=\varepsilon_y=\varepsilon_t$ and $\mu_x=\mu_y=\mu_t$. The anisotropic definition also differs from the traditional isotropic definition where ε_r and μ_r are the same in all three Cartesian directions defined by a single value. Anisotropic media yield characteristics such as conformal surfaces, the focusing and refraction of electromagnetic waves as they propagate through a medium, high impedance surfaces for artificial magnetic conductors as well as high index, low loss, and lightweight ferrite materials. The following sections aim to discuss in more detail some RF applications directly affected by the incorporation of anisotropic media and will also present a rigorous derivation of the wave equation and dispersion relationships for anisotropic magneto-dielectric media. All results agree with those presented by Meng, et. al [7, 8]. Furthermore, setting $\underline{\underline{\mu_r}} = \underline{\underline{I}}$, where $\underline{\underline{I}}$ is the identity matrix, yields results that agree with those presented by Pozar and Graham for anisotropic dielectric media [6, 9].

2. Applications of Anisotropy in Radio Frequency (RF) Devices

Traditionally, the study of anisotropic properties for RF was limited to a narrow application space where traditional ferrites exhibiting natural anisotropy were the enabling technology. These types of applications included isolators, absorbers, circulators and phase shifters [10, 11]. Traditional ferrites are generally very heavy and very lossy at microwave frequencies and those are the two main limiting factors narrowing their use in RF devices; however, propagation loss is an important asset to devices such as absorbers. Anisotropy itself leads to propagation of an electromagnetic wave in different directions. This phenomenon is already important in devices such as circulators and isolators [10]. For phase shifters and other control devices, changing the bias field across the ferrite controls the microwave signal [11]. However, newer versions of some of these devices, utilizing FETs and diodes in the case of phase shifters, rely on isotropic media to enable higher efficiency devices.

As early as 1958, Collin showed that at microwave frequencies, where the wavelength is larger, it is possible to fabricate artificial dielectric media having anisotropic properties [3]. This has led some to investigate known theoretical solutions to typical RF problems, such as a microstrip patch antenna, and extend them utilizing anisotropic wave propagation in dielectric media [5, 6]. The anisotropic dielectric antenna shows interesting features of basic antenna applications featuring anisotropic substrates. While these solutions establish a framework for electromagnetic wave propagation in anisotropic media, they simplify the problem by necessarily setting $\mu_r=1$ and only focusing on dielectric phenomena of anisotropy.

The proliferation of metamaterials research over the last few decades exemplifies the growing interest in the area of artificial media [12-22]. Metamaterials incorporate the use of artificial microstructures made of subwavelength inclusions implemented with periodic and/or multilayered structures known as unit cells [13]. These devices operate where the wavelength is much larger than the characteristic dimensions of the unit cell elements. One characteristic feature of some types of metamaterials is wave propagation anisotropy [14]. Applications such as directive lensing [15], cloaking [16], electronic beam steering [17], metasurfaces [18], anisotropic oscillators [19], and optics [20] among others all utilize anisotropic media as the enabling technology. More recent research focuses on the proper modeling schemes for computational electromagnetics [21] or examining phenomenon such as scattering off of anisotropic objects [22].

A class of engineered materials exists, known as MetaFerrites, that exhibits positive refractive index, anisotropy, and magneto-dielectric properties with reduced propagation loss at microwave frequencies compared to traditional ferrites. These materials show the unique ability to provide broadband impedance matches for very low profile antennas by exploiting the inherent anisotropy to redirect surface waves thus improving the impedance match of the antenna when very close to a ground plane. Designers have demonstrated antenna profiles at 200 MHz to 500 MHz on the orders of a fortieth of a wavelength using these anisotropic media with over an octave of bandwidth and positive realized gain [2, 23].

3. Anisotropic Wave Equation

The recent development of low loss anisotropic magneto-dielectrics greatly expands the current antenna design space. Here we present a rigorous derivation of the wave equation and dispersion relationships for anisotropic magneto-dielectric media. All results agree with those presented by Meng, et. al [7, 8]. Furthermore, setting $\underline{\underline{\mu_r}} = \underline{\underline{I}}$, where $\underline{\underline{I}}$ is the identity matrix, yields results that agree with those presented by Pozar and Graham for anisotropic dielectric media [6, 9]. Incorporating a fully developed derivation of anisotropic properties of both $\underline{\underline{\varepsilon_r}}$ and $\underline{\underline{\mu_r}}$ expands upon the simplification imposed by both Pozar and Graham that uses an isotropic value of $\mu_r=1$. An expansion on the results of Meng, et. al given in section 4 develops the waveguide theory including a full modal decomposition utilizing the biaxial definition of anisotropy versus their simplified uniaxial definition. The derivation of anisotropic rectangular waveguide resonance in section 6 differs from that of Meng, et. al by addressing the direct relationship between boundary conditions and the existence of birefringence. Furthermore, the analysis of anisotropic properties is not restricted to double negative materials, which is the case for the Meng, et. al studies.

In order to solve for the propagation constants we will need to formulate the dispersion relationship from the anisotropic wave equation. This allows us to solve for the propagation constant in the normal direction of the anisotropic medium. We start with the anisotropic, time harmonic form of Maxwell's source free equations for the electric and magnetic fields

$$\nabla x\underline{\boldsymbol{E}} = -j\omega\mu_o\underline{\underline{\boldsymbol{\mu}}}_r\cdot\underline{\boldsymbol{H}}, \tag{1}$$

$$\nabla x\underline{\boldsymbol{H}} = j\omega\varepsilon_o\underline{\underline{\boldsymbol{\varepsilon}}}_r\cdot\underline{\boldsymbol{E}}, \tag{2}$$

where ω is the frequency in radians, ε_o is the permittivity of free space, μ_o is the permeability of free space, $\underline{\boldsymbol{E}}=\underline{\boldsymbol{x}}_oE_x+\underline{\boldsymbol{y}}_oE_y+\underline{\boldsymbol{z}}_oE_z$, and $\underline{\boldsymbol{H}}=\underline{\boldsymbol{x}}_oH_x+\underline{\boldsymbol{y}}_oH_y+\underline{\boldsymbol{z}}_oH_z$. We define $\underline{\underline{\boldsymbol{\mu}}}_r$ and $\underline{\underline{\boldsymbol{\varepsilon}}}_r$ as

$$\underline{\underline{\varepsilon}}_r = \begin{bmatrix} \varepsilon_x & 0 & 0 \\ 0 & \varepsilon_y & 0 \\ 0 & 0 & \varepsilon_z \end{bmatrix}, \tag{3}$$

$$\underline{\underline{\mu}}_r = \begin{bmatrix} \mu_x & 0 & 0 \\ 0 & \mu_y & 0 \\ 0 & 0 & \mu_z \end{bmatrix}. \tag{4}$$

Combining (1-4) yields

$$\begin{aligned} &\underline{x}_o\left(\frac{dE_Z}{dy}-\frac{dE_Y}{dz}\right)+\underline{y}_o\left(\frac{dE_X}{dz}-\frac{dE_Z}{dx}\right)+\underline{z}_o\left(\frac{dE_Y}{dx}-\frac{dE_X}{dy}\right), \\ &\quad = -j\omega\mu_o\left(\mu_xH_X\underline{x}_o+\mu_yH_Y\underline{y}_o+\mu_zH_Z\underline{z}_o\right) \end{aligned} \tag{5}$$

$$\begin{aligned} &\underline{x}_o\left(\frac{dH_Z}{dy}-\frac{dH_Y}{dz}\right)+\underline{y}_o\left(\frac{dH_X}{dz}-\frac{dH_Z}{dx}\right)+\underline{z}_o\left(\frac{dH_Y}{dx}-\frac{dH_X}{dy}\right). \\ &\quad = j\omega\varepsilon_o\left(\varepsilon_xE_X\underline{x}_o+\varepsilon_yE_Y\underline{y}_o+\varepsilon_zE_Z\underline{z}_o\right) \end{aligned} \tag{6}$$

Using the radiation condition, we assume a solution of $\underline{\boldsymbol{E}}(\underline{\boldsymbol{r}})=\underline{\boldsymbol{E}}(x,y)\exp(-jk_zz)$. Now isolate the individual components of (5) by taking the dot product with $\underline{\boldsymbol{x}}_o$, $\underline{\boldsymbol{y}}_o$, and $\underline{\boldsymbol{z}}_o$ respectively. This operation yields the following equations

$$dE_z/dy - jk_zE_y = -j\omega\mu_o\mu_xH_x, \tag{7}$$

$$jk_zE_x - dE_z/dx = -j\omega\mu_o\mu_yH_y, \tag{8}$$

$$dE_y/dx - dE_x/dy = -j\omega\mu_o\mu_zH_z. \tag{9}$$

The same procedure assuming a solution of $\underline{\boldsymbol{H}}(\underline{\boldsymbol{r}})=\underline{\boldsymbol{H}}(x,y)\exp(-jk_zz)$ for (6) yields

$$dH_z/dy - jk_zH_y = j\omega\varepsilon_o\varepsilon_xE_x, \tag{10}$$

$$jk_zH_x - dH_z/dx = j\omega\varepsilon_o\varepsilon_yE_y, \tag{11}$$

$$dH_y/dx - dH_x/dy = j\omega\varepsilon_o\varepsilon_zE_z. \tag{12}$$

Equations (7-12) allow us to solve for the transverse field components of the electric and magnetic fields in terms of the derivatives of H_z and E_z as

$$E_x = -\frac{j}{k_o^2\mu_y\varepsilon_x-k_z^2}\left(\omega\mu_o\mu_y\frac{dH_z}{dy}+k_z\frac{dE_z}{dx}\right), \tag{13}$$

$$E_y = \frac{j}{k_o^2\mu_x\varepsilon_y-k_z^2}\left(\omega\mu_o\mu_x\frac{dH_z}{dx}-k_z\frac{dE_z}{dy}\right), \tag{14}$$

$$H_x = \frac{j}{k_o^2\mu_x\varepsilon_y-k_z^2}\left(\omega\varepsilon_o\varepsilon_y\frac{dE_z}{dy}-k_z\frac{dH_z}{dx}\right), \tag{15}$$

$$H_y = -\frac{j}{k_o^2\mu_y\varepsilon_x-k_z^2}\left(\omega\varepsilon_o\varepsilon_x\frac{dE_z}{dx}+k_z\frac{dH_z}{dy}\right). \tag{16}$$

Now that we have relationships for the transverse field components, we can solve (1) and (2) for $\underline{\boldsymbol{H}}$ and $\underline{\boldsymbol{E}}$ respectively

$$\underline{H} = -\underline{\mu}_r^{-1}\cdot\left(\nabla x\underline{E}\right)/\left(j\omega\mu_o\right), \tag{17}$$

$$\underline{E} = \underline{\varepsilon}_r^{-1}\cdot\left(\nabla x\underline{H}\right)/\left(j\omega\varepsilon_o\right). \tag{18}$$

Taking the cross product of both sides and substituting (1) for $\nabla x\underline{\boldsymbol{E}}$ and (2) for $\nabla x\underline{\boldsymbol{H}}$ yields [8]

$$\nabla\times\underline{E} = \nabla\times\underline{\varepsilon}_r^{-1}\cdot\left(\nabla x\underline{H}\right)/\left(j\omega\varepsilon_o\right), \tag{19}$$

$$\nabla\times\underline{H} = -\nabla\times\underline{\mu}_r^{-1}\cdot\left(\nabla x\underline{E}\right)/\left(j\omega\mu_o\right), \tag{20}$$

$$\nabla x\underline{\mu}_r^{-1}\cdot\left(\nabla x\underline{E}\right) = -k_o^2\underline{\varepsilon}_r\cdot\underline{E}, \tag{21}$$

$$\nabla x\underline{\mu}_r^{-1}\cdot\left(\nabla x\underline{E}\right) = -k_o^2\underline{\varepsilon}_r\cdot\underline{E}. \tag{22}$$

Equations (21) and (22) represent the vector wave equations in an anisotropic medium.

4. Dispersion Equation for H_z

We expand the $\nabla x\underline{\boldsymbol{H}}$ term of (22) in terms of (13-16), and take the dot product with $\underline{\underline{\boldsymbol{\varepsilon}}}_r^{-1}$

$$\begin{aligned} \nabla x\Bigg[&\frac{\underline{x}_o}{\varepsilon_x}\left(\frac{dH_Z}{dy}-\frac{dH_Y}{dz}\right)+\frac{\underline{y}_o}{\varepsilon_y}\left(\frac{dH_X}{dz}-\frac{dH_Z}{dx}\right) \\ &+\frac{\underline{z}_o}{\varepsilon_z}\left(\frac{dH_Y}{dx}-\frac{dH_X}{dy}\right)\Bigg] = k_o^2\underline{\mu}_r\cdot\underline{H} \end{aligned}, \tag{23}$$

Evaluating (23) and taking the dot product with the $\underline{\boldsymbol{z}}_o$ direction allows us to isolate the H_z component of the magnetic field on the right hand side of the equation

$$-\frac{1}{\varepsilon_y}\frac{d^2H_z}{dx^2}-\frac{1}{\varepsilon_x}\frac{d^2H_z}{dy^2}+\frac{1}{\varepsilon_y}\frac{d^2H_x}{dxdz}+\frac{1}{\varepsilon_x}\frac{d^2H_y}{dydz} = k_o^2\mu_zH_z, \tag{24}$$

Setting E_z=0, if we differentiate (13) by $d^2/dxdz$ and (14) by $d^2/dydz$, keeping in mind that $d/dz = -jk_z$, we arrive at the following result

$$\begin{aligned} &\frac{k_z^2}{\varepsilon_y\left(k_z^2-k_o^2\varepsilon_y\mu_x\right)}\frac{d^2H_z}{dx^2}+\frac{k_z^2}{\varepsilon_x\left(k_z^2-k_o^2\varepsilon_x\mu_y\right)}\frac{d^2H_z}{dy^2} \\ &\quad -\frac{1}{\varepsilon_y}\frac{d^2H_z}{dx^2}-\frac{1}{\varepsilon_x}\frac{d^2H_z}{dy^2} = k_o^2\mu_zH_z \end{aligned}, \tag{25}$$

Combining the d^2H_z/dx^2 and d^2H_z/dy^2 terms in (25) gives the following second order differential dispersion equation for H_z

$$\frac{k_o^2\mu_x}{k_o^2\mu_x\varepsilon_y-k_z^2}\frac{d^2H_z}{dx^2}+\frac{k_o^2\mu_y}{k_o^2\mu_y\varepsilon_x-k_z^2}\frac{d^2H_z}{dy^2}+k_o^2\mu_zH_z = 0. \tag{26}$$

5. Dispersion Equation for E_z

We expand the $\nabla x\underline{\boldsymbol{E}}$ term of (21) in terms of (13-16), and take the dot product with $\underline{\underline{\boldsymbol{\mu}}}_r^{-1}$

$$\begin{aligned} \nabla x\Bigg[&\frac{\underline{x}_o}{\mu_x}\left(\frac{dE_z}{dy}-\frac{dE_y}{dz}\right)+\frac{\underline{y}_o}{\mu_y}\left(\frac{dE_x}{dz}-\frac{dE_z}{dx}\right) \\ &+\frac{\underline{z}_o}{\mu_z}\left(\frac{dE_y}{dx}-\frac{dE_x}{dy}\right)\Bigg] = k_o^2\underline{\varepsilon}_r\cdot\underline{E} \end{aligned}. \tag{27}$$

Evaluating (27) and taking the dot product with the $\underline{z}_o$ direction allows us to isolate the E_z component of the magnetic field on the right hand side of the equation

$$-\frac{1}{\mu_y}\frac{d^2E_z}{dx^2}-\frac{1}{\mu_x}\frac{d^2E_z}{dy^2}+\frac{1}{\mu_y}\frac{d^2E_x}{dxdz}+\frac{1}{\mu_x}\frac{d^2E_y}{dydz}=k_o^2\varepsilon_z E_z . \quad (28)$$

Setting H_z=0, if we differentiate (15) by $d^2/dxdz$ and (16) by $d^2/dydz$, keeping in mind that $d/dz = -jk_z$, and plug the results into (28), then we arrive at the following result

$$\begin{gathered}\frac{k_z^2}{\mu_y\left(k_z^2-k_o^2\mu_y\varepsilon_x\right)}\frac{d^2E_z}{dx^2}+\frac{k_z^2}{\mu_x\left(k_z^2-k_o^2\mu_x\varepsilon_y\right)}\frac{d^2E_z}{dy^2}\\ -\frac{1}{\mu_y}\frac{d^2E_z}{dx^2}-\frac{1}{\mu_x}\frac{d^2E_z}{dy^2}=k_o^2\varepsilon_z E_z\end{gathered} . \quad (29)$$

Combining the d^2E_z/dx^2 and d^2E_z/dy^2 terms in (38) gives the following second order differential dispersion equation for E_z

$$\frac{k_o^2\varepsilon_x}{k_o^2\mu_y\varepsilon_x-k_z^2}\frac{d^2E_z}{dx^2}+\frac{k_o^2\varepsilon_y}{k_o^2\mu_x\varepsilon_y-k_z^2}\frac{d^2E_z}{dy^2}+k_o^2\varepsilon_z E_z=0 . \quad (30)$$

6. Suppression of Birefringence in a Rectangular Waveguide

Birefringence is a characteristic of anisotropic media where a single incident wave entering the boundary of an anisotropic medium gives rise to two refracted waves as shown in Figure 1 or a single incident wave leaving gives rise to two reflected waves as shown in Figure 2. We call these two waves the ordinary wave and the extraordinary wave. For low order modes, and especially the first mode, a rectangular waveguide suppresses the birefringence inherent to anisotropic media by suppressing propagation in the vertical direction. This suppression assumes that the dimensions of the waveguide are such that the horizontal dimension is at least twice the size of the vertical dimension [10]. To see how the geometry of the waveguide acts to cancel out the extraordinary wave, we need to solve for k_z from our dispersion equations.

Equations (26) and (30) yield the following solutions in unbounded anisotropic media restricted by the radiation condition (previously stated in section 3)

$$E_z(x,y,z)=E_o e^{-j\left(k_x x+k_y y+k_z z\right)}, \quad (31)$$

$$H_z(x,y,z)=H_o e^{-j\left(k_x x+k_y y+k_z z\right)} . \quad (32)$$

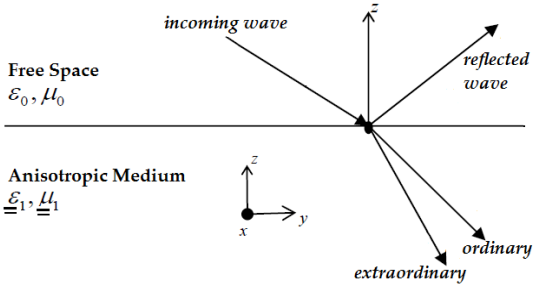


Figure 1: Free space plane wave incident on an anisotropic boundary [9].

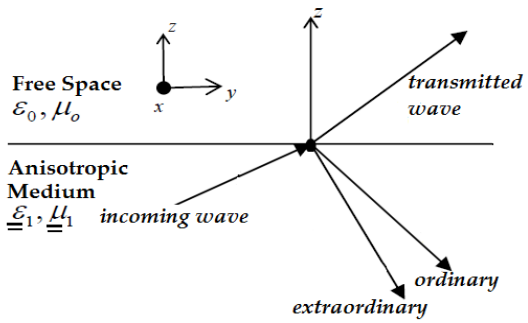


Figure 2: Anisotropic plane wave incident on a free space boundary [9].

Plugging (31) into (30) (equivalently we could substitute (32) into (26)) yields a polynomial equation whose solutions give the values of k_z in the anisotropic medium. Noting that $d^2/dx^2 = -k_x^2$ and $d^2/dy^2 = -k_y^2$, (30) simplifies to

$$\frac{k_o^2k_x^2\varepsilon_x E_z}{\left(k_o^2\mu_y\varepsilon_x-k_z^2\right)}+\frac{k_o^2\varepsilon_y k_y^2E_z}{\left(k_o^2\mu_x\varepsilon_y-k_z^2\right)}-k_o^2\varepsilon_z E_z=0 . \quad (33)$$

Dividing out the $k_o^2E_z$ term and multiplying through by both denominators gives us the following factored polynomial

$$\begin{gathered}\left(k_o^2\mu_y\varepsilon_x-k_z^2\right)\left(k_o^2\mu_x\varepsilon_y-k_z^2\right)\varepsilon_z-k_x^2\varepsilon_x\left(k_o^2\mu_x\varepsilon_y-k_z^2\right)\\ -\varepsilon_y k_y^2\left(k_o^2\mu_y\varepsilon_x-k_z^2\right)=0\end{gathered} . \quad (34)$$

Finally, multiplying out (34) yields a fourth order polynomial whose roots yield the four values of k_z describing the ordinary wave and extraordinary wave in the positive and negative propagation directions

$$\begin{gathered}k_z^4\mu_z+\left[k_x^2\mu_x+k_y^2\mu_y-(\varepsilon_x\mu_y+\varepsilon_y\mu_x)k_o^2\mu_z\right]k_z^2\\ +\left[k_o^4\varepsilon_x\varepsilon_y\mu_x\mu_y\mu_z-k_o^2k_x\varepsilon_x\mu_y\mu_x-k_o^2\varepsilon_y k_y\mu_x\mu_y\right]=0\end{gathered} . \quad (35)$$

Equation (35) is directly responsible for the existence of the two ordinary and extraordinary waves that are characteristic of the birefringence phenomenon. In an isotropic medium, the resulting polynomial for k_z is a second order polynomial, which yields only the values for the positive and negative propagation of the single ordinary wave. A fourth order polynomial allows for positive and negative propagation of both the ordinary and extraordinary wave shown in Figures 1 and 2.

The suppression of birefringence requires (35) to reduce to a second order polynomial. This is clearly not possible in an unbounded anisotropic medium as currently formulated. However, the boundary conditions in a rectangular waveguide assume that the first resonance suppresses the propagation constant in the vertical direction [10]. In other words, k_y=0 and d^2/dy^2=0, and this completely eliminates the second E_z term in (33). This simplifies (34) to

$$-\frac{k_o^2k_x^2\varepsilon_x}{k_o^2\mu_y\varepsilon_x-k_z^2}E_z+k_o^2\varepsilon_z E_z=0 , \quad (36)$$

which resembles the form of the traditional second order wave equation. Solving for k_z thus leads to the following second order polynomial

$$k_z^2 = \varepsilon_x \left(k_o^2 \mu_y - k_x^2 / \varepsilon_z \right). \quad (37)$$

Since (37) yields the positive and negative square root of the right-hand-side as the two solutions for k_z, this indicates the existence of only a single propagating wave. The suppression of the k_y term in the first resonant mode of the rectangular waveguide yields a second order differential equation for the wave number in the propagation direction, thereby eliminating the property of birefringence for this case. For a practical validation of this theory, the author would point the reader to reference [2]. This paper shows an antenna application that exploits the suppression of anisotropic birefringence to facilitate wideband impedance matching of a cavity backed antenna.

The reader should remember that (37) applies only to the first order resonance mode of the rectangular waveguide. As higher order modes are introduced propagation birefringence will exist for some modes depending on the boundary conditions. A similar analysis should be carried out for each propagating mode.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] G. Mitchell, "Suppression of anisotropic birefringence in a rectangular waveguide" in Proceedings of the Applied Computational Electromagnetics Society (ACES) Symposium, 2018. https://doi.org/10.23919/ROPACES.2018.8364258

[2] G. Mitchell and W. Wasylkiwskyj, "Theoretical anisotropic transverse resonance technique for the design of low profile wideband antennas" IET Microwaves, Antennas and Propagation, 10(5), 487-493, 2016. https://doi.org/10.1049/iet-map.2015.0470

[3] R. Collin, "A Simple Artificial Anisotropic Medium" IRE Transactions on Microwave Theory and Techniques, 6(4), 206-209, 1958. https://doi.org/10.1109/TMTT.1958.1125216

[4] M. Pardavi-Horvath, "Microwave Applications of Soft Ferrites" Journal of Magnetism and Magnetic Materials, 215-216(1), 171-183, 2000. https://doi.org/10.1016/S0304-8853(00)00106-2

[5] K. Buell, H. Mosallaei, K. Sarabandi, "A Substrate for Small Patch Antennas Providing Tunable Miniaturization Factors" IEEE Transactions on Microwave Theory and Techniques, 54(1), 135-146, 2006. https://doi.org/10.1109/TMTT.2005.860329

[6] D. Pozar, "Radiation and Scattering from a Microstrip Patch on a Uniaxial Substrate" IEEE Transactions on Antennas and Propagation, 35(6), 613-621, 1987. https://doi.org/10.1109/TAP.1987.1144161

[7] F.Y. Meng, Q. Wu, L. W. Li, "Transmission Characteristics of Wave Modes in a Rectangular Waveguide filled with Anisotropic Metamaterial" Applied Physics A: Materials Science and Processing, 94(4), 747-753, 2009. https://doi.org/10.1007/s00339-008-5057-2

[8] F.Y. Meng, Q. Wu, J.H. Fu, "Miniaturized Rectangular Cavity Resonator based on Anisotropic Metamaterials Bilayer" Microwave and Optical Technology Letters, 50(8), 2016-2020, 2008. https://doi.org/10.1002/mop.23556

[9] J. Graham, "Arbitrarily Oriented Biaxially Anisotropic Media: Wave Behavior and Microstrip Antennas," Ph.D Thesis, Syracuse University, 2012.

[10] D. Pozar, Microwave Engineering: 3rd edition, John Wiley and Sons, 2005.

[11] W.J. Ince, E. Stern, "Non-Reciprocal Remanence Phase Shifters in Rectangular Waveguide" IEEE Transactions on Microwave Theory and Techniques. 15(2), 87-95, 1967. https://doi.org/10.1109/TMTT.1967.1126381

[12] D. Torrent, J. Sanchez-Dehesa, "Radial Wave Crystals: Radially Periodic Structures from Anisotropic Metamaterials for Engineering Acoustic or Electromagnetic Waves" Physics Review Letters, 103(6), 2009. https://doi.org/10.1103/PhysRevLett.103.064301

[13] J. Sanchez-Dehesa, D. Torrent, J. Carbonell, "Anisotropic metamaterials as sensing devices in acoustics and electromagnetism" in Proceedings of the International Society for Optics and Photonics (SPIE), 2012. https://doi.org/10.1117/12.916043

[14] Y.G Ma, P. Wang, X. Chen, C.K. Ong, "Near-field plane-wave-like beam emitting antenna fabricated by anisotropic metamaterial" Applied Physics Letters, 94(4), 2009. https://doi.org/10.1063/1.3077128

[15] Q. Cheng, "Directive Radiation of Electromagnetic Waves Based on Anisotropic Metamaterials" in Proceedings of IEEE Asia-Pacific Conference on Antennas and Propagation, Singapore, Singapore, 2012.

[16] D. Schurig, J. Mock, B. Justice, S. Cummer, J. Pendry, A. Starr, D. Smith, "Metamaterial electromagnetic cloak at microwave frequencies" Science, 314(5801), 977-980, 2006. https://doi.org/10.1126/science.1133628

[17] J. Wong, K. Balmain, "A beam-steerable antenna based on the spatial filtering property of hyperbolically anisotropic metamaterials" in Proceedings of IEEE International Symposium of the Antennas and Propagation Society, Honolulu, HI, USA, 2007. https://doi.org/10.1109/APS.2007.4396438

[18] T. Cai, G.M. Wang, "Polarization-controlled Bifunctional Antenna based on 2-D Anisotropic Gradient Metasurface" in Proceedings of IEEE Conference on Microwave and Millimeter Wave Technology, Beijing, China, 2016. https://doi.org/10.1109/ICMMT.2016.7761687

[19] W. Zhong, X. Meng, Q. Hong, "Radiation power of half wave oscillator antenna in magnetic anisotropic medium" in 2018 International Conference on Electronics Technology (ICET), Chengdu, China, 2018. https://doi.org/10.1109/ELTECH.2018.8401446

[20] M. Born, E. Wolf, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light: 7th edition, Cambridge University Press, 1999.

[21] O. Ozgun, M. Kuzuoglu, "Finite element modeling of anisotropic half-space problems by a simple mesh truncation scheme" in 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, San Diego, CA, USA 2017. https://doi.org/10.1109/APUSNCURSINRSM.2017.807283

[22] J. Zhuo, F. Han, L. Ye, Z. Yu, Q.H. Liu, "Simulation of Electromagnetic Scattering of 3-D Inhomogeneous Biaxial Anisotropic Magnetodielectric Objects Embedded in Uniaxial Anisotropic Media by the Mixed-Order BCGS-FFT Method" IEEE Transactions on Microwave Theory and Techniques, 66(8), 3745-3755, 2018. https://doi.org/10.1109/TMTT.2018.2840984

[23] G. Mitchell, S. Weiss, "An Overview of ARL's Low Profile Antenna Work utilizing Anisotropic MetaFerrites", in Proceedings of the IEEE International Symposium on Phased Array Systems and Technology, Waltham, MA, USA, 2016. https://doi.org/10.1109/ARRAY.2016.7832581

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 3, 63-68 (2019)
www.astesj.com
Special Issue on Advanced Electrical and Communication Technologies

ASTES Journal
ISSN: 2415-6698

Novel Design of Multiband Microstrip Patch Antenna for Wireless Communication

Youssef Rhazi[*,1], Outman El Bakkali[2], Youssef El merabet[3], Mustpaha Ait lafkih[1], Seddik Bri[4], Mohamed Nabil Srifi[2]

[1]*Laboratory of Automatic, Energy Conversion and Microelectronics (LACEM), University of Sultan Moulay Slimane, Faculty of Sciences and Technology, B.P: 523 Beni-Mellal 23000, Morocco*

[2]*Electronics and Telecommunication Systems, Research Group National, ENSA, Kénitra, Morocco*

[3]*Laboratoire LASTID, Faculté des Sciences, Université Ibn Tofail, BP 133, 14000 Kénitra, Morocco*

[4]*Faculty of Sciences, Moulay Ismail University, Meknes, Morocco*

ARTICLEINFO

Article history:
Received: 09 January, 2019
Accepted: 07 May, 2019
Online: 23 May, 2019

Keywords:
Multiband patch antenna
Operating frequency 4.2 GHz
FR4-epoxy substrate
HFSS

ABSTRACT

This paper presents a novel six band frequency reconfigurable antenna for 2.4 GHz (Lower Worldwide Interoperability for Microwave Access (WiMAX)), 5.3 GHz (Wireless Local Area Network (WLAN)) and 9.1-10.2 GHz (X-band) frequency bands. The proposed antenna has a compact size of 22mm×30mm at lower resonance of 4.2 GHz and is printed on FR4 material with height h =1.6 mm, loss tangent δ= 0.02 and dielectric constant ϵ_r=4.4. Multiband phenomenon in the designed antenna is reached by inserting a circular hole inside a rectangular patch antenna and rectangular slots in the ground plane. During simulation, the designed antenna exhibits hexa band with S11<-10 dB bandwidth of about 4.76% (4.1- 4.3 GHz), 4.71% (5.21- 5.43 GHz), 16.27% (6.55-7.25 GHz), 1.83% (7.02-7.15 GHz), 0.87% (9.07-9.15 GHz) and 4.90% (10.02- 10.5 GHz) under simulation. We used HFSS (high frequency structured simulator) software for simulation of antennas and to find out the results. We keep changing the design of the antenna, as our objective is to achieve miniature antenna with better performance than traditional one.

1 Introduction

Current and future technology trends in wireless communication have increased the demand for patch antennas that can work at various bands with sufficient bandwidth. Microstrip antennas are one of the basic components required for wireless Communication. In the recent years, there has been a rapid and continuous growth in wireless communication. Nowadays there is a growing demand for efficient mobile device and good performance communication networks, thus requiring more efficiency in the antenna design. In recent years, active researches in wireless communication focus on reducing the number of antennas for a large variety of applications within a single system which becomes strongly recommended due to physical limits in the installation space. Hence, it is often desirable to design a single antenna that could work for various application. Such design requirements have encouraged antenna researchers to design a multiband antenna with the aim to improve the wireless quality and increase the application coverage [1]–[4]. It is noted that the discovery of the multiband nature constitutes one maiden revolution in the design of modern antennas. Besides covering various interesting and challenging applications, multiband antennas also find their success in the field of cost, size and high data rate features [1]. Multiband antennas present the ability to be easily integrated with control circuits and switching circuits while offering more excellent reconfiguration. In [1], the developed multiband patch antenna resonates for seven different frequencies, while covering X and C band frequencies. The authors in [5] designed an antenna with a compact size but which operates at four bands. Another four band antenna for wireless applications is developed in [3] but it has a O-shape multiband integrated

[*]Corresponding Author: Youssef Rhazi, email : y.rhazi@usms.ma

https://dx.doi.org/10.25046/aj040310

wideband monopole antenna. The multiband antenna developed in [6] delimits, likewise, the space requirement with a compact size of 20×20 mm^2. In [7], the authors proposed an antenna for WLAN/WiMAX applications which has a compact size of 30×28 mm^2. The proposed antenna is limited to four band of operations. The compact metamaterial antenna developed in [8] has a size of 25×22 mm^2 but limited for only triple band operations. The authors in [9], have designed a compact monopole patch antenna which has a size of 35×45 mm^2 but which also limited to triple band of operations. The authors in [10] and [11] have developed antennas with large size which are also limited to dual and triple band operation respectively.

Although multiband operation and frequency reconfigurability are provided by the above studied antennas, they still suffer from some critical limitations in terms of number of operating bands, size and antenna gain. This paper aims at developing a compact multiband reconfigurable antenna such that it can be readily integrated with switching circuit, wireless end terminal devices. The proposed antenna design, which operates at six active bands of as 4.2, 5.3, 6.8, 7.1, 9.1 and 10.2 GHz with a compact size of only 28×30 mm^2, consists of a circular slot in the radiating patch. The slotting technique is employed to alter the surface current path which makes the proposed antenna to operate at six bands. The designed antenna, which resonates for 6 different frequencies which covers modern wireless services such as Bluetooth and WLAN (Wireless Local Area Network) frequencies, utilizes microstripline feed with double stubs to yield multiband such as 4.1- 4.3 GHz, 5.21- 5.43 GHz, 6.55-7.25 GHz, 7.02-7.15 GHz, 9.07-9.15 GHz and 10.02- 10.5 GHz. In order to produce the source signal, we have employed, like the works presented in [11]–[17], an insert fed method. As will be shown latter, the proposed antenna design is able to achieve, for all resonant frequencies, a reflected power which is less than -14 dB. Furthermore, the simulated results showed that the proposed antenna has VSWR which varies in the range [1-2], and very good performance in terms of directivity and gain are achieved. Our design procedure, which uses both transmission-line and cavity models, was fine-tuned by the full-wave model using HFSS (Ansoft High Frequency Structural Simulator) version 13.

2 Antenna Design

A basic multiband patch antenna is initially developed, which enables independent control in multi bands with good radiation properties and matching while using a simple geometry and feeding technique. However, it is limited to four band of operation and simulated results show that it presents impedance matching is of poor quality. To deal with this weakness and thus to attain improved impedance matching in the entire multiband, a new compact multiband reconfigurable antenna is designed, in which a circular hole is added inside the patch with different values of radius R.

The ground plane has dimensions (Ls×Ws) and it is separated from the coupled element by a gap h. For a regular rectangular patch without slot [14]–[18], its resonant frequency of TM_{mn} mode is given by:

$$f_{mn} = \frac{C}{2\sqrt{\mu_r \epsilon_r}} \sqrt{(\frac{m}{W})^2 + (\frac{n}{L})^2} \quad (1)$$

where C is the light velocity in free space, μ_r is the equivalent permeability and ϵ_r is the equivalent permittivity. By selecting the feed location, the first two modes TM_{10} and TM_{11} can be excited in the study. We found that by increasing h, the resonant frequency shifts to low frequency and the bandwidth of low frequency band becomes narrow while the bandwidth of high frequency band becomes broad. The antenna geometry is shown in Figures 1 and 2. Detailed dimensions are listed in Table 1.

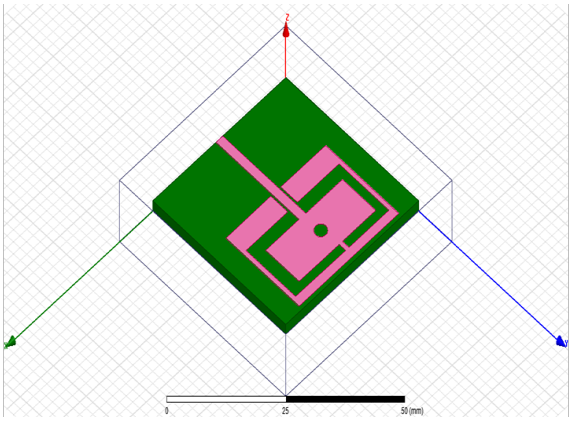


Figure 1: Proposed microstrip patch antenna.

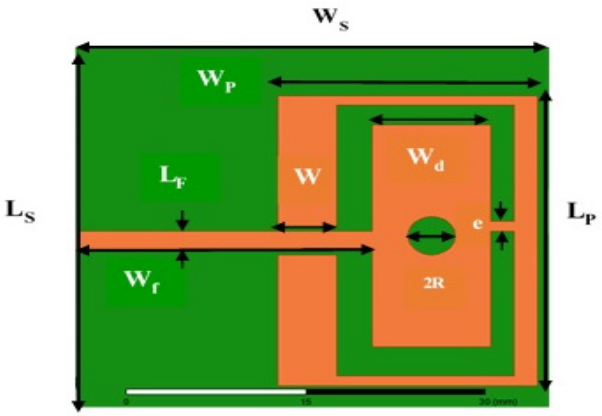


Figure 2: Design of the proposed microstrip patch antenna.

3 Simulation Results

The reflection coefficient is a parameter which is used to quantify how much of an electromagnetic wave is reflected back at antenna terminals in the transmission medium. An S11 value (cf. Eq. 2) is measured in dB and is negative, and expresses the ratio of reflected

power (P_{ref}) to incident one (P_{in}) at port 1, if S11<-10 dB then 90% of power excited is transmitted.

Table 1: Proposed antenna parameters (mm).

Parameters	Values (mm)
Patch length (Lp)	30
Patch width (Wp)	22
Feed width (Lf)	2
Feed length (Wf)	25
Ground length (Ls)	40
Ground width (Ws)	40
Height (h)	1.6
e	1
R	0.75

$$S_{11}(dB) = 20log(\frac{P_{ref}}{P_{in}}) \quad (2)$$

Figure 3 shows the result of the comparison study for different values of the parameter R. It is apparent from Figure 3 that, on the one hand, the widest frequency of multiband antenna is getting for R=0.75 mm and on the other hand, a very low return loss can be recorded at the frequency of 6.8 GHz. However, simulated results depict that a poor impedance matching in the upper Ultra Wideband range from 9.07-9.15 GHz and 10.02-10.5 GHz is obtained.

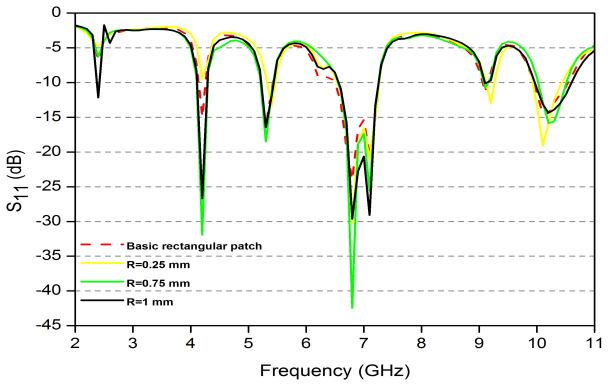


Figure 3: Return loss of proposed patch antenna for different values of R.

Figure 4 illustrates the performance of VSWR for the proposed microstrip patch antenna, which lies between 1 and 2 for all resonant frequencies with minimum reflected power which is inferior than -14 dB. The conclusion derived from the analysis of both Fig. 3 and 4, i.e., simulated results of return loss and VSWR respectively, confirms that the designed multiband patch antenna ensures obtaining good performance.

The simulation results of gain of the proposed microstrip patch antenna are shown in Figure 5. The results emerged from Figure 5 confirm the fact that the first 3 resonant frequencies, i.e., 4.1- 4.3 GHz, 5.21-5.43 GHz and 6.55-7.25 GHz, are approximately omni-directional pattern while the remaining ones, i.e., 7.02-7.15 GHz, 9.07-9.15 GHz and 10.02- 10.5 GHz. are directional pattern. Table 2 reports the obtained performances of bandwidth, return loss, VSWR and Gain, under several resonant frequencies.

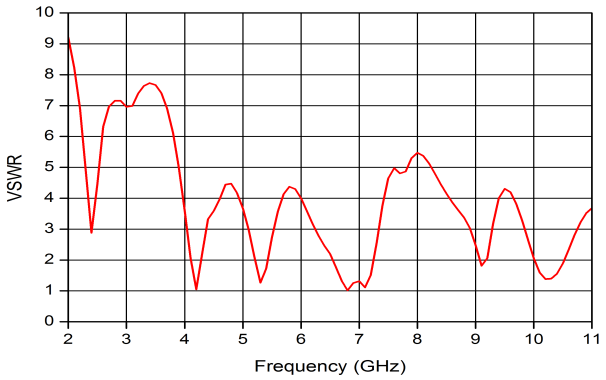


Figure 4: Simulated results of proposed patch antenna(R=0.75 mm): VSWR.

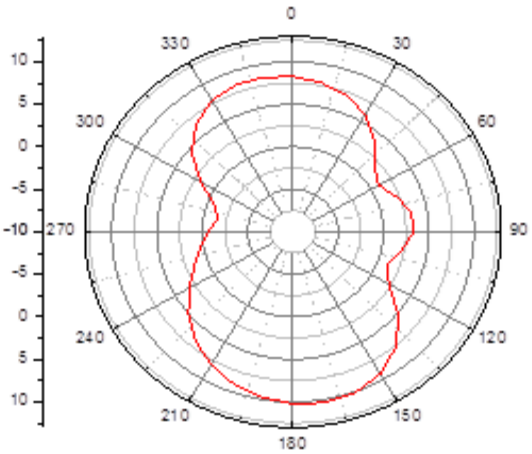


Figure 5: Gain of proposed patch antenna (R=0.75 mm).

Figure 6 shows the 3D radiation pattern for different resonance frequencies : (a) for 4.2 GHz, (b) for 5.3 GHz, (c) for 6.8 GHz, (d) for 7.1 GHz, (e) 9.1 GHz and (f) for 10.2 GHz. As can be seen, simulated results show similar uni-directional radiation patterns and very low cross polarization level. At the six frequencies, the radiation pattern achieves the front-back ratio which is superior than 10 dB.

In order to further analyze the effectiveness of the proposed antenna design, the current distribution is also investigated. Figure 7 illustrates the current distribution simulated on the radiating element for different resonance frequencies : (a) for 4.2 GHz, (b) for 5.3 GHz, (c) for 6.8 GHz, (d) for 7.1 GHz, (e) for 9.1 GHz and (f) for 10.2 GHz. It is found that due to the insertion a circular hole inside the rectangular patch antenna and rectangular slots in the ground plane, surface current density is varied significantly. It is flowed an entire

Table 2: Bandwidth, Return loss, VSWR and Gain for various resonant frequencies.

Resonance frequency (GHz)	Bandwidth (%)	Return loss (dB)	VSWR	Gain (dB)
4.2	4.76	-31,90	1.05	0.25
5.3	4.71	-18.49	1.27	0.43
6.8	16.27	-42.45	1.01	2.73
7.1	1.83	-25.52	1.11	1.98
9.1	0.87	-10.76	1.81	1.77
10.2	4.9	-15.82	1.38	1.63

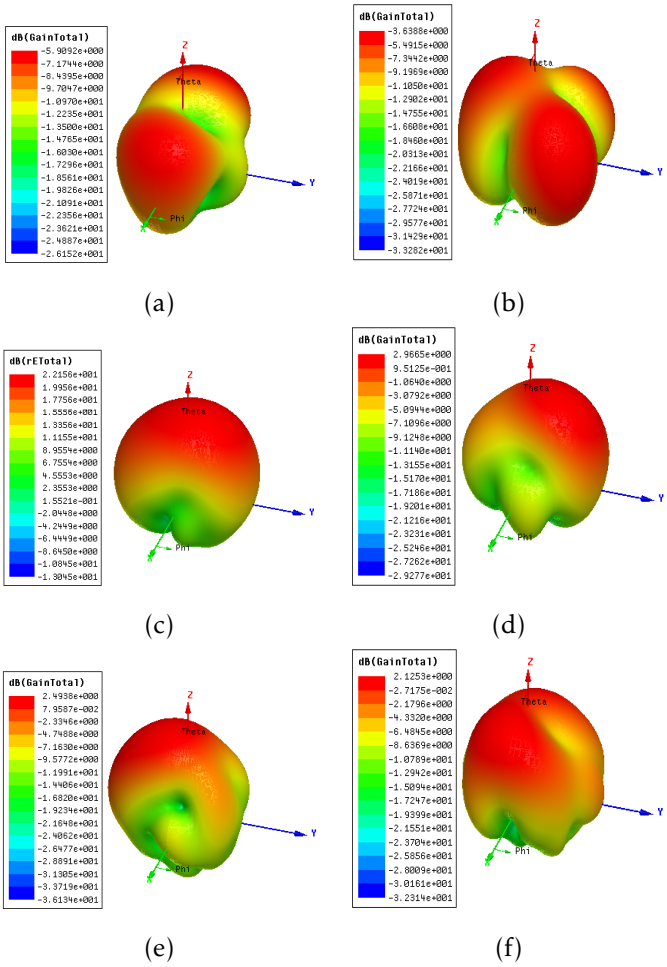


Figure 6: The radiation pattern in 3D of proposed patch antenna (R=0.75 mm) at (a) 4.2 GHz, (b) 5.3 GHz, (c) 6.8 GHz, (d) 7.1 GHz, (e) 9.1 GHz and (f) 10.2 GHz

patch at 4.20 GHz and 6.8 GHz. It also emerges that it mainly flows on feed line and notch cut of the patch antenna at 5.3 GHz, 7.1 GHz and 11.47 GHz. It flows on edge and feed line of the proposed patch antenna at 9.1 GHz and 10.2 GHz. In addition, it can be observed that the current density is more near the edge rectangular slot WxLp as illustrated in Fig. 7.(a) and (b), which exhibits the resonance at 4.2 and 5.3 GHz, respectively. At these frequencies, the current in the front patch is employed to adjust these resonances. Near the microstrip line, higher current density is observed which is responsible for the resonance at 6.8 GHz and 7.1 GHz, respectively, as illustrated in Fig. 7.(c) and (d). The larger current path near the slots WxLp and microstrip line exhibit the resonance at 10.2 GHz, as depicted in Fig. 7.(e) and (f). Note that, the introduction of the circular hole inside a rectangular patch antenna modifies the surface current distribution of the proposed design, so that the total current length path increases, allowing the designed antenna

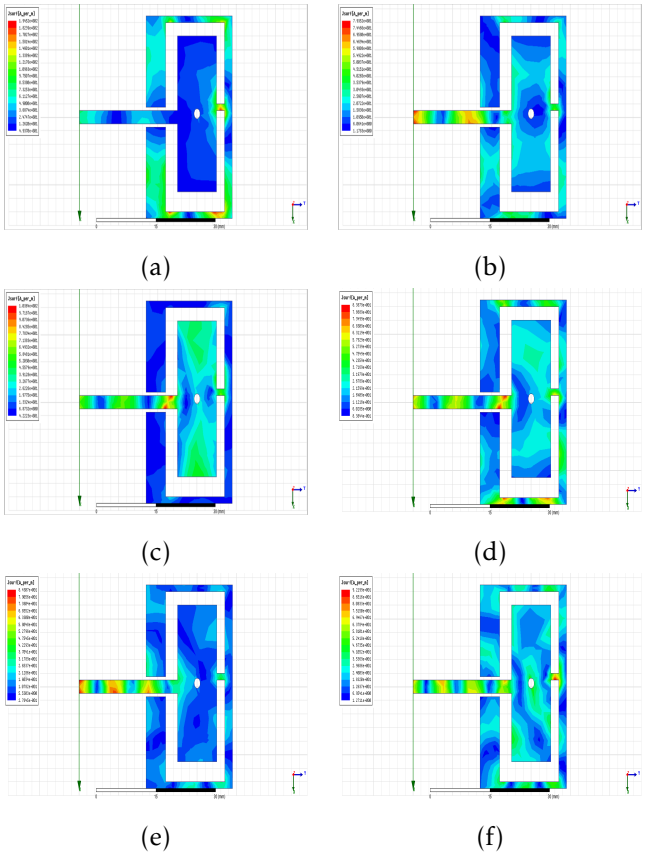


Figure 7: Simulated results of the surface current distribution for the antenna at (a) 4.2 GHz, (b) 5.3 GHz, (c) 6.8 GHz, (d) 7.1 GHz, (e) 9.1 GHz and (f) 10.2 GHz.

to operates at 4.2 GHz, 5.3 GHz, 6.8 GHz, 7.1 GHz, 9.1 GHz and 10.2 GHz frequency bands respectively, as illustrated in Fig. 7.

4 Conclusion

In this work, we designed a new multiband microstrip patch antenna for wireless communication. The proposed antenna covers multiple frequencies (i.e., the frequency range between 2 GHz and 11 GHz). Simulation results showed that the return loss for all resonant frequencies is less than -14 dB and the proposed design achieves omnidirectional and bi-directional radiation pattern. Furthermore, the achieved peak gain is superior than 5 dB and the simulation results provide the better outcome for wireless communication.

References

[1] N. Prema, Anil kumar., "Design of multiband microstrip patch antenna for C and X band", Optik **127** (2016) 88128818.

[2] Z.H. Li, Y.L. Xue, Z.Q. Deng, et al., "Study on optical switching effect of photonic crystals with negative effective index of refraction", Optik 120 **12**(2009) 605609.

[3] K.Srivastava, A.Kumar, et al., "Multiband Integrated Wideband Antenna for Bluetooth/WLAN Applications", International Journal of Electronics and Communications, S1434-8411(18)30036-0.

[4] Ji-jun Wang, Zhi-pan Zhu, Yu-xin Sun, Ting-gen Shen, Lei-lei Gong, "Study on left-handed effect of composite helices and its application in square frame patch antennas", Optik **124** (2013) 51895192.

[5] Chun-Xu Mao, Steven Gao et al., "A Novel Multiband Directional Antenna for Wireless Communications", OI 10.1109/LAWP.2016.2628715, IEEE Antennas and Wireless Propagation Letters, 2016.

[6] Khalid Hati et al., "A Novel Multiband Patch Antenna Array for Satellite Application", INTER-ENG 2016, 181(2017) 496 -502.

[7] T. Ali, M. Muzammil Khaleeq, R.C. Biradar, "A Multiband Reconfigurable slot antenna for Wireless Applications", International Journal of Electronics and Communications (2017), doi: https://doi.org/10.1016/j.aeue.2017.11.033.

[8] T. Ali, Biradar R.C. "A compact multiband antenna using rectangular stub loaded with metamaterial for IEEE 802.11N and IEEE 802.16E", Microw Opt Technol Lett 2017; **59**(4):1000-1006.

[9] Ting Wu, Xiao-Wei Shi, Ping Li, HaoBai. "Tri-band microstrip-fed monopole antenna with dual polarisation characteristics for WLAN and WiMAX applications", Electron Lett 2013;**49**(25):1597-1598.

[10] JaswinderKaur, Rajesh Khanna. "Development of dual-band microstrip patch antenna for WLAN/MIMO/WIMAX/AMSAT/WAVE applications", Microw Opt. TechnolLett2014;**56**(4):988-993.

[11] Shan Shan Huang, Jun Li, Jian Zhong Zhao. "Design of a compact triple-band monopole planar antenna for wlan/wimax applications", Prog Electromagn Res C 2014; 48:29-35.

[12] M. Salehi, A. Tavakoli, "A novel low mutual coupling microstrip antenna array design using defected ground structure", Int. J. Electron. Comm. **60** (10) (2006) 718723.

[13] F.Y. Zulkifli, E.T. Rahardjo, D. Hartanto, "Mutual coupling reduction using dumbbell defected ground structure for multiband microstrip antenna array", Prog. Electromagn. Res. Lett. **13** (2010) 2940.

[14] M.S. Alam, M. Tariqul, H. Arshad, "Gain enhancement of a multiband resonator using defected ground surface on epoxy woven glass material", Sci. World J. **6** (2014), 159468 8, http://dx.doi.org/10.1155/2014/159468.

[15] J. Pei, A.-G. Wang, S. Gao, W. Leng, "Miniaturized triple-band antenna with a defected ground plane for WLAN/WiMAX applications", IEEE Antennas Wireless Propag. Lett. **10** (2011) 298301.

[16] K. He, S. Gong, F. Gao, Low-profile wideband unidirectional patch antenna with improved feed structure, Electron. Lett. **51** (4) (2015) 317319.

[17] Youssef Rhazi et al., "Effect of Microstrip Antenna Feeding in the K-band",International Journal of Engineering and Technology, **4**(6): 515-522, 2013, ISSN 0975-4024.

[18] Verma, A. K. and Z. Rostamy, "Resonance frequency of uncovered and covered rectangular microstrip patch using modified Wolff model," IEEE Trans. Microwave Theory Tech., **41**, 109-116, 1993.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 69-74 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Multi-Step Iteration Algorithm of Total Asymptotically Quasi-Nonexpansive Maps

Salwa Salman Abed*, Zahra Mahmood Mohamed Hasan

Department of Mathematics, college of Education for Pure Sciences (Ibn Al-Haitham) / University of Baghdad, Baghdad, Iraq

ARTICLE INFO

Article history:
Received: 14 February 2019
Accepted: 28 April, 2019
Online: 21 May, 2019

Keywords:
Banach space,
total asymptotically quasi-nonexpansive map,
weak convergence,
strong convergence,
common fixed points

ABSTRACT

In Banach spaces an iteration algorithm for two finite families of total asymptotically quasi-nonexpansive maps is introduced. Weak and strong convergence theorems of this algorithm to approximation common fixed points are proved by using suitable conditions. As well as, numerical example by using Mat-lab is given.

1. Introduction and Preliminaries

This paper was originally published in the Conference: 2018 International Conference on Advanced Science and Engineering (ICOASE), Iraq [1]. It is well known that the concept of asymptotically nonexpansive introduced by Goebel and Kirk [2]. Additionally, every asymptotically nonexpansive map of a Banach space has a fixed point is proved. In [3], Petryshyn and Williamson proved the weak and strong convergence for quasi-nonexpansive map by using a sufficient and necessary condition. Alber [4], a new class of asymptotically nonexpansive is introduced. As well as, approximating methods for finding their fixed points are studied. In 2014, G. S. Saluja [5] established the strong and weak convergence for approximating common fixed point for generalized asymptotically quasi-nonexpansive maps in a Banach space.

Very recently, In [6], the authors proposed an implicit iteration for two finite families of generalized asymptotically quasi-nonexpansive maps. As well as, some strong convergence theorems are established. It is useful to point out our findings in this area which appeared in [7].

Let B be a non-empty closed convex subset of a real Banach space M and T be a self-map of B. The set of all fixed points denoted by F(T). A self-map T from B into M is called nonexpansive map [2] if

$$\|Ta - Tb\| \leq \|a - b\| \; for \; all \; a, b \in B$$

and is called quasi- nonexpansive map [6] if $F(T) \neq \emptyset \; and$

$$\|Ta - a^*\| \leq \|a - a^*\|$$

for all $a \in B$ and for all $a^* \in F(T)$.

A Banach space M is satisfying:

- "Opial's condition if for each sequence (a_n) in M, is weak convergence to a implies that $\lim_{n\to\infty} inf\|a_n - a\| < \lim_{n\to\infty} inf \; \|a_n - b\|$ for all $b \in M \; with \; a \neq b$".
- "Kadec-Klee property if for each sequence (a_n) in M is weak convergence to (a) together with $\|a_n\|$ converges strongly to $\|a\|$ imply that (a_n) is strong convergence to a point $a \in M$ [7]".

The aim of this paper, an iterative scheme for two families of total asymptotically quasi-nonexpansive maps is established. The strong and weak convergence theorems of this scheme for approximation of common fixed points in Banach space by using suitable conditions are established. For this purpose, let us recall the following definitions and lemmas.

Definition (1.1): "A map T is named asymptotically nonexpansive [1] if there is a sequence (f_n) in $[0, +\infty)$ with $\lim_{n\to\infty} f_n = 0$ and $\|T^n a - T^n b\| \leq (1 + f_n)\|a - b\|$, for all $a, b \in B, n = 1,2,..$.

*Salwa Salman Abed, Email: salwaalbundi@yahoo.com

If $F(T) \neq \emptyset$ there is a sequence (f_n)in $[0, +\infty)$ with $\lim_{n\to\infty} f_n = 0$ and $\|T^n a - a^*\| \leq (1 + f_n)\|a - a^*\|$,
for all $a \in B, a^* \in F(T)$ and $n = 1, 2, \ldots$.Therefore, T is named asymptotically quasi-nonexpansive map [10]".

Definition (1.2): "A map T is named total asymptotically nonexpansive map [4] if there are null sequences of postive real numbers $(f_n)_{n=1}^{\infty}, (g_n)_{n=1}^{\infty}, n \geq 1$ and nondecreasing continuous function $\psi: [0, \infty) \to [0, \infty)$ with $\psi(0) = 0$ such as $\forall a, b \in B$

$$\|T^n a - T^n b\| \leq \|a - b\| + f_n \psi\|a - b\| + g_n".$$

T is named total asymptotically quasi-nonexpansive map if $F(T) \neq \emptyset$, there are null sequences of postive real numbers $(f_n)_{n=1}^{\infty}$, $(g_n)_{n=1}^{\infty}, n \geq 1, \sum_{n=1}^{\infty} f_{in} < \infty$ and $\sum_{n=1}^{\infty} g_{in} < \infty$, and nondecreasing continuous function $\psi: [0, \infty) \to [0, \infty)$ with $\psi(0) = 0$ such as $\forall a \in B$ and $a^* \in F(T)$
$\|T^n a - a^*\| \leq \|a - a^*\| + f_n \psi\|a - a^*\| + g_n$.
If $g_n = 0, \forall n = 1, 2, \ldots$ therefore T is asymptotically quasi-nonxpansive map.

Definition (1.3)[11]: "Let B be a nonempty closed convex subset of a Banach space M. A self-map T is named uniformily K-Lipschitzain if there exists a constant $K > 0$ such that $\|T^n a - T^n b\| \leq K\|a - b\|, \ \forall a, b \in B$.

Definition (1.4)[12]: "A map $T: B \to M$ is named demi-closed with respect to $b \in M$ if for each sequence (a_n) in B, (a_n) is weak convergence to a and $T(a_n)$ is strong convergence to b. Hence $a \in B$ and $T(a) = b$. If $(I - T)$ is demiclosed which means if (a_n) is weak convergence to a in B and $(I - T)$ is strong convergence to 0. Therefore $(I - T)(a) = 0$".

Note: Now to explain the relation between the above definitions:

$$\begin{array}{ccccc} Nonexpansive & \Rightarrow & Asymptotically\ nonexpansive & \Rightarrow & Total\ asymptotically\ nonexpansive \\ \Downarrow & & \Downarrow & & \Downarrow \\ Quasi-nonexpansive & \Rightarrow & Asymptotically\ quasi-nonexpansive & \Rightarrow & Total\ asymptotically\ quasi-nonexpansive \end{array}$$

Lemma (1.5)[13]:" Let M be a uniformly convex Banach space and $0 < L \leq t_n \leq K < 1, \forall\, n \in N$. Presume that (a_n) and (b_n) are two sequences of M such as:

$$\lim_{n\to\infty} \|a_n\| \leq r, \lim_{n\to\infty} \|b_n\| \leq r \ \ and \ \ \lim_{n\to\infty} \|t_n a_n + (1 - t_n) b_n\| = r$$

hold for some $r \geq 0$. Thus $\lim_{n\to\infty} \|a_n - b_n\| = 0$".

Lemma (1.6)[14]: "Let $(\mu_n)_{n=1}^{\infty}, (\sigma_n)_{n=1}^{\infty}$ and $(e_n)_{n=1}^{\infty}$ be sequences of postive numbers accomplishing the following inequality:

$$\mu_{n+1} \leq (1 + \sigma_n)\mu_n + e_n \ , \forall\, n \geq 1$$

if $\sum_{n=1}^{\infty} \sigma_n < \infty$ and $\sum_{n=1}^{\infty} e_n < \infty$, then $(\mu_n)$is bounded and $\lim \mu_n$ exists. In additional if, $\liminf_{n\to\infty} \mu_n = 0$ then $\lim_{n\to\infty} \mu_n = 0$".

Lemma (1.7)[15]: "Let B be a nonempty convex subset of a uniformly convex Banach space. Therefore there exists a strictly nondecreasing continuous function $f: [0, \infty) \to [0, \infty)$ with $f(0) = 0$ such as for each Lipschitzain map $T: B \to B$ with Lipschitz constant K:

$$\|tTa + (1 - t)Tb - T(ta + (1 - t)b\| \leq Kf^{-1}\left(\|a - b\| - \frac{1}{K}\|Ta - Tb\|\right),$$

$\forall\, a, b \in B$ and $\forall t \in [0,1]$".

Lemma (1.8)[12]: "Let M be a uniformly convex Banach space and its dual M^* accomplishing the Kadec-Klee property. Presume that (a_n) bounded sequence in M such as $\lim_{n\to\infty} \|ta_n + (1 - t)p_1 - p_2\|$ exists $\forall\, t \in [0,1]$ and $p_1, p_2 \in W_w(a_n)$, thus $p_1 = p_2$".

2. Main Results

Let B be a nonempty closed convex subset of a Banach space M and $\{T_j, S_j, \forall j = 1, 2, \ldots, k\}$ be two families of total asymptotically quasi-nonexpansive self-maps. We define the iteration algorithm (a_n) as follows:

$$\begin{aligned} &a_1 \in B \\ &a_{n+1} = (1 - \alpha_{jn})S_j^n a_n + \alpha_{jn} T_j^n b_{jn} \\ &b_{jn} = (1 - \alpha_{jn})S_j^n a_n + \alpha_{jn} T_j^n b_{(j-1)n} \\ &b_{(j-1)n} = (1 - \alpha_{(j-1)n})S_{j-1}^n a_n + \alpha_{(j-1)n} T_{j-1}^n b_{(j-2)n} \\ &\cdot \\ &\cdot \\ &\cdot \\ &b_{2n} = (1 - \alpha_{2n})S_2^n a_n + \alpha_{2n} T_2^n b_{1n} \\ &b_{1n} = (1 - \alpha_{1n})S_1^n a_n + \alpha_{1n} T_1^n b_{0n} \end{aligned} \quad (1)$$

Where $b_{0n} = a_n$ and $(\alpha_n)_{n=1}^{\infty}$ are sequences in [0, 1].

Lemma (2.1): Let B be a nonempty closed convex subset of a normed space M and $T_j, S_j, j = 1, 2, \ldots, k$ be two family of total asymptotically quasi-nonexpansive self-maps of B. Presume that $F(T_j, S_j) \neq \emptyset$ and the sequence (a_n) be as shown in step (1). Then:
i- There are sequences (u_n) and (v_n) in $[0, \infty)$ such as $\sum_{n=1}^{\infty} u_n < \infty, \sum_{n=1}^{\infty} v_n < \infty$ and
$\|a_{n+1} - a^*\| \leq (1 + u_n)^{j+1}\|a_n - a^*\| + v_n^{j+1}$,
$\forall\, a^* \in F(T_j, S_j)$ and $\forall\, n$.
ii- There exist constants $J_1, J_2 > 0$ such as
$\|a_{n+p} - a^*\| \leq J_1\|a_n - a^*\| + J_2$, $\forall\, a^* \in F(T_j, S_j)$
and $n, p = 1, 2, \ldots$..
If there is $Z > 0$ such that $\psi(\lambda_j) \leq Z\lambda_j, j = 1, 2, \ldots, k$.
Proof: i- Let $a^* \in F, u_n = \max_{1\leq j\leq k} f_{jn}$ and $v_n = \max_{1\leq j\leq k} g_{jn}$.

Now, we have

$$\begin{aligned} \|b_{1n} - a^*\| &= \|(1 - \alpha_{1n})S_1^n a_n + \alpha_{1n} T_1^n a_n - a^*\| \\ &\leq (1 - \alpha_{1n})\|S_1^n a_n - a^*\| + \alpha_{1n}\|T_1^n a_n - a^*\| \\ &\leq (1 - \alpha_{1n})\{\|a_n - a^*\| + f_{1n}\psi\|a_n - a^*\| + g_{1n}\} \\ &\quad + \alpha_{1n}\{\|a_n - a^*\| + f_{1n}\psi\|a_n - a^*\| + g_{1n}\} \\ &\leq (1 - \alpha_{1n})(1 + f_{1n}Z)\|a_n - a^*\| + (1 - \alpha_{1n})g_{1n} \\ &\quad + \alpha_{1n}(1 + f_{1n}Z)\|a_n - a^*\| + \alpha_{1n} g_{1n} \\ &\leq (1 + f_{1n}Z)\|a_n - a^*\| + g_{1n} \\ &\leq (1 + u_n)\|a_n - a^*\| + v_n \end{aligned}$$

Assume that $\|b_{jn} - a^*\| \leq (1 + u_n)^j\|a_n - a^*\| + v_n^j$
Therefore,

$$\|b_{(j+1)n} - a^*\| = \|(1 - \alpha_{(j+1)n})S_{j+1}^n a_n + \alpha_{(j+1)n} T_{j+1}^n b_j - a^*\|$$

$$\leq (1-\alpha_{(j+1)n})\|S_{j+1}^n a_n - a^*\| + \alpha_{(j+1)n}\|T_{j+1}^n b_j - a^*\|$$
$$\leq (1-\alpha_{(j+1)n})(1+f_{(j+1)n}Z)\|a_n - a^*\| + (1-\alpha_{(j+1)n})g_{(j+1)n} + \alpha_{(j+1)n}(1+f_{(j+1)n}Z)\|b_j - a^*\| + \alpha_{(j+1)n}g_{(j+1)n}$$
$$\leq (1-\alpha_{(j+1)n})(1+u_n)\|a_n - a^*\| + v_n + \alpha_{(j+1)n}(1+u_n)(1+u_n)^j\|a_n - a^*\| + \alpha_{(j+1)n}(1+u_n)v_n^j$$
$$\leq (1+u_n)^{j+1}\|a_n - a^*\| + v_n^{j+1}$$

Thus, by induction, we obtain

$$\|b_{jn} - a^*\| \leq (1+u_n)^j\|a_n - a^*\| + v_n^j \quad (2)$$

for all $j = 1,2,\dots,k$.

Now, by (2), we get

$$\|a_{n+1} - a^*\| \leq (1-\alpha_{jn})\|S_j^n a_n - a^*\| + \alpha_{jn}\|T_j^n b_{jn} - a^*\|$$
$$\leq (1-\alpha_{jn})(1+u_n)\|a_n - a^*\| + v_n + \alpha_{jn}(1+u_n)(1+u_n)^j\|a_n - a^*\| + \alpha_{jn}(1+u_n)v_n^j$$
$$\leq (1+u_n)^{j+1}\|a_n - a^*\| + v_n^{j+1}$$

ii- By using part (i), we get

$$\|a_{n+p} - a^*\| \leq (1+u_{n+p-1})^{j+1}\|a_{n+p-1} - a^*\| + v_{n+p-1}^{j+1}$$
$$\leq e^{(1+u_{n+p-1})^{j+1}\|a_{n+p-1}-a^*\|} + e^{v_{n+p-1}^{j+1}}$$
$$\leq e^{(j+1)u_{n+p-1}\|a_{n+p-1}-a^*\|} + e^{(j+1)v_{n+p-1}}$$
$$\leq e^{(j+1)\sum_{k=1}^{n+p-1}u_k}\|a_n - a^*\| + e^{(j+1)\sum_{k=1}^{n+p-1}v_k}$$
$$\leq J_1\|a_n - a^*\| + J_2.$$

Setting $J_1 = e^{(j+1)\sum_{k=1}^{n+p-1}u_k}$ and $J_2 = e^{(j+1)\sum_{k=1}^{n+p-1}v_k}$.

Lemma (2.2): Let B be a nonempty closed convex subset of a normed space M and T_j, S_j , $j = 1, 2, \dots, k$ be two families of total asymptotically quasi-nonexpansive self-maps of B. Presume $F(T_j, S_j) \neq \emptyset$ and (a_n) be as shown in step (1). Therefore, $\lim_{n\to\infty}\|a_n - a^*\|$ exists for all $a^* \in F(T_i, S_i)$.

Proof: By Lemma (2.1.i)

$$\|a_{n+1} - a^*\| \leq (1+u_n)^{j+1}\|a_n - a^*\| + v_n^{j+1}$$
$$\leq (1+u_n)\|a_n - a^*\| + v_n$$

and $\sum_{n=1}^{\infty}u_n < \infty, \sum_{n=1}^{\infty}v_n < \infty$. So by Lemma (1.6.i), we get $\lim_{n\to\infty}\|a_n - a^*\|$ exists for all $a^* \in F(T_j, S_j)$.

Lemma (2.3): Let B be a nonempty closed convex subset of a Banach space M and T_j, S_j , $j = 1, 2, \dots, k$ be two families of Lipschitzain and total asymptotically quasi-nonexpansive self-maps of B. Let (a_n) be as shown in step (1). Therefore, for all $a_1^*, a_2^* \in F(T_j, S_j)$, the limit $\lim_{n\to\infty}\|ta_n + (1-t)a_1^* - a_2^*\|$ exists for all $t \in [0,1]$.

If there is $Z > 0$ such that $\psi(\lambda_j) \leq Z\lambda_j, j = 1, 2, \dots, k$.

Proof: By using Lemma (1.6),we have $\lim_{n\to\infty}\|a_n - a^*\|$ exists $\forall a^* \in F(T_j, S_j)$ and (a_n) is bounded. Let $\gamma_n(t) = \|ta_n + (1-t)a_1^* - a_2^*\|$, $\forall t \in [0,1]$.

Therefore, $\lim_{n\to\infty}a_n(0) = \|a_1^* - a_2^*\|$ and $\lim_{n\to\infty}a_n(1) = \|a_n - a_2^*\|$ exist by Lemma (2.2).

Then, for $t \in [0,1]$ and for all $a \in B$, we define the map $R_n: B \to B$ by:

$$b_{1n} = (1-\alpha_{1n})S_1^n a_n + \alpha_{1n}T_1^n b_{0n}$$
$$b_{2n} = (1-\alpha_{2n})S_2^n a_n + \alpha_{2n}T_2^n b_{1n}$$
$$\vdots$$
$$b_{jn} = (1-\alpha_{jn})S_j^n a_n + \alpha_{in}T_j^n b_{(j-1)n}$$
$$R_n a = (1-\alpha_{jn})S_j^n a_n + \alpha_{jn}T_j^n b_{jn}$$

Now,

$$\|R_n a - R_n c\| \leq (1-\alpha_{jn})\|S_j^n a - S_j^n c\| + \alpha_{jn}\|T_j^n b_{jn} - T_j^n d_{jn}\|$$
$$\leq (1-\alpha_{jn})(1+f_{jn}Z)\|a - c\| + (1-\alpha_{jn})g_{jn} + \alpha_{jn}(1+f_{jn}Z)\|b_{jn} - d_{jn}\| + \alpha_{jn}g_{jn}$$
$$\leq (1+u_n)\|a - c\| + v_n$$
$$\leq (1+f_nZ)\|a - b\| + g_n$$

with $\sum_{n=1}^{\infty}u_n < \infty$, $\sum_{n=1}^{\infty}v_n < \infty$ and $s_n = 1 + u_n$, it follows that $s_n \to 1$ as $n \to \infty$.

Setting $W_{n,m} = R_{n+m-1}R_{n+m-1}\dots R_n$ and $b_{n,m} = \|W_{n,m}(ta_n + (1-t)a_1^* - (tW_{n,m}a_n + (1-t)a_1^*)\|$.

Thus,

$$\|W_{n,m}a - W_{n,m}b\| = \left\| \begin{matrix} R_{n+m-1}R_{n+m-2}\dots R_n(a) \\ -R_{n+m-1}R_{n+m-2}\dots R_n(b) \end{matrix} \right\|$$
$$\leq s_{n+m-1}\left\| \begin{matrix} R_{n+m-2}\dots R_n(a) \\ -R_{n+m-2}\dots R_n(b) \end{matrix} \right\| + v_{n+m-1}$$
$$\leq \prod_{j=n}^{n+m-1}s_j\|a - c\| + \sum_{j=n}^{n+m-1}v_j$$
$$= A_n\|a - c\| + \sum_{j=n}^{n+m-1}v_j$$

for all $a, c \in B$, where $A_n = \prod_{j=n}^{n+m-1}s_j$, $W_{n,m}a_n = a_{n+m}$ and $a^* = a^*$ for all $a^* \in F(T_i, S_i)$.

Hence,

$$\gamma_{n+m}(t) = \|ta_{n+m} + (1-t)a_1^* - a_2^*\|$$
$$= \left\| \begin{matrix} tW_{n,m}a_n + ((1-t)a_1^* - a_2^* + W_{n,m}(ta_n + (1-t)a_1^*) - a_2^* \\ +a^* - a^* \\ -W_{n,m}(ta_n + (1-t)a_1^*) - a_2^* \end{matrix} \right\|$$
$$\leq b_{n,m} + \|W_{n,m}(ta_n + (1-t)a_1^*) - a_2^*\|$$
$$\leq b_{n,m} + A_n\gamma_n(t) + \sum_{j=n}^{n+m-1}v_j$$

By using Lemma (1.7), we have

$$b_{n,m} \leq Kf^{-1}(\|a_n - a^*\| - \frac{1}{K}\|W_{n,m}a_n - W_{n,m}a^*\|)$$
$$\leq Kf^{-1}(\|a_n - a^*\| - \frac{1}{K}(\|a_{n+m} - a^*\| - \|W_{n,m}a_n - a^*\|)$$

and $(b_{n,m})$ converges uniformly to zero. Since $\lim_{n\to\infty}A_n = 1$ and $\lim_{n\to\infty}v_n = 0$, we get

$$\lim_{n\to\infty}\lim_{m\to\infty}sup\gamma_{n+m} \leq \lim_{n\to\infty}\lim_{m\to\infty}b_{n,m} + \lim_{n\to\infty}inf\gamma_n(t) = \lim_{n\to\infty}inf\gamma_n(t)$$

Thus, $\lim_{n\to\infty}\gamma_n(t)$ exists for all $t \in [0,1]$.

Theorem (2.4): Let B be a nonempty closed convex subset of a uniformly convex Banach space $M, T_j, S_j, \forall j = 1, 2, \dots, k$ be two families of Lipschiztain and total asymptotically quasi-nonexpansive self-maps of B and the sequence (a_n) be as shown in step (1). If there is $Z > 0$ such that $\psi((\lambda_j) \le Z\lambda_j, j = 1, 2, \dots, k$. Then, $\lim_{n\to\infty}\|T_j^n a_n - a_n\| = 0 = \lim_{n\to\infty}\|S_j^n a_n - a_n\|, \forall j$.

Proof: By Lemma (2.2), $\lim_{n\to\infty}\|a_n - a^*\|$ exists. Assume that

$$\lim_{n\to\infty}\|a_n - a^*\| = e, \qquad \forall e \ge 0.$$

If e=0, the proof is straight forward.

Now, suppose $e > 0$. We get

$$a_{n+1} = (1 - \alpha_{jn})S_j^n a_n + \alpha_{jn}T_j^n b_{jn}$$

and

$$e = \|a_{n+1} - a^*\| = \|(1 - \alpha_{jn})(S_j^n a_n - a^*) + \alpha_{jn}(T_j^n b_{jn} - a^*)\|.$$

Since $\|S_j^n a_n - a^*\| \le K\|a_n - a^*\| \le e$

and $\|T_j^n b_{jn} - a^*\| \le K\|b_{jn} - a^*\| \le K\{(1 + u_n)^j\|a_n - a^*\| + v_n^j\} \le e.$

Therefore, by Lemma (1.5), we get

$$\lim_{n\to\infty}\|S_j^n a_n - T_j^n b_{jn}\| = 0$$

Next,

$$\begin{aligned}\|b_{jn} - a_n\| &\le (1 - \alpha_{jn})\|S_j^n a_n - a_n\| + \alpha_{jn}\|T_j^n b_{(j-1)n} - a_n\| \\ &\le (1 - \alpha_{jn})(1 + f_{jn}Z)\|a_n - a_n\|(1 - \alpha_{jn})g_{jn} \\ &\quad + \alpha_{jn}(1 + f_{jn}Z)\|b_{(j-1)n} - a_n\| + \alpha_{jn}g_{jn} \\ &\le g_{jn} + \alpha_{jn}(1 - \alpha_{(j-1)n})(1 + f_{jn}Z) \\ &\quad \|S_{j-1}^n a_n - a_n\| + \alpha_{jn}\alpha_{(j-1)n}(1 + f_{jn}Z) \\ &\quad \|T_{j-1}^n b_{(j-2)n} - a_n\| \\ &\le g_{jn} + \alpha_{jn}(1 - \alpha_{(j-1)n})(1 + f_{jn}Z)g_{jn} \\ &\quad + \alpha_{jn}\alpha_{(j-1)n}(1 + f_{jn}Z)(1 \\ &\quad + f_{(j-1)n}Z)\|b_{(j-2)n} - a_n\| \\ &\quad + \alpha_{jn}\alpha_{(j-1)n}(1 + f_{jn}Z)g_{(j-1)n} \\ &\quad . \\ &\quad . \\ &< \infty\end{aligned}$$

Since $Z > 0, \sum_{n=1}^{\infty} f_{jn} < \infty$ and $\sum_{n=1}^{\infty} g_{jn} < \infty$, hence $\lim_{n\to\infty}\|b_{jn} - a_n\| = 0$.

Then,

$$\begin{aligned}\|S_j^n a_n - a_n\| &\le \|S_j^n a_n - T_j^n b_{jn}\| + \|T_j^n b_{jn} - a_n\| \\ &\le \|S_j^n a_n - T_j^n b_{jn}\| + K\|b_{jn} - a_n\| \to 0 \\ &\quad as\ n \to \infty.\end{aligned}$$

And

$$\begin{aligned}\|T_j^n a_n - a_n\| &\le \|T_j^n a_n - S_j^n a_n\| + \|S_j^n a_n - a_n\| \to 0 \\ &\quad as\ n \to \infty.\end{aligned}$$

Theorem (2.5): Let B be a nonempty closed convex subset of a Banach space and $T_j, S_j, j = 1, 2, \dots, k$ be two families of total asymptotically quasi-nonexpansive self-maps of B. Presume that $F(T_j, S_j) \ne \emptyset$ and (a_n)be as shown in step (1)is strong convergence to a common fixed point of T_j and S_j iff $\lim_{n\to\infty} inf\ d(a_n, F) = 0$, where $d(a, F) = inf_{a^* \in F}\|a - a^*\|$.

Proof: To show $\lim_{n\to\infty} inf\ d(a_n, F) = 0$ implies that (a_n) is strong convergence to a common fixed point of $T_j, S_j, j = 1, 2, \dots, k$, since by (2)

$$\begin{aligned}d(a_{n+1}, F) &\le (1 + u_n)^{j+1}d(a_n, F) + v_n^{j+1} \\ &\le (1 + u_n)d(a_n, F) + v_n\end{aligned}$$

By Lemma (1.6), we get $\lim_{n\to\infty} a_n$ exists and $\lim_{n\to\infty} inf a_n = 0$.

Hence, $\lim_{n\to\infty} a_n = 0$.

Next, to prove the sufficiency, firstly show that (a_n) Cauchy sequence. By using Lemma (2.1.ii), we get

$$\|a_{n+p} - a^*\| \le J_1\|a_n - a^*\| + J_2 \qquad (3)$$

$\forall a^* \in F(T_j, S_j), n = p = 1,2, \dots$

Since $\lim_{n\to\infty} a_n = 0, \forall \epsilon > 0, \exists N$ such that

$$d(a_n, F) \le \frac{\epsilon}{3J_1} - \frac{J_2}{J_1}, \qquad \forall n \ge N$$

therefore, there is $l \in F(T_i, S_i)$ such that

$$\|a_N - l\| \le \frac{\epsilon}{2J_1} - \frac{J_2}{J_1} \qquad (4)$$

From (3) and (4), $\forall n \ge N$, we get

$$\begin{aligned}\|a_{n+p} - a_n\| &\le \|a_{n+p} - l\| + \|a_n - l\| \\ &\le J_1\|a_N - l\| + J_2 + J_1\|a_N - l\| + J_2 \\ &\le J_1\frac{\epsilon}{2J_1} - J_2 + J_2 + J_1\frac{\epsilon}{2J_1} - J_2 + J_2 \\ &= \epsilon\end{aligned}$$

Then, (a_n) is a Cauchy sequence and converges to $c \in M$.

Lastly, to show that $c \in F(T_j, S_j)$, for any $\epsilon^* > 0$, there is N_1such that

$$\|a_n - c\| \le \frac{\epsilon^*}{2(2 + f_jZ)} - \frac{3g_j}{(2 + f_jZ)}, \forall n \ge N_1 \qquad (5)$$

Since $\lim_{n\to\infty} a_n = 0$ implies that $N_2 \ge N_1$ such that $d(a_n, F) \le \frac{\epsilon^*}{3(4+3f_jZ)}, \forall n \ge N_2$.

Thus, $\exists\ l_1 \in F(T_j, S_j)$ such that

$$\|a_N - l_1\| \le \frac{\epsilon^*}{2(4+3f_jZ)} \qquad (6)$$

From (5) and (6) for any $T_j, \forall j = 1, 2, \dots, k$, we obtain

$$\begin{aligned}\|T_j c - c\| &\le \|jc - l\| + 2\|T_j a_{N_2} - l_1\| + \|a_{N_2} - l\| \\ &\quad + \|a_{N_2} - c\| \\ &\le \|c - l_1\| + f_j\psi\|p - l_1\| + g_j + 2\|a_{N_2} - l_1\| \\ &\quad + 2f_j\psi\|a_{N_2} - l\| + 2g_j + \|a_{N_2} - l_1\| \\ &\quad + \|a_{N_2} - c\| \\ &\le (1 + f_jZ)\|c - l_1\| + 2(1 + f_jZ)\|a_{N_2} - l_1\| + 3g_j \\ &\quad + \|a_{N_2} - l_1\| + \|a_{N_2} - c\| \\ &\le (1 + f_jZ)\|a_{N_2} - c\| + (1 + f_jZ)\|a_{N_2} - l_1\| \\ &\quad + 2(1 + f_jZ)\|a_{N_2} - l_1\| + 3g_j + \|a_{N_2} - l_1\| \\ &\quad + \|a_{N_2} - c\| \\ &\le (2 + f_jZ)\|a_{N_2} - c\| + (4 + 3f_jZ)\|a_{N_2} - l_1\| \\ &\quad + 3g_j \\ &\le (2 + f_jZ)\frac{\epsilon^*}{2(2 + f_jZ)} - 3g_j + 3g_j \\ &\quad + (4 + 3f_jZ)\frac{\epsilon^*}{2(4 + 3f_jZ)} \\ &= \epsilon^*\end{aligned}$$

Therefore, $\|T_j c - c\| = 0\ \forall i$ which means $T_j c = c, \forall j = 1, 2, .., k$.

By using the same above argument we can prove $\|S_j c - c\| = 0, \forall j$.Thus $c \in F(T_j, S_j)$.

Theorem (2.6): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, j = 1, 2, \dots, k$ be two families of Lipschitzain and total asymptotically quasi-nonexpansive self-maps of B. If M accomplishes Opial 's condition and the maps $I - T_j$ and $I - S_j, j = 1,2, \dots, k$ are demiclosed to zero, therefore (a_n) be as shown in step (1) is weak convergence to a common fixed point of T_j and $S_j, j = 1, 2, .., k$.

Proof: Let $a^* \in F(T_j, S_j)$. By Lemma (2.2), $\lim_{n\to\infty} \|a_n - a^*\|$ exists.

By Theorem (2.4), we have

$$\lim_{n\to\infty} \left\|T_j^n a_n - a_n\right\| = 0 = \lim_{n\to\infty} \left\|S_j^n a_n - a_n\right\|$$

for j=1, 2,…, k. Since by the supposition the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demiclosed to zero, therefore $T_j a^* = a^*$ and $S_j a^* = a^*$, that means $a^* \in F(T_j, S_j)$.

Next, to prove (a_n) converges weakly to a^*. Assume there is other subsequence $(a_{ni})\ of\ (a_n)$ is weak convergence to $b^* \in F(T_j, S_j)$ and $a^* \neq b^*$. By using the same argument as above we can show that $b^* \in F(T_j, S_j)$.

Now, to prove the uniqueness, assume $a^* \neq b^*$. Therefore, by using Opial 's condition, we obtain:

$$\begin{aligned}\lim_{n\to\infty} \|a_n - a^*\| &= \lim_{nj\to\infty} \left\|a_{nj} - a^*\right\| \\ &< \lim_{nj\to\infty} \left\|a_{nj} - b^*\right\| \\ &= \lim_{ni\to\infty} \|a_{ni} - b^*\| \\ &< \lim_{ni\to\infty} \|a_{ni} - a^*\| \\ &= \lim_{n\to\infty} \|a_n - a^*\|\end{aligned}$$

This is contradiction, therefore $a^* \neq b^*$. Hence, (a_n) is weak convergence to a^*.

Theorem (2.7): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, j = 1, 2, \dots, k$ be two families of Lipschitzain and total asymptotically quasi-nonexpansive self-maps of B. If the dual space $M^*\ of\ M$ has the Kadec-klee property and the maps $I - T_j, I - S_j, j = 1,2, \dots, k$ are demi-closed to zero, therefore, (a_n) be as shown in step (1) is weak convergence to a common fixed point of T_j and S_j.

Proof: As showed by Lemma (2.2), that $\lim_{n\to\infty} \|a_n - a^*\|$ exists. Since (a_n) is bounded in B and M is reflexive. Therefore, there exists a subsequence $(a_{ni})\ of\ (a_n)$ which is weak convergence to a point$a^* \in B$. By Theorem (2.4)

$$\lim_{n\to\infty} \left\|T_j^n a_n - a_n\right\| = 0 = \lim_{n\to\infty} \left\|S_j^n a_n - a_n\right\|$$

∀j=1, 2, …, k .

Since by the supposition the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demi-closed to zero. Thus, $a^* \in F(T_j, S_j)$. Now, to prove (a_n) is weak convergence to a point a^*. Presume that (a_{nk}) is other subsequence of (a_n) that is weak convergence to a point $b^* \in F(T_j, S_j)$. By using the same argument as above, we obtain $b^* \in F(T_j, S_j)$.

Therefore, by Lemma (2.3) $\lim_{n\to\infty} \|ta_n + (1-t)a^* - b^*\|$ exists for all $t \in [0,1]$.

By Lemma (1.8) $a^* = b^*$. As a result, the sequence (a_n) is weak convergence to the point $a^* \in F(T_j, S_j)$.

The following corollaries are special cases

Corollary (2.8): Let B be a nonempty closed convex subset of a Banach space and $T_j, S_j, j = 1, 2, \dots, k$ be two families of total asymptotically nonexpansive self-maps of B. Presume that $F(T_j, S_j) \neq \emptyset$ and $\sum_{n=1}^{\infty} f_n < \infty, \sum_{n=1}^{\infty} g_n < \infty$.

Presume that (a_n) be as shown in step (1) is strong convergence to a a common fixed point of T_j and S_j iff $\lim_{n\to\infty} inf\ d(a_n, F) = 0$, where $d(a, F) = inf_{a^* \in F} \|a - a^*\|$.

Corollary (2.9): Let T_j, S_j, B, f_j and $g_j,\ j = 1, 2, \dots, k$ be as in corollary (2.8). Therefore (a_n) be as shown in step (1) is strong convergence to $a^* \in F(T_j, S_j)$ iff $(a_{ni})\ of\ (a_n)$ that converges to a^*.

Corollary (2.10): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, \forall j = 1, 2, \dots, k$ be two families of Lipschitzain and total asymptotically nonexpansive self-maps of B. If the dual space $M^*\ of\ M$ has the Kadec-klee property and the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demi-closed to zero, therefore (a_n) be as shown in step (1) is weak convergence to a common fixed point of $T_i\ and\ S_i$.

Corollary (2.11): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, \forall j = 1, 2, \dots, k$ be two families of total asymptotically nonexpansive self-maps of B. If M accomplishes Opial 's condition and the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demi-closed to zero, therefore, (a_n) be as shown in step (1) is weak convergence to a common fixed point of $T_j, S_j, \forall j = 1, 2, .., k$.

Corollary (2.12): Let B be a nonempty closed convex subset of a Banach space and $T_j, S_j, j = 1, 2, \dots, k$ be two families of asymptotically quasi-nonexpansive self-maps of B. Presume that $F(T_j, S_j) \neq \emptyset$ and $\sum_{n=1}^{\infty} f_n < \infty$. Presume that (a_n) be as shown in step (1) is strong convergence to a common fixed point of T_j and S_j iff $\lim_{n\to\infty} inf\ d(a_n, F) = 0$, where $d(a, F) = inf_{a^* \in F} \|a - a^*\|$.

Corollary (2.13): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, \forall j = 1, 2, \dots, k$ be two families of Lipschitzain and asymptotically quasi-nonexpansive self-maps of B. If the dual space $M^*\ of\ M$ has the Kadec-klee property and the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demi-closed to zero, therefore, (a_n) be as shown in step (1) is weak convergence to a common fixed point of T_j and S_j.

Corollary (2.14): Let B be a nonempty closed convex subset of a uniformly convex Banach space and $T_j, S_j, \forall j = 1, 2, \dots, k$ be two families of Lipschitzain and asymptotically quasi-nonexpansive selfmaps of B. If M accomplishes Opial 's condition and the maps $I - T_j$ and $I - S_j, \forall j = 1,2, \dots, k$ are demi-closed to zero, therefore, (a_n) be as shown in step (1) is weak convergenve to a common fixed point of T_j and $S_j, \forall j = 1, 2, .., k$.

3. Numerical Example

We illustrate our results by the following

Example (3.1): Let $T_j, S_j: R \to R, \forall j = 1, 2, \ldots, k$ be two maps such that $T_j a = \frac{2a}{3^j}$ and $S_j a = \frac{46a}{2^j}, \forall\, a \in R$. Choose $\alpha_{jn} = \frac{n}{4(n+1)}\ \forall\, n$ with initial value $a_1 = 15$. Let (a_n) be the sequence. According to Table 1 and Figure 1, it eases to sight that (a_n) converges to the fixed point $a^* = 0$.

Table 1: Numerical results corresponding to $a_1 = 15$ for 36 steps.

n	Iteration (1)	n	Iteration (1)
1	15.0000	**13**	0.0184
2	9.4401	**14**	0.0102
3	5.6594	**15**	0.0056
4	3.3084	**16**	0.0031
5	1.9044	**17**	0.0017
6	1.0849	**18**	0.0009
7	0.6134	**19**	0.0005
8	0.3448	**20**	0.0003
9	0.1930	**21**	0.0002
10	0.1076	**22**	0.0001
11	0.0599	**23**	0.0000
12	0.0332	**24**	0.0000

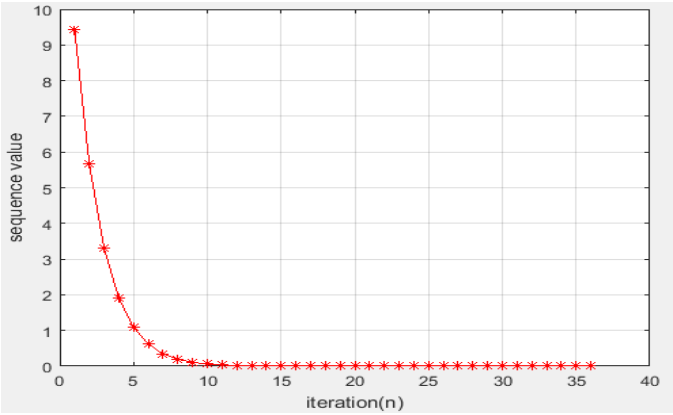


Figure. 1. Convergence behavior corresponding to $a_1 = 15$ for 36 steps.

4. Conclusion

We study the strongly and weakly convergence of new type of finite-step iteration processes under total asymptotically quasi-nonexpansive maps, see Theorems (2.4)-(2.6). Our results are generalizing and unifying the results of others who have been referred to in the references.

5. Open Problem

Recently, S.S. Abed has been defined as the following type of generalizations of total asymptotically quasi-nonexpansive[16]: Let $\mathcal{A}$ be a subset of real Banach space a set–valued map $G: \mathcal{A} \to 2^{\mathcal{A}}$ is called the general asymptotic set-valued if for each $x \in \mathcal{A}$ there exists null non- negative real sequences $\{a_n\}$ and $\{b_n\}$ such that

$$\mathfrak{D}(G^n x, G^n y) \leq \|x - w\| + a_n\, \mu(\|x - w\|) + b_n \qquad (7)$$

for any $y \in \mathcal{A}$, $w \in G^n y$ and $\mu: R^+ \to R^+$ with $\mu(0) = 0$. One can study convergence theorems in (1) and in [17, theorem (11)] for families of general asymptotic set-valued maps. As well as possible to demonstrate new results in the case of other spaces as a modular space [18].

Acknowledgment

The authors wish to express his thanks to the referees for their helpful advice.

References

[1] S.S.Abed, Z.M Mohamed Hasan "Convergence theorems of a finite-step iteration algorithm under two finite families of total asymptotically quasi-nonexpansive maps" IEEE Xplore Digital Library Internationl Conference on Advanced Science and Engineering (ICOASE), , 2018.

[2] U K.Goebel, W. A. Kirk, "Fixed point theorem for asymptotically nonexpansive mappings" Proc. Amer. Math. Soc. 35, 171-174, 1972.

[3] W. V. Petryshyn, T. E. Williamson, "Strong and weak convergence of the sequence of successive approximations for quasi-nonexpansive mappings", J. Math. Anal. Appl. 43, 459-497, 1973.

[4] YA. I. Alber, C. E. Chidume, H. Zegeye, " Approximating fixed points of total asymptotically nonexpansive mappings" Fixed point theory Appl. 2006, article ID 10673.

[5] G. S. Saluja, "Convergence to common fixed points for generalized asymptotically quasi-nonexpansive mappings" Bull. Soc. Math. 4, 69-79, 2014.

[6] G. S. Saluja, "Strong convergence theorems for two finite families of generalized asymptotically quasi-nonexpansive mappings with application" Math. Moravica, 22(1), 1-14, 2018.

[7] S.S.Abed, Z.M Mohamed Hasan,."Common fixed point of a finite-step iteration algorithm under total asymptotically quasi-nonexpansive maps" is accepted in Baghdad Science Journal.

[8] J. B. Diaz, F. B. Metcalf, "On the structure of the set of subsequential limit points of successive approximation" Bull. Amer. Math. Soc., 73, 516–519, 1967.

[9] D. R, Sahu, D. ORegan, R. P. Agarwal ,"Fixed point theorey for Lipschitzain-type mappings with applications, Topological fixed point theory and its applications" doi:10.1007/978-387-75818-3-1.

[10] Q. H. Liu, "Iterative sequences for asymptotically quasi-nonexpansive mappings" J. Math. Anal. Appl. 259(1), 1–7, 2001.

[11] A. Abkar, M. Shekarbaigi,"A novel iteration algorithm applied to totally asymptotically nonexpansive mappings in CAT(0) spaces" Mathematices 5,14, 2017, doi:10.3390/math5010014.

[12] F.E.Browder," Semicontractive and semiaccretive nonlinear mappings in Banach spaces" Bull. Amer. Math. Soc. 74, 660-665, 1968.

[13] A.Sharma, M.Imdad, "Approximating fixed points of generalized nonexpansive mappings Via faster iteration schemes" Fixed point theory, 4(4), 605-623, 2014.

[14] C. E. Chidume, E.U. Ofoedu, "Approximation of common fixed points for finite families of total asymptotically nonexpansive mappings" J. Math. Anal. Appl. 333, 128-141, 2007,doi:10.1016/j.jmaa.2006.09.023.

[15] J.G. Falset, W. Kaczor, T. Kuczumow, S. Reich, " Weak convergence theorem for asymptotically nonexpansive mappings and semigroups", Nonlinear Anal. 43, 377- 401, 2001.

[16] S.S.Abed, "Approximating fixed points of the general set –valued asymptotically mappings " to appear.

[17] S.S.Abed, Z.M Mohamed Hasan "Convergence comparison of two schemes for common fixed points with an application" accepted in Ibn Al Haitham journal for pure and applied sciences.

[18] S.S.Abed, K. E. Abdul Sada, " Common fixed points in modular spaces", IHSCICONF 2017, Ibn Al Haitham journal for pure and applied sciences.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 75-87 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Development of Smart Technology for Complex Objects Prediction and Control on the Basis of a Distributed Control System and an Artificial Immune Systems Approach

Samigulina Galina[1], Samigulina Zarina[2,*]

[1]*Intellectual control systems and forecasting, Institute of information and computing technologies, 050010, Kazakhstan*

[2] *Faculty of Information Technologies, Kazakh British Technical University, 050000, Kazakhstan*

ARTICLE INFO

Article history:
Received: 07 February 2019
Accepted: 03 May, 2019
Online: 21 May, 2019

Keywords :
Intellectual control systems
Complex object
Artificial Immune System
Distributed control system
Multi-agent Smart system

ABSTRACT

This paper is an extension of work originally presented in 2018 Global Smart Industry Conference (GloSIC). Researches are devoted to the development of Smart technology for complex objects control and prediction on the basis of a distributed Honeywell DCS control system of the TengizChevroil enterprise using the example of a technological process of medium pressure gas cleaning. The article describes how on the basis of the multi-algorithm approach there was developed a modified algorithm based on modern artificial intelligence methods in order to select informative features (principal component method, Random Forest algorithm, particle swarm algorithm) and artificial immune systems (clonal selection) solving the image recognition problem and predicting the state of a complex control object. There was conducted a comparative analysis of the simulation results using the example of real production data (daily data of sensors from an average pressure absorber).

1. Introduction

Nowadays, the efficiency of industrial enterprises depends on many factors: material and technical support (installation of modern production equipment, modernization and reconstruction of the existing technical base), economic factors (growth of production and raw materials costs), management factors (information support for decision-making), and etc [1]. Therefore, modern researches in the field of development of automated complex objects control systems are aimed at increasing the reliability and at reducing the influence of various negative factors on the production process.

The implementation of modern Smart technology enables the possibility to improve production by effective prediction of the technical state of the equipment and the timely resolution of problems at the stage of their occurrence in real time. A characteristic feature of the complex objects control is the need to process a huge array of production data, as well as prediction and decision-making under the condition of parameters uncertainty.

In order to solve such problems at the complex objects control process, there were proven modern methods of artificial intelligence with the following advantages: the ability to process multidimensional data, self-organization and relative ease of implementation.

Widespread use has control systems based on neural networks (NN). For example, in [2] there is presented the use of a decentralized robot control system based on a neural network. There were proposed four decentralized control systems which can identify the dynamics of the robot's movement. Training of each neural network is performed online using the advanced Kalman filter. The researches [3] are devoted to the development of a combined adaptive neural network and to the nonlinear prediction model for production processes control.

The effectiveness of the proposed method was tested on a system of reactors with continuous mixing. In work [4] there is presented a PID controller based on a neural network for the pH neutralization process control system. The classical PID control cannot adapt to changes in the process, and therefore a PID controller with a signal neuron was developed. The simulation results showed the reliability and adaptability of the proposed control system.

In addition to neural networks there are used swarm intelligence (SI) algorithms for complex objects control. For example, in work [5] there is presented the application of the

*Corresponding Author Zarina Samigulina, e-mail: zarinasamigulina@mail.ru
www.astesj.com
https://dx.doi.org/10.25046/aj040312

particle swarm algorithm (PSO) for the optimal choice of process control parameters. In researches [6] there was used an ant colony optimization at setting the PID controller for a conical tank. The proposed approach has shown better results compared to setting of a closed loop based on the Ziegler-Nichols method and to setting of a controller in a closed loop system. In work [7] there is presented the development of a control system for autonomous mobile robots based on a bee colony algorithm and a genetic algorithm. Researches [8] are devoted to applying the method of bacteria movement optimization at the development of a PID controller by a two-wheeled robot with a two-way movement mechanism for a working environment with limited space.

Also there is widely used an approach based on fuzzy logic for complex objects control. Another study a control system for mobile robots with nonholonomic wheels moving on a plane on the basis of a fuzzy controller in order to achieve a given trajectory of motion. The problem of rounding obstacles is considered. Parameters of PID controller are calculated on the basis of fuzzy logic. There are presented simulation results based on a real vehicle. The researches are devoted to cascade control based on a fuzzy PID controller for a conveyor for weighing materials. The article presents a non-linear model of a conveyor system of continuous weighing, taking into account the delay time of material transportation. There has been developed a PID controller based on fuzzy logic with two control loops: an internal speed control loop and an external mass flow control loop. The results of simulation and experiment show that the proposed algorithm for fuzzy PID control can improve the quality of material mass flow control in an industrial conveyor system of continuous weighing using a programmable logic controller.

In work [9] there is proposed an innovative approach for fuzzy control systems setting with low parametric sensitivity based on the Gray Wolfs Optimization Algorithm (GWO). The servo system is controlled on the basis of the Takagi-Sugeno-Kang fuzzy PI controller (TSK PI-FC). The GWO algorithm is used in solving the optimization problem. Article [10] is devoted to an optimized fuzzy-genetic algorithm (F-GA) for controlling the process of production of aluminum foam with open pores. The application of the F-GA method allows to take into account the uncertainty associated with the model and the statistical variability of the production process within one model.

Nowadays, the researches in the field of the use of the bio-inspired approach of artificial immune systems (AIS) for complex objects control are also promising. In work [11] there is presented AIS application at PID controller parameters setting for nonlinear control systems. Researches [12] are devoted to the development of a complex object control algorithm based on an adaptive artificial immune system. The simulation was carried out on the example of a reactor with a mixing function. The article [13] provides an AIS application for control a group of robots. The work [14] is devoted to the use of a meta-heuristic AIS for signal evaluation in identifying the contour of the angle control of an unmanned aerial vehicle. In the article [15] there is used the AIS algorithm for preliminary signal processing at detection of failures for autonomous underwater vehicles. Article [16] is devoted to applying the AIS approach (negative selection algorithm) in order to control the combustion process in the chamber with the implementation on the microcontroller. In researches [17] there was developed an immune reconfigurable controller based on four modules for the immune system modeling: the observation process, response, mechanism of memory and self-training. The proposed algorithm is able to detect drive failures and to provide failure-tolerant control for the systems with multiple inputs and outputs.

In addition to bioinspired methods, machine learning algorithms for complex objects control and solving the problem of identifying informative features have become widespread. The most promising is the Random Forest (RF) algorithm. For example, in [18], RF is used to detect defective plates by chemical characteristics in the production of semiconductors. The article [19] presents the use of algorithms for machine learning of RF, decision trees, etc. in heavy industry for analysis of large amounts of production data. The work [20] is devoted to the use of the Random Forest algorithm in sensor failures diagnostics. The algorithm is used to reduce memory requirements in the process of assessing the state of the equipment in real time.

Therefore, the analysis of the literature proves the relevance of the development of modern Smart technologies and prediction of complex objects control based on industrial enterprises using artificial intelligence methods.

The following structure of the article is proposed. The second section contains the formulation of the research problem. The third section presents the development of Smart technology for complex objects control and prediction based on the Honeywell Experion Process Knowledge System (PKS) distributed control system (DCS) and the approach of artificial immune systems; reviewed the operation of the DCS at the Tengizchevroil enterprise, and presented the general structure of the Smart technology. The fourth section is devoted to the development of a modified algorithm of artificial immune systems. The fifth section describes the multi-agent Smart system for complex object prediction and control under industrial operation conditions. The sixth section contains the simulation results. Description of a real complex medium-pressure absorber, its principle of operation, readings from sensors, etc. The simulation results are presented in the form of graphs, as well as a comparative analysis of the effectiveness of the proposed algorithms. The conclusion describes the main findings and a list of references.

2. Problem statement

Since the complex objects control with the presence of parameters uncertainty in the conditions of industrial operation is a time-consuming task due to the need to process multidimensional data in real time, the use of promising heuristic algorithms of artificial intelligence is important.

The research problem statement is formulated as follows: it is necessary to develop a modern Smart technology for complex industrial objects control based on a multi-algorithm approach. Within the framework of the multi-algorithm approach to synthesize a modified algorithm of artificial immune systems in order to solve the problem of prediction and control. In order to organize an autonomous, flexible and decentralized control system based on the proposed technology, it is necessary to develop a multi-agent Smart system for implementation in real production based on TengizChevroil enterprise, with the aim of working with Honeywell PKS distributed control systems.

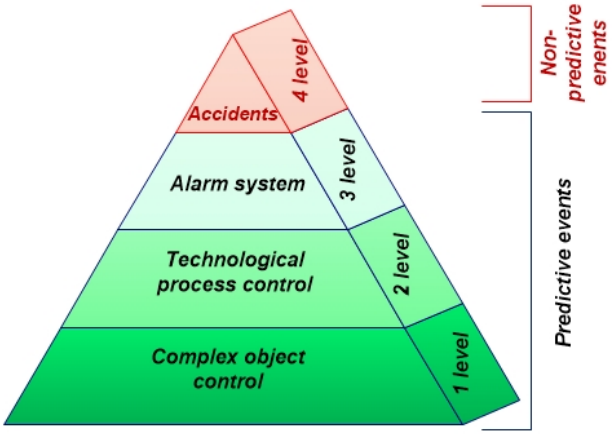


Figure 1: Honeywell DCS process control system with independent protection levels

3. Development of Smart technology for complex objects prediction and control on the basis of a distributed control system and an artificial immune systems approach

3.1. Distributed control system at TengizChevroil enterprise

TengizChevroil is the largest company engaged in the development and production of oil in the territory of the Republic of Kazakhstan together with the American company Chevron. In TengizChevroil there is used modern industrial equipment from Honeywell, in particular, Honeywell PKS distributed control systems (DCS). Nowadays, the Experion PKS system is the best in the field of distributed control systems and received an award from the magazine "Control Engineering". The company created a consortium for emergency situations control, aimed at exploring the causes of emergency situations and developing solutions aimed at reducing them. Therefore, the researches on improving the quality of equipment at the enterprise are relevant.

The Honeywell PKS distributed control system contains several independent levels of protection (Figure 1), designed to work with predictable and unpredictable events [23].

Predicted events include the following levels of independent protection:

1 level. Maintains "normal" operating conditions of the process. The distributed control system is engaged in the collection and processing of technological parameters, as well as their adjustment to ensure a given mode.

2 level. This level requires decision-making from the operator if the alarm is triggered and the first level is unable to maintain normal operating conditions of the control object.

3 level. This level indicates that the control object has left the "normal conditions" of operation mode and a controlled stop is required. The Safety Instrumented Systems (SIS) is activated. The SIS security system also includes: Emergency Shutdown System (ESD), Fire & Gas System (F&G), Safety Shutdown System (SSD) и т.д.

4 level. The level refers to unpredictable events, such as the occurrence of fires, gas leaks and other major accidents. A distributed system cannot predict the occurrence of such accidents, but it is designed for quick response and minimization of damage.

Information on the operation of the technological process is stored in the historical module (History Module, HM) of Honeywell DCS [23]. Developed Smart technology is aimed at working with the predicted events of the Honeywell DCS, with obtaining data for processing using the HM.

3.2. Smart technology based on Honeywell PKS distributed control system

Let consider the general structure of the proposed Smart-technology for complex objects prediction and control. The Honeywell PKS distributed control system consists of nodes described by controllers, operator consoles, control devices, data storage equipment, etc. The information transfer between nodes occurs through a local and universal control network. Information is collected in a historical module that is connected to a local control network, and a process history database is formed here.

The developed Smart-technology is designed to work with the predicted events of Honeywell PKS (Figure 1). The process history database containing current information on the behavior of a complex control object (readings from sensors, monitoring systems, etc.) is processed on the basis of a multi-algorithm approach (Figure 2).

Then a database of optimal parameters is formed for prediction and control on the basis of AIS algorithm, clonal selection. Depending on the nature of the processed data in the modified algorithm based on the AIS method, the most effective algorithm is selected that shows the best predictive result.

Further in the paper, as an example, we will consider the complex control object of the Tengizchevroil enterprise: a fragment of the Installation 300 for cleaning gases from acidic components [21].

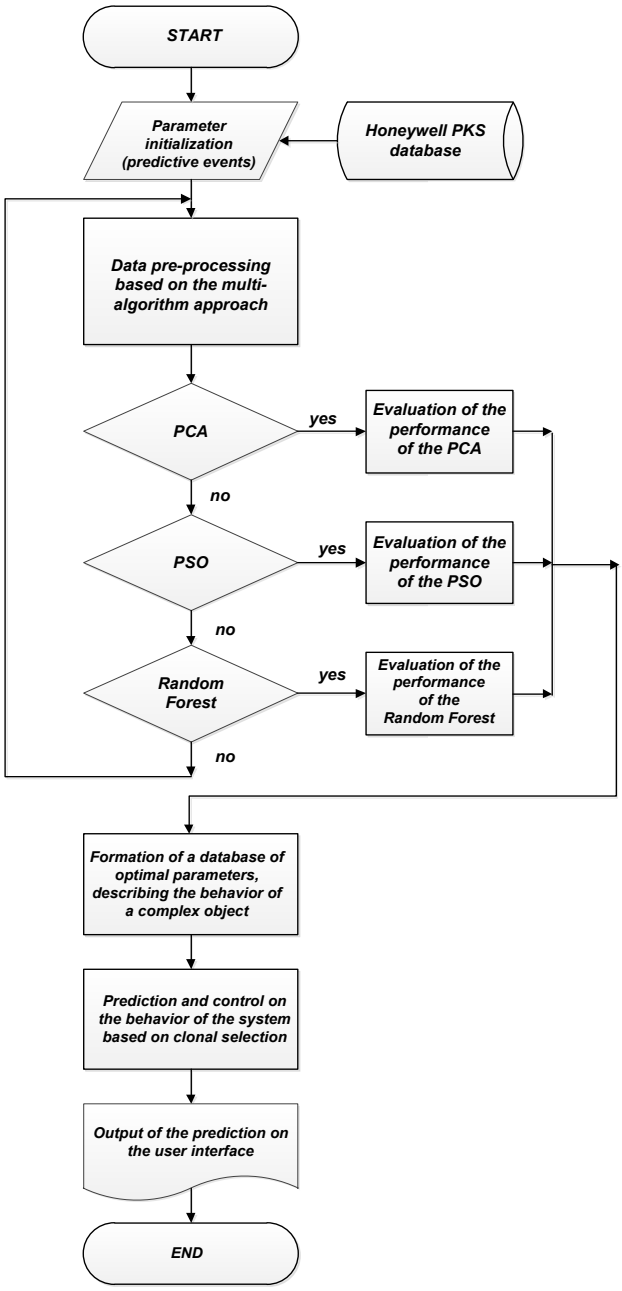


Figure 2: The structure of the developed Smart technology for complex objects prediction and control based on Honeywell PKS

4. Modified artificial immune algorithm based on clonal selection

Nowadays, the approach of artificial immune systems is a promising area of research. This method has the following advantages: adaptability, memory, dynamic arrangement of elements, ability to self-organization, etc. There are several main areas of AIS research: negative selection algorithms (NSA), immune system algorithms (INA), danger theory algorithms (DTA), and clonal selection algorithms (CSA) [22].

Let consider in more detail the clonal selection algorithm [23-25]. When a foreign antigen enters the body, special cells B-lymphocytes produce antibodies A_b. Each cell produces only one type of antibody that is relatively specific for the antigen. At binding antibodies (receptors), as well as using T-helper cells, the antigen stimulates B cells to multiply and to mature into plasma cells. Different cell divisions produce clones, by which there is meant a set of cells that are descendants from a single cell. Plasma cells most actively produce antibodies B-lymphocytes that divide rapidly. Antibodies A_b divide much more slowly. Lymphocytes can also differentiate in addition to plasma cells into long-lived memory cells - B - memory cell. Memory cells circulate through the body through the blood, lymph and tissue. In a secondary immune response, they can turn into large lymphocytes capable of producing antibodies with high affinity. These cells are pre-selected for the specific antigen that stimulated the primary response. The clonal selection algorithm is based on the following operations: selection, cell cloning and mutation. The selection procedure is carried out for cells with higher affinity and subsequent cloning, while other cells undergo a mutation process, which helps to improve their quality.

The clonal selection algorithm is presented as follows [23]:

Algorithm 1

Step 1. Generation of a set of possible solutions P, composed of a subset of memory cells M, added to the remaining population P_r.

$$P = P_r \cdot M \tag{1}$$

Step 2. Determination of the best individuals n, the population P_n based on the value of affinity (affinity).

Step 3. Cloning of n best individuals of a population, with the generation of a temporary population of clones C.

Step 4. Representation of a clone population to a hypermutation procedure, where hypermutation is proportional to the affinity of the antibody and antigen. Creation of a population of mature antibodies C^*.

Step 5. Repeated selection of improved individuals from the population C^* in order to form a set of memory M. Some members of the population P may be replaced by other improved population individuals C^*.

Step 6. Replacement of antibodies d by the new ones. Cells with lower affinity have a higher probability of replacement.

Modern methods of artificial intelligence are well established for the development of complex objects control systems, but unfortunately these approaches are not universal, and need to be customized for specific cases, taking into account the features of the object and the system parameters. The quality of forecasting and management directly depends on the processed data. The higher the quality of the data is, the more accurate is the forecast. Therefore, the application of the multi-algorithm approach is actual, within the framework of which the procedure of preliminary production data processing is carried out on the basis of modern statistical and intellectual algorithms. As a result of

research several methods are considered. Algorithms that will give the best predictive result are used for further research. Preliminary production data processing within the framework of the proposed Smart technology is carried out on the basis of the following approaches:

- Statistical approach - factor analysis and principal component method;

Statistical approaches are often used to process production data. In order to solve the problem of reduction of non-informative parameters during the study of complex industrial facilities, factor analysis (principal component method) has proven itself well. Among the advantages of factor analysis we can note the ability to analyze hidden (latent) relationships between parameters, as well as the relative simplicity of implementation [26]. As a disadvantage there should be noted the difficulties in the interpretation of the obtained results.

The algorithm of factor analysis consists of the following steps. The entire data set must be rotated counterclockwise, with the first axis associated with the maximum dispersion, and each subsequent one with the others. According to the results of the data rotation, the parameters that lie closer to the origin of coordinates are subject to reduction, as uninformative.

- Swarm intelligence - particle swarm optimization algorithm;

Let consider the concept of swarm intelligence (SI) [27]. The term SI was proposed by Beni and Wang in application to cellular robots and implied the collective behavior of a decentralized self-organizing system. Then it was used in the theory of artificial intelligence and describe the social behavior of living organisms (particle swarm method, ant algorithm, bee algorithm, wolf pack algorithm, bat algorithm, etc.).

In the proposed Smart technology, in the framework of the multi-algorithm approach, a particle swarm optimization algorithm is considered in detail.

A feature of the PSO method is that during each iteration, the solution is evaluated using the objective function to determine its suitability [28]. Each solution is represented as a particle in the search space. Particles move in the search space in order to find the maximum value of the objective function. As the particles move, each particle has a position in the search space, speed, and an individual best position. The PSO algorithm consists of the following steps [26, 27]:

Algorithm 2 Particle Swarm Optimization

Step1 . Initialization of a particles swarm.

Step 2. Evaluation of the suitability of each particle.

Step 3. Calculation of historically optimal positions of individuals.

Step 4. Calculation of historically optimal positions of the swarm.

Step 5. Speed and position update of particles according to the equation of speed and position. Repeat of this iteration until the stop condition is reached.

The speed of each particle is calculated by the formula:

$$\upsilon_i(t+1) = \varpi\upsilon_i(t) + c_1 \cdot r_1 \cdot (\hat{x}_i(t) - x_i(t)) + \\ + c_2 \cdot r_2 \cdot (g(t) - x_i(t)) \quad (2)$$

where i - particle index; ϖ - inertia coefficient $0.8 \le \varpi \le 1.2$; c_1, c_2 - acceleration coefficients ($0 \le c_1, c_2 \le 2$); r_1, r_2 - random values lying in the range $0 \le r_1, r_2 \le 1$; $\upsilon_1(t)$ - particle speed in time t; $x_1(t)$ – particle position in time t; $g(t)$ – the best swarm solution in time t.

The change in the position of each particle of the swarm is calculated by the following formula:

$$x_i(t+1) = x_i(t) + \upsilon_i(t+1) \quad (3)$$

The advantages of the PSO algorithm include: a simple implementation; small amount of adjustable parameters; parallel computations are possible; with a certain setting has the ability to fast convergence; no matches and mutations; has small requirements for computer memory and fast calculation speed; can be used in case when the representation of a mathematical model is impossible.

- Random Forest (RF) algorithm, etc.

Nowadays, the RF algorithm is a promising method of machine learning, capable of solving the problems of image recognition, prediction and selection of informative features [28]. A random forest is created on the basis of decision trees and has the same set of hyperparameters, while in RF the process of searching for the root node and the separation of object nodes is performed randomly [29, 30]. The more trees in the forest, the more accurate the simulation result are.

The advantages of the RF algorithm are following: there is no risk of retraining of the model; the algorithm has flexibility and high accuracy; no need to scale data; individual decision trees can be trained in parallel; data processing ability with a large amount of attributes, etc.

The operation of the Random Forest algorithm is presented below:

Algorithm 3 Random Forest

Step 1. Arbitrary selection of f signs from the total number of F signs.

Step 2. Calculation of a *node* among f features using the best separation point.

Step 3. Separate a *node* on child nodes using the best partition.

Step 4. Repeat steps 1-3 until reaching the value of the number one.

Step 5. Creation of a forest by n repeating of the steps 1-4 in order to create the n number of trees.

Based on the previously discussed methods, the modified AIS algorithm based on clonal selection can be described as follows (Figure 2).

Algorithm 4 Modified artificial immune algorithm

1. Connection to the historical module of the Honeywell PKS distributed system.
2. Data collection (readings from sensors) describing the behavior of a complex control object.
3. Formation of a database of system parameters.
4. Pre-processing of the obtained database, rationing, centering and filling in the missing data.
5. Formation of the optimal set of parameters due to the reduction of non-informative data based on the multi-algorithm approach.
5.1 Reduction of non-informative parameters based on the statistical approach, the principal component method.
5.1.1 Evaluation of the performance of the PCA algorithm.
5.1.2 Formation of a database of optimal parameters DB_PCA, describing the behavior of a complex object.
5.2 Reduction of non-informative parameters based on swarm intelligence, particle swarm algorithm.
5.2.1 Evaluation of the performance of the PSO algorithm.
5.2.2 Formation of a database of optimal parameters DB_PSO, describing the behavior of a complex object.
5.3 Reduction of non-informative parameters based on the Random Forest algorithm.
5.3.1 Evaluation of the performance of the Random Forest algorithm.
5.3.2 Formation of a database of optimal parameters DB_RF, describing the behavior of a complex object.
6. The solution of the problem of image recognition on the basis of clonal selection. Processing the received databases (DB_PCA, DB_PSO, DB_RF) and prediction.
6.1 Selection of an efficient algorithm that shows the best prognostic result.
6.2 Decision-making on the complex object control.

The proposed technology in industrial operation is most conveniently implemented on the basis of a multi-agent approach. Further there is presented a developed multi-agent Smart system for complex objects prediction and control.

5. Multi-agent object control Smart-system

Nowadays, multi-agent systems (MAS) are widely used in the field of industrial automation and artificial intelligence. A multi-agent system consists of independent agents with autonomy, decentralization, and flexibility. At developing complex industrial systems using MAS, the process of modeling and control is greatly facilitated. The paper [31] presents the MAS technology for organization of a flexible production process of customized products, through cooperation between companies, for the integration of heterogeneous production systems. Article [32] presents the use of a multi-agent artificial immune system for detecting and classifying network intrusions. Researches [33] are devoted to a decentralized multi-agent control system based on artificial immune systems. This system consists of the following target agents: the agent of self-organization, the agent of interaction, the agent of following specific goals and has the properties of flexibility, the ability of agents to work autonomously, the ability to optimally distribute the system resource, multifunctionality and self-organization.

Based on the foregoing, the development of a multi-agent Smart system for complex objects control at TengizShevroil for the integration of the developed modified artificial intelligence algorithms into production is relevant. Figure 3 shows the structure of the multi-agent Smart system.

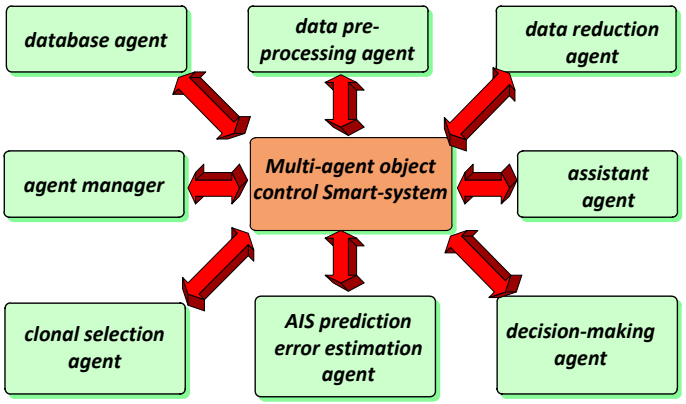


Figure 3: Structure of multi-agent smart system for complex objects control

Table 1 describes the main agents and their functions.

Table 1: Functions of agents of multi-agent smart system

Agent name	Function of the agents
Database agent	A database agent is needed to organize data readout from the Honeywell PKS historical module of the distributed control system. The generated database consists of sensors and monitoring systems, which allow to assess the state of the technological control object (normal operation, work with deviations, an accident, etc.).
Data preprocessing agent	The agent of preliminary data processing is responsible for rationing, centering, filling in the missing data in the database of parameters of the control object. The agent interacts directly with the database agent.
Data reduction agent	The data dimension reduction agent performs information processing based on the multi-algorithm approach for reducing non-informative parameters describing the behavior of a complex object.
Clonal selection agent	The agent implements the approach of artificial immune systems, the clonal selection algorithm for prediction of the behavior of a complex object.
AIS prediction error estimation agent	This agent is designed to select the algorithm with the best predictive result.
Decision-making agent	The decision support agent generates recommendations for process control based on the obtained data on the results of the

	prediction based on the modified AIS algorithm.
Assistant agent	The agent contains reference information about the algorithms and principles of data processing. Assistant functions are necessary for correct input of information into the system.
Agent manager	Agent Manager coordinates the work of other agents.

6. Simulation results

6.1. Technological process of oil gases cleaning from acidic components at installation 300

Researches were conducted on the real production data of the TengizChevroil company. The enterprise consists of several complex technological lines (CTL) and installations (Figure 4).

As the control object we will consider the Installation 300, which is intended for the petroleum gases cleaning from acid components: hydrogen sulfide, carbon dioxide and carbon dioxide. Acidic components are removed using a solution of diethanolamine [21]. The raw materials of the U300 are high and medium pressure oil gases coming from Installation 200 (U200) intended for gas separation. The final products of the U300 are: purified high-pressure gas entering the Installation 700 for fractionation; refined medium pressure gas returned to the U200 for compression; acid gas with a high content of hydrogen sulfide, is sent to the installation 400 for the production of sulfur. In general, the U300 consists of the following main components: high-pressure gas cleaning, medium-pressure gas cleaning, dietalomine regeneration, amine filtration, chemical supply, etc.

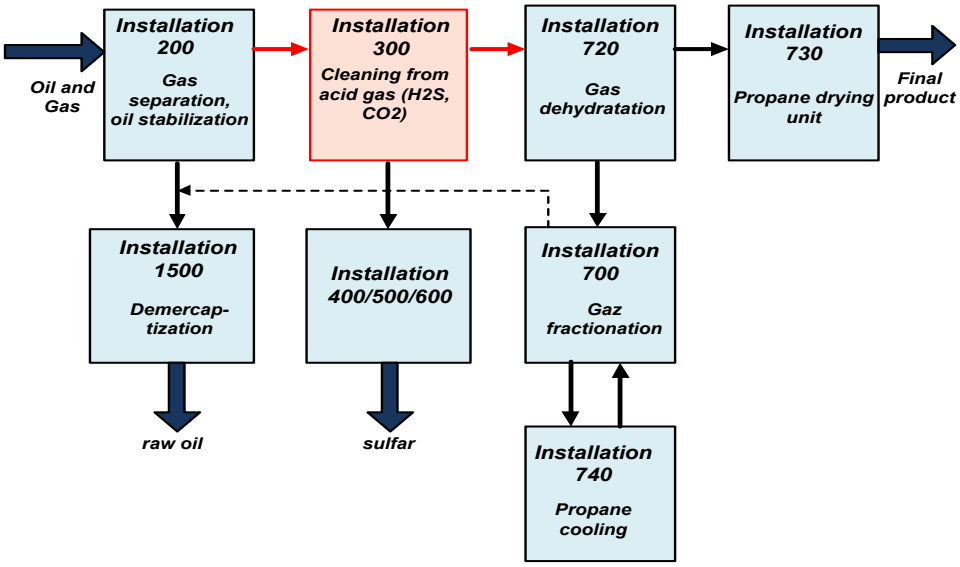


Figure 4: The structure of the Tengizchevroil enterprise

Table 2: Specification of medium pressure gas cleaning equipment

Designation on the diagram	Unit Name
D-302	Medium Pressure Absorber
F-328	Medium Pressure Separator and Coalescer
F-324	Medium Pressure Filter-Separator
F-320	Separator
F-319	Separator
F-303	Saturated Amine Tank
F-203	Low pressure separator
GC-200	First stage line
GC-201	Second stage suction line
EA-302	Regenerator reboolers (4 heat exchangers)
EA-304	Regenerated / saturated amine heat exchanger

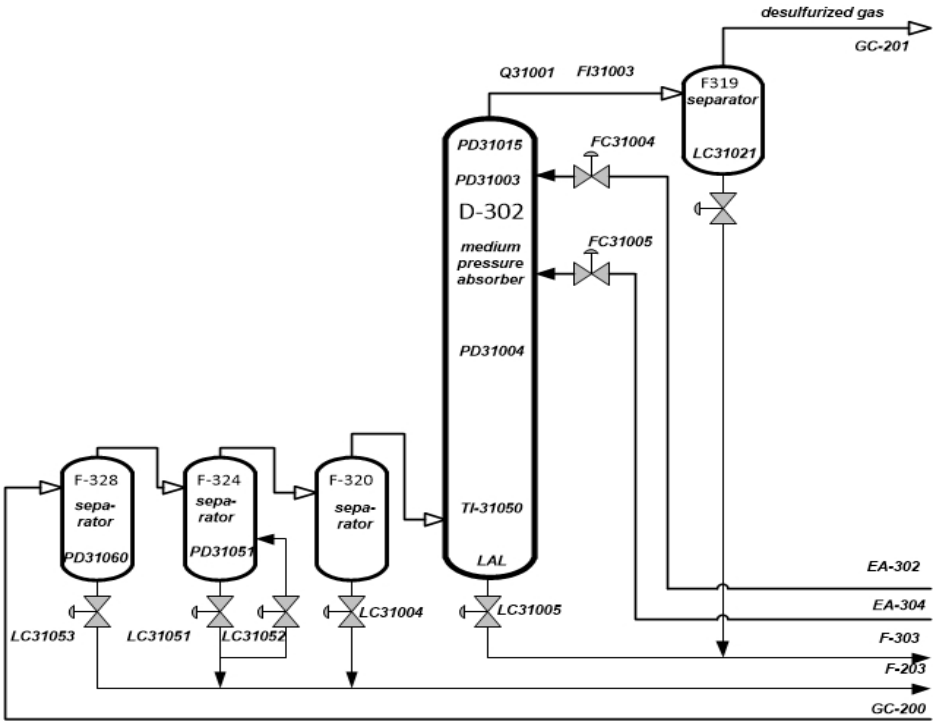


Figure 5: Fragment of the U300 technological scheme, medium pressure gas cleaning

The technological process of medium pressure gas cleaning in the absorber D-302 (medium pressure absorber) was chosen as a complex control object.

6.2. Medium Pressure Gas Cleaning of the Installation 300

Let consider a fragment of the U300, a medium pressure gas cleaning process (Figure 5, Table 2). The flow of raw gas enters the medium pressure absorber D-302 from the U200. In separators F-328, F-324, F-320 the hydrocarbon condensate is removed from the raw gas and transferred to the low-pressure separator F-203. This process is controlled by level controllers: LIC31053, LIC31051, LIC31052 and LIC31004.

Next, gas from the F-320 enters D-302. In the medium pressure absorber, the absorption process of H_2S and CO_2 also takes place due to the contact with two streams of the solution of regenerated diethanolamine [21], the temperature of which should not be lower than 540° in order to prevent condensation of hydrocarbons. In a medium pressure absorber, a pressure is approximately 24.5 bar. Gas temperature is measured using TI31050.

The absorber D-302 in the upper part consists of 11 valve plates, the pressure drops on which are measured using PDI31015, PDI31003 and PDI31004. In the case of a high pressure differential, plates may be blocked or foamed in the absorber. In order to prevent foaming, the antifoam injection points are located on the supply line of the regenerated amine in the column to the flow regulators: FIC31004 and FIC31005. The output of saturated dietalomine to the F-303 receiver is regulated by the LIC31005 regulator.

Next, the purified gas passes through an F-319 separator, where about 99% of the amine solution is removed in droplet particles of 10 microns and more in size and discharged to F-303. The output of the amine solution in F-303 is controlled by the LIC31021 level controller. Sulfured gas from F-319 is directed to the suction line of the second stage of GC-201. The consumption of diethanolamine in the upper section is regulated by the flow regulator FIC31004, and in the lower section by the FIC31005. The consumption of purified gas from D-302 is measured using FI31003, the concentration of H_2S is measured by QRAH31001. A total of 457 sensors are located on the U300, information from which is collected on the basis of the distributed Honeywell DCS control system. Sensor specifications are presented in Table 3.

As an example, let consider daily measurements from the U300 medium-pressure gas cleaning unit. A fragment of the parameters database is presented in Table 4.

6.3. Simulation of the developed Smart technology on the example of a complex object of the medium pressure absorber

Modern distributed control systems, such as the Honeywell DCS during the operation collect a huge amount of production data, the processing of which in industrial operation conditions is a complex task. Particular attention is paid to industrial safety, in connection with which there are different priorities of the alarms: emergency, high, low, logging and sending to the printer, writing to the log, sending to the printer and no action [21]. In this case, there are three alarms states: on, off, disabled. Therefore, for the application of the developed Smart technology there are selected 5 classes, depending on the state of the object for an alarm:

1 class. The alarm priority - emergency, the alarm state is on;
2 class. The alarm priority - high, the alarm state is on;
3 class. The alarm priority - low, the alarm state is off;

Table 3: Specification of sensors for the U300 circuit fragment, medium pressure gas cleaning block

№	Number of position	Purpose of the device	Measure	Measurement range	Manufacturer
D1	*LIC31053*	buoy level gauge	kg/1007 мм	0,046-0,75	Fisher 2390B-249BF
D2	*LIC31051*	buoy level gauge	kg/559 мм	0,051-0,843	Fisher DLC3010-249BF
D3	*LIC31052*	buoy level gauge	kg/631 мм	0,058-0,949	Fisher 2390B-249BF
D4	*LIC31021*	buoy level gauge	kg/500 мм	0,029-1,054	Fisher 2390B-249BF
D5	*LIC31004*	buoy level gauge	kg/500 мм	0,034-0,724	Fisher 2390B-249BF
D6	*LIC31005*	difference converter	mbar	0-294,2	Rosemount 3051CD2
D7	*TI31050*	temperature converter	°C	0-100	Honeywell STT350-0-EP0
D8	*PDI31003*	pressure drop	mbar	0-280	Rosemount 3051CD3A
D9	*PDI31004*	pressure drop	mbar	0-180	Rosemount 3051CD2A
D10	*FIC31004*	consumption regulator with indicator	mbar	0-600	Rosemount G1151DP4E
D11	*FIC31005*	consumption regulator with indicator	mbar	0-1124	Rosemount G1151DP5E
D12	*QRAH31001*	analyzer of H_2S in gas	ppm H_2S	0-50	Tracor Atlas 722R/102
Dn	...	…	…	…	…

Table 4: Fragment of the parameters database of medium pressure gas purification process of U300

№	*TI31050*	*LIC31053*	*PDI31003*	*…*	*QRAH31001*
1	85,563	41,815	229,829	…	44,509
2	57,315	7,017	243,282	…	43,887
3	96,944	92,453	186,916	…	47,597
4	85,257	44,324	232,445	…	237,979
5	85,844	47,465	233,306	…	46,529
6	87,093	50,434	247,162	…	45,659
7	83,714	33,813	215,779	…	49,894
8	86,189	40,041	228,184	…	46,105
9	85,688	41,387	228,406	…	49,168
10	83,381	32,606	213,857	…	45,197
11	82,75	31,564	212,562	…	44,564
12	84, 76	30,187	211,865	…	43,765
…	…	…	…	…	…
800	74,379	51,453	220,456	…	25,648

4 class. The alarm priority - writing to the log and sending to the printer, the state of the alarm - prohibited;
5 class. The alarm priority - no action, the state of the alarm is prohibited.

Remark. The creation of classical mathematical models for complex control objects is a time consuming task. As a mathematical model for research, there is considered a time series consisting of records from the medium-pressure absorber sensors.

As a modeling environment, modern Rapid Miner software products and WEKA operators were used. Let consider a fragment of the DB_D302 database describing the state of the complex control object of the absorber D-302 at medium pressure gases cleaning on the U300 consisting of 12,000 copies of data. Figure 6 shows the visualization of results from the DB_D302 database sensors According to the 4th algorithm, the database DB_D302 is pre-processed.

Then, within the framework of the multi-algorithm approach, the PCA algorithm is applied in order to reduce the low-informative features that characterize the behavior of the complex control object *D-302.* Figure 7 shows the simulation results in the form of a graph of attributes ranking by degree of importance, as well as the visualization of the *DB_D302* database after the non-informative parameters reduction. Similarly, the reduction of low-information descriptors is performed using the Random Forest algorithm (Figure 8) and the PSO swarm intelligence algorithm (Figure 9).

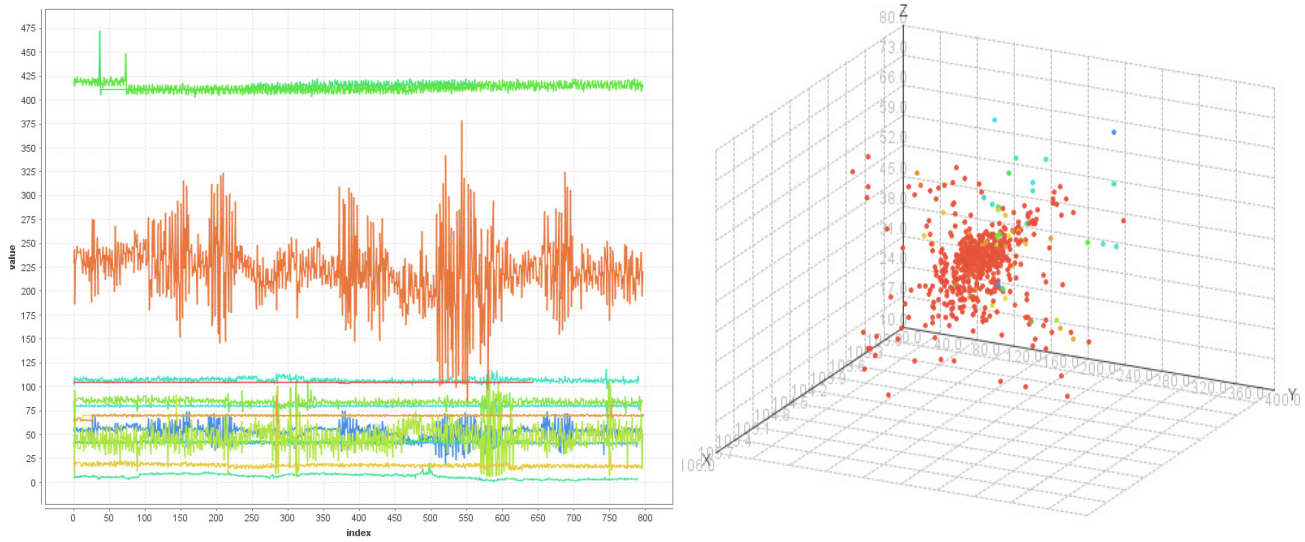


Figure 6: Fragment of *DB_U300_D302* database visualization in 2D and 3D form

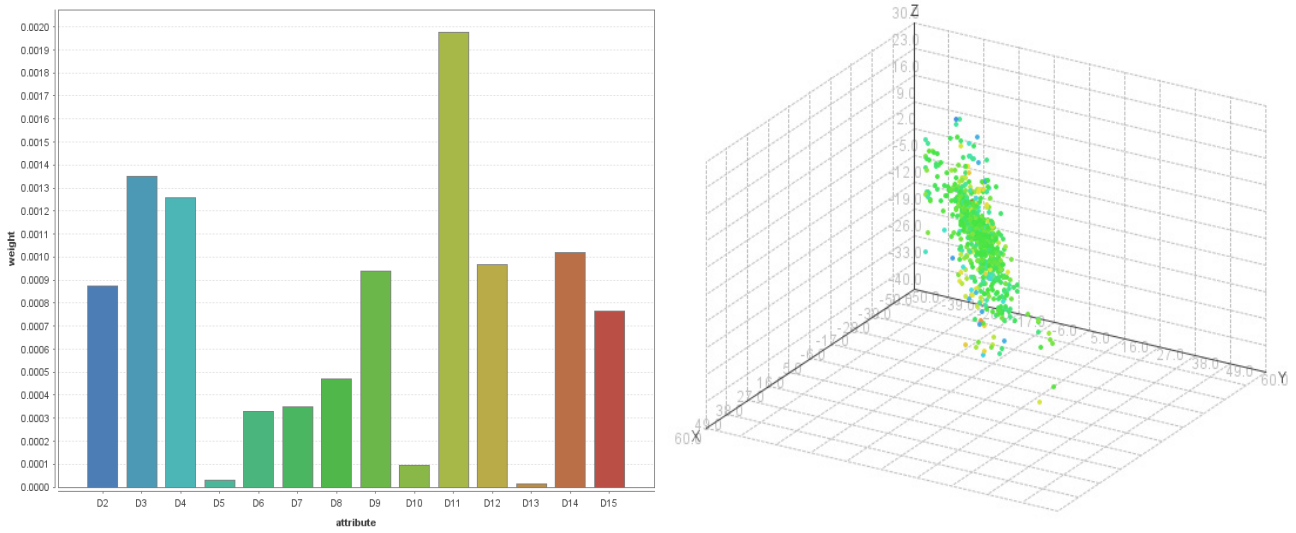


Figure 7: Low-informative parameters reduction using PCA

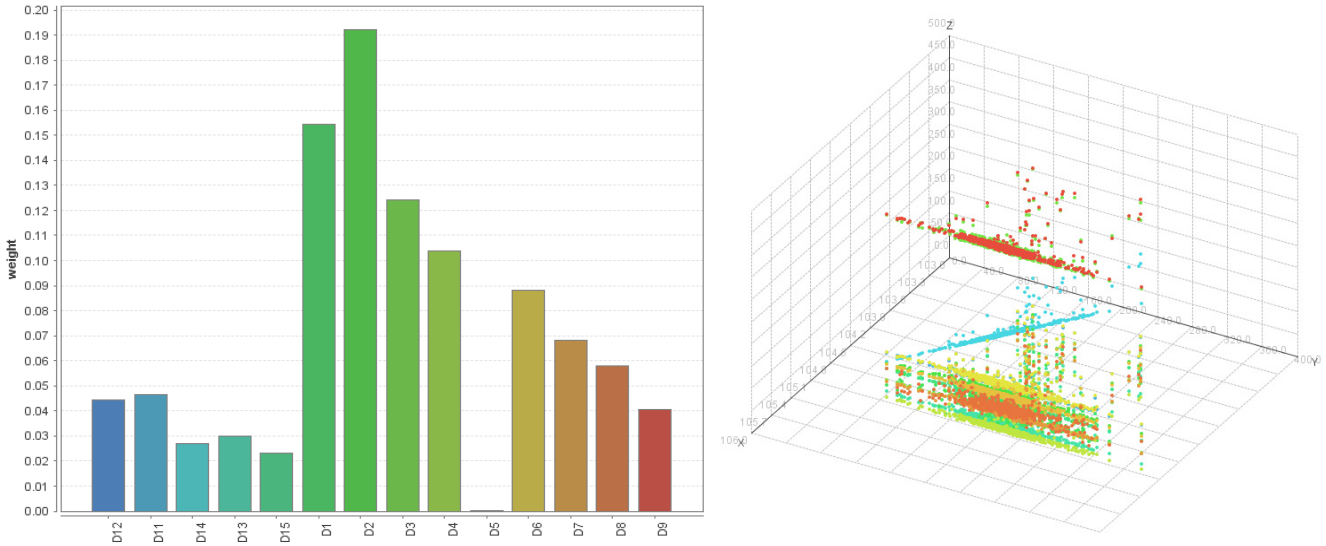


Figure 8. Low-informative parameters reduction using Random Forest

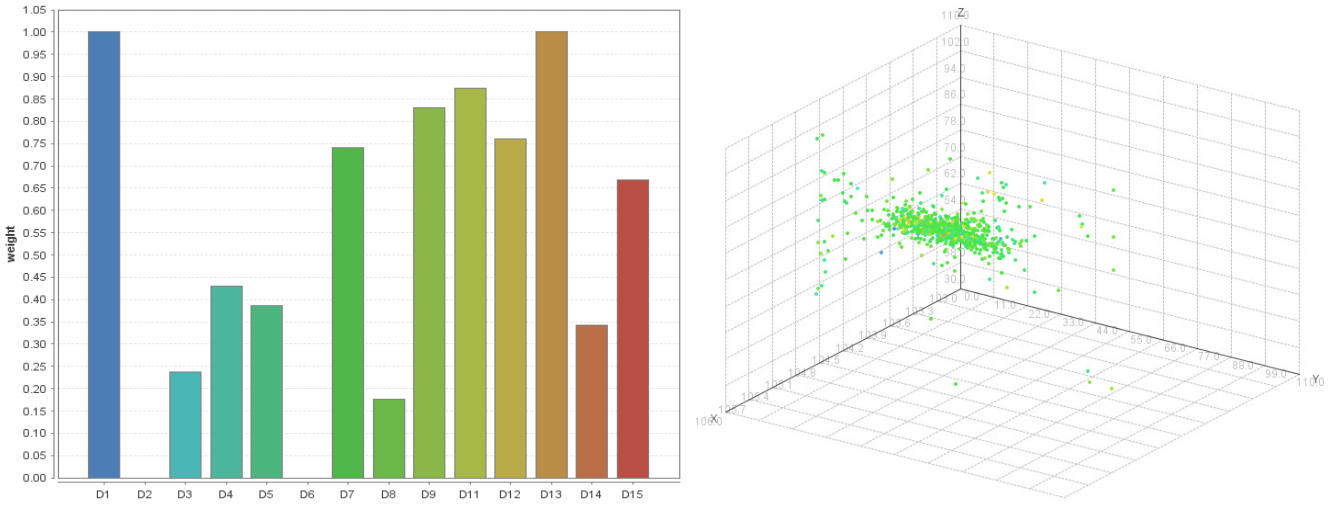


Figure 9. Low-informative parameters reduction using PSO

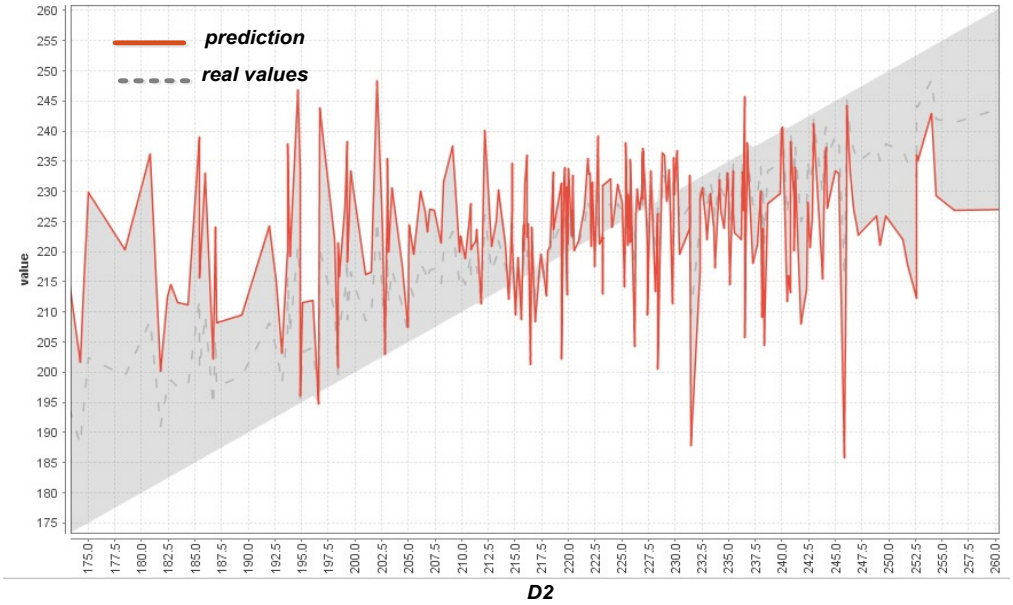


Figure 10. Prediction based on modified algorithm *PSO - Clonal Selection*

Further, according to the 4th algorithm, after the selection of informative features, the task of classification is solved using the AIS algorithm, based on clonal selection. A comparative analysis of the simulation results is presented in Table 5.

Table 5: Comparative analysis of image recognition results based on clonal selection

Feature selection method	PCA	RF	PSO
Performance of feature selection	87,67%	95,45%	94,56%
Runtime of feature selection	1 min 25 sec	7 min 34 sec	1 min 4 sec
Performance of clonal selection classification	85,53%	86,74%	93,67%
Runtime of feature selection	56 sec	2 min 59 sec	34 sec

Because the Honeywell DCS distributed control system from the Installation 300 collects a huge amount of production data, the required time to simulate multidimensional data is a necessary criterion for evaluating the effectiveness of the modified algorithm.

In the case of the use of intelligent algorithms in the development of control systems, it is impossible to use classical sustainability criteria in order to assess the effectiveness of the developed Smart technology. The effectiveness of the heuristic algorithms is carried out using the calculation of such parameters as [35]:

- *accuracy* - accuracy (relative amount of correctly classified examples, percentage of correct predictions);
- *classification_error* - the percentage of incorrect predictions;
- *kappa* - kappa statistics calculation;

- *weighted_mean_precision* - weighted average value for each class of measurement accuracy;
- *spearman_rho* - rank correlation between actual and predictable labels using Spearman's rho measure, which shows a linear relationship between two variables;
- *kendall_tau* - rank correlation using Kendall's tau measure, showing the strength of the relationship between two variables;
- *absolute_error* - average absolute deviation of the predicted values from the actual value;
- *relative_error* - average absolute deviation of the forecast from the actual value divided by the actual value;
- *relative_error_lenient* - average relative error, which shows the average absolute deviation of the forecast from the actual value divided by the maximum of the actual value and the forecast;
- *relative_error_strict* - average severe relative error;
- *normalized_absolute_error* - normalized absolute error;
- *root-relative-squared error* - averaged relative quadratic error, etc.

Table 5 shows the final accuracy parameter showing the percentage of correct predictions by the proposed methods.

Therefore, according to the research results within the multi-algorithm approach, the best prognostic result for solving the problem of prediction and control of the process of medium-pressure gases cleaning based on real data of the control object D-302 is a modified algorithm *PSO - Clonal Selection*. A visualization of a fragment of the prediction results based on the *PSO - Clonal Selection* algorithm is presented on Figure 10.

7. Conclusion

As a result of the research, there was proposed a Smart technology for complex objects control based on Honeywell's distributed control system and an artificial immune system approach. The scientific novelty of the proposed technology is following:

1. The technology allows the prediction and control of complex objects of the oil and gas industry based on the intellectual analysis of historical data obtained from the distributed control system of Honeywell DCS.

2. There was developed a modified algorithm of artificial immune systems on the basis of a multi-algorithm approach in order to improve the quality of prediction of the behavior of complex objects under industrial operation.

3. There has been created a multi-agent Smart system that allows the integration of the proposed modified artificial intelligence algorithms into real industrial control systems.

4. Presented the simulation results on the basis of a real complex facility of the TengizShevroil enterprise (medium pressure absorber) of a fragment of the installation line 300 for gas cleaning from acid components.

Therefore, the proposed Smart technology allows the implementation of artificial intelligence algorithms based on real industrial production, in case when the classical methods of control theory are not efficient and unsuitable for working with large amount of data.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The work was carried out under the grant of the KN MES of the Republic of Kazakhstan on the topic: AP05130018 "Development of cognitive Smart-technology for intellectual complex objects control systems based on artificial intelligence approaches" (2018-2020).

References

[1] D. O'Brien, "Smart operations", Pulp and Paper Canada, 2018.
[2] D. Yongsheng, Z. Lei, H. Kuangrong, Bio-Inspired clloaborative intelligent control and optimization, Springer, 2018.
[3] R. Hernandez, M. Lopez-Franco, E. Sanchez, A. Alanis, J. Ruz-Hernandez, Decentralized neural control, application to robotics, Springer, 2017.
[4] T. Wang, H. Gao, J. Qiu, "A combained adaptive neural network and nonlinear model predictive control for multirate networked industrial process control" IEEE Transactions on Neural Networks and Learning Systems, 27(2), 416-425, 2016. http://doi.org/10.1109/TNNLS.2015.2411671
[5] G. Yan, H. Fang, X. Chen, L. Shi, "Process control parameters optimal selection based on particle swarm algorithm", in 8th International Symposium on Computational Intelligence and Design, Hangzhou, China, 2015. http://doi.org/10.1109/ISCID.2015.195
[6] K. Lakshami Narayana, V. Naveen Kumar, M. Dhivya, R. Prejila Raj, "Application of ant colony optimization in tuning a PID controller to a conical tank", 8(s2), 217-223, 2015. http://doi.org/10.17485/ijst/2015/v8iS2/61635
[7] A. Ramzy, A. Aldair, A. Almousawi, "Design an optimal PID controller using artificial bee colony and genetic algorithm for autonomous mobile robot", International Journal of Computer Applications, 100(16), 8-16, 2014. http://doi.org/10.5120/17607-8016
[8] K. Goher, S. Fadlallah, "Bacterial foraging-optimized PID control of a two-wheeled machine with a two-directional handling mechanism", Robotics and Biomimetics, 4(1), 2-19, 2017. http://doi.org/10.1186/s40638-017-0057-3.
[9] R. Precup, R. David, E. Petriu, "Grey wolf optimizer algorithm – based tuning of fuzzy control systems with reduced parametric sensitivity", IEEE Transactions on Industrial Electronics, 64(1), 527-534, 2017. http://doi.org/10.1109/TIE.2016.2607698
[10] G. Ponticelli, S. Guarino, V. Tagliaferri, O. Giannini, "An optimized fuzzy-genetic algorithm for metal foam manufacturing process control", The International Journal of Advanced Manufacturing Technology, 1-12, 2018. http://doi.org/ 10.1007/s00170-018-2942-5
[11] S. Mustapha, K. Faycal, S. Mohammed, "Application of artificial immune algorithm-based optimisation in tuning a PID controller for nonlinear systems", Int. J. Automation and Control, 9(3), 186-200, 2015. http://doi.org/ 10.1504/IJAAC.2015.070955
[12] Y. Peng, X. Luo, W. Wei, "A new control method based on artificial immune adaptive strategy", Electronika ir elektrotechnika, 19(4), 3-8, 2013. http://doi.org/10.5755/j01.eee.19.4.1246
[13] J. Daudi, "An overview of application of artificial immune system in swarm robotic systems", Advances in Robotics & Automation, 4(127), 1-6, 2015. http://doi.org/10.4172/2168-9695.1000127
[14] H. Rezende, M. Silva, M. Santos, L. Honório, L. Silva, V. Vidal, J. Ribeiro, A. Cerqueira, A. Pancoti, B. Reginax, "Signal Estimation for UAV Control Loop Identification Using Artificial Immune Systems", in 22nd International conference on system theory, control and computing, 579-584, 2018. http://doi.org/ 10.1109/ICSTCC.2018.8540706
[15] F. Yao, F. Wang, M. Zhang, "Weak thruster fault detection for autonomous underwater vehicle based on artificial immune and signal pre-processing", Advances in mechanical engineering, 10(2), 1-13, 2018. https://doi.org/10.1177/1687814018758739
[16] J. Marciniak, K. Wawryn, P. Widulinski, "An Artificial Immune Negative Selection Algorithm to Control Water Temperature in the Outlet of the Chamber", in International Conference on Signals and Electronic Systems, 236-241, 2018. https://doi.org/10.1109/ICSES.2018.8507293
[17] Y. Ding, Nan Xu, S. Dai, L. Ren, K. Hao, B. Huang, "An immune system inspired reconfigurable controller", IEEE Transactions on Control Systems

Technology, 24(5), 1875 – 1882, 2016. https://doi.org/10.1109/TCST.2016.2514847
[18] L. Puggini, J. Doyle, S. McLoone, "Fault detection using random forest similarity distance", IFAC-PapersOnLine, 48(21), 583-588, 2015. https://doi.org/10.1016/j.ifacol.2015.09.589
[19] K. Hanson, S. Yella, M. Doughert, H. Fleyeh, "Machine learning Algorithms in Heavy Process Manufacturing", American Journal of Intelligent Systems, 6(1), 1-13, 2016. https://doi.org/10.5923/j.ajis.20160601.01
[20] L. Yang, J. Zhang, F. Deng, J. Chen, "Sensor fault diagnosis based on on-line random forests", in 35th Chinese Control Conference, 1934-1768, 2016. https://doi.org/10.1109/ChiCC.2016.7553992
[21] Technological regulations for the technological process of hydrocarbon gases cleaning at the plant 300. TengizShevroil LLP, 2016.
[22] J. Brownlee, "Clonal selection algorithms", CIS Technical Report 070209A, 2007, 1-13.
[23] L. de Castro, F. von Zuben, "The Clonal Selection Algorithm with Engineering Applications", Proceedings of Workshop on Artificial Immune Systems and Their Applications, 36-37, 2000.
[24] U. Berna, S. Kulturel-Konak, "A review of clonal selection algorithm and its applications," Artificial Intelligence Review, Springer, 36(2), 117–138, 2011. https://doi.org/10.1007/s10462-011-9206-1
[25] J. Brownlee, "Clever Algorithms, Nature-Inspired Programming Recipes", 2011.
[26] J. Kennedy and R.C. Eberhard, "Particle swarm optimization", in IEEE International Conference on Neural Networks, 4, 1942-1948, 1995. https://doi.org/10.1109/ICNN.1995.488968
[27] J. Blondin, "Particle Swarm Optimization Applications in Parameterization of Classifiers", Armstrong Atlantic State University, 1-33.
[28] G.A. Samigulina, Z.I. Samigulina, "Modified immune network algorithm based on the Random Forest approach for the complex objects control", Artificial Intelligence Review, 2018. https://doi.org/10.1007/s10462-018-9621-7
[29] L. Breiman, A. Cutler, "Random Forest", Berkeley, 2005.
[30] S.P. Chistyakov, "Random forests: a review", in proceedings of the Karelian Scientific Center of the Russian Academy of Sciences, 1, 117-136, 2013.
[31] L. Salazar, F. Mayer, D. Schütz, B.Vogel-Heuser, "Platform Independent Multi-Agent System for Robust Networks of Production Systems", IFAC Proceedings Volumes, 51(11):1261-1268, 2018. https://doi.org/ 10.1016/j.ifacol.2018.08.359.
[32] A. Sayed, S. Hanafi, A. Hassanien, "Multi-agent Artificial Immune System for Network Intrusion Detection and Classification", Proceeding of International Joint Conference SOCO'14, 145-154.
[33] S. German, S. Shin, A. Tsourdos, "Immune-system inspired approach for decentralized multi-agent control," in IEEE 24th Mediterranean Conference on Control and Automation (MED), 1020 – 1025, 2016. https://doi.org/10.1109/MED.2016.7536048
[34] G.A. Samigulina, Zh.A. Massimkanova, "Multi-agent system for forecasting based on modified algorithms of swarm intelligence and immune network modeling," in 12th International Conference Agents and Multi-agent Systems: Technologies and Applications (KES-AMSTA-18), 199-208, 2018. https://doi.org/10.1007/978-3-319-92031-3_19
[35] M. Hofmann, R. Klinkenberg, "RapidMiner: Data Mining Use Cases and Business Analytics Applications", 2014, 48-49.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 88-99 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Analysis of Fuzzy Data for the Selection of Ingredients of the Polymer Composition by Technical Functions in Quality Management

Ilya Vasilyevich Germashev[*,1], Tatyana Petrovna Mashihina[1], Evgeniya Vyacheslavovna Derbisher[2], Vyacheslav Evgenyevich Derbisher[3], Mikhail Alexeevich Kharitonov[4]

[1] *Department of Mathematical Analysis and Function Theory, Volgograd State University, 400062, Russia*

[2] *Department of Analytical, Physical Chemistry and Polymers Physico-Chemistry, Volgograd State Technical University, 400005, Russia*

[3] *Department of Technology of Macromolecular and Fibrous Materials, Volgograd State Technical University, 400005, Russia*

[4] *Department of Fundamental Informatics and Optimal Control, Volgograd State University, 400062, Russia*

ARTICLE INFO

Article history:
Received: 07 April, 2019
Accepted: 08 May, 2019
Online: 21 May, 2019

Keywords:
Data analysis
Fuzzy numbers
Polymer composites
Mathematical support
Intelligent system
Quality
Optimal choice
Management
Physical and chemical properties

ABSTRACT

The task of fuzzy mathematical modeling of the composition of a polymer composition is singled out and formulated, as one of the types of chemical-technological systems for creating the theoretical foundations for managing the physicochemical properties of the material produced from this composition. Fuzzy data on the technical functions and concentration of individual ingredients that included in the composition in the form of fuzzy numbers are used as the initial formalized information at modeling the polymer composition. As a result, a parametric space is formed that describes the multicomponent system under study. The index of compliance of each parameter of the polymer composition with the requirements for the physicochemical properties and quality of the polymer material is calculated. This allows you to go to the relative dimensionless real values. Weighted voting procedure is used as an aggregate function. The obtained numerical value reflects the integral measure of the compliance of the chemical-technological system with the requirements for the polymer composition.

An example of the choice of ingredients of the polymer composition by calculating using fuzzy numbers is presented. In addition to the formulation of the problem, the study develops the mathematical software for the intellectual system supporting the research and production activities of the chemical technologist to substantiate the decision to choose a polymer composition.

Accepted abbreviations: PC — polymer composition, CTS — chemical-technological system, PCM — polymer composite material, FS — fuzzy set, IS — information system, DNA — Deoxyribonucleic acid

1. Introduction

This study is an extension of work originally presented in International Russian Automation Conference [1]. The study gives a new vector of application of fuzzy mathematics methods in the analysis of complex multifactor technical systems, including chemical-technological system (CTS), for their identification and pre-experimental research. These systems, as a rule, are described by fuzzy information: the presence of linguistic descriptions, intervals, lack of data, incompatible research results, conflicting expert opinions, etc. Under these conditions, mathematical modeling using this information also has a fuzzy nature and a multiplicity of solutions. Nevertheless, the use of fuzzy mathematics, fuzzy modeling allows at the preliminary stage to choose the optimal and economical strategy of the experiment and evaluate the expected technical result, which is directly related to product quality. It should be noted that the use of fuzzy set (FS) for the preparation of initial information in the simulation of CTS has recommended itself from the positive side, as a form of theoretical

[*]Ilya Vasilyevich Germashev, Department of Mathematical Analysis and Function Theory, Volgograd State University, Prospect Universitetskii, 100, Volgograd, 400062, Russia, germashev@volsu.ru

analysis of polymer composition (PC) and finding the relationship between the components and properties of the polymer system as a whole. Here, a prerequisite is the identification and ranking of the technical functions of the ingredients of the PC. Managing the properties of polymer composite material (PCM) is an important task of chemical technology, it can be extended to: oil and gas processing (primary separation, refining, age hardening), petrochemical and organic synthesis, biotechnology, the technology of high-molecular substances, the technology of processing plant and animal raw materials and etc. [2].

The formation of the PC optimization problem and its solution is based on the "composition – property – quality – application" dependence and provides for the selection of a certain number of ingredients with the necessary technical functions using fuzzy modeling. In this paper, this idea is extended to the PCM analysis.

Raw PC, in other words, mixtures, are complex systems. They are based on polymer matrices in which various ingredients are added. The interaction between such matrices, their structure and composition also depend on the chemical and technological processes of their processing. PC-derived composites are also complex heterogeneous systems and can be divided into five leading classes:

1. Matrix, consisting of a continuous phase (matrix) and a dispersed phase (discrete particles) (dispersed — hardened composite materials).
2. Compositions with fibrous fillers.
3. Mixed compositions.
4. Compositions having an interpenetrating structure of two or more continuous phases.
5. Layered PCMs.

Considering the special interest of researchers and the market of consumers of polymer products to nanotechnology, let us mention the polymer matrix of the "polymer matrix – additive (filler)" composition, in which 2D and 3D nano-ingredients are used as fillers — graphene and carbon nanotubes. Here, mathematical modeling can be conducted based on the analysis of ingredients in the direction of predicting the level of physical and mechanical properties, depending on the number and size of nanoparticles and the parameters of the polymer matrix of a stable configuration. This direction is at the beginning of development, therefore, in most cases, problems are solved on the basis of experiments [3, 4]. Here the applicability of fuzzy mathematics is quite possible when calculating and analyzing results. Thus, the PC is ultimately a multi-component, inhomogeneous system, the characteristics of which do not allow to calculate the exact values of technical effects by its component. The versatility of the situation emphasizes the fact that in addition to a large set of active and inactive fillers, modern technology with a different purpose gives rise to the use of new modifying additives, a brief description of which is proposed below. In parentheses are their possible concentrations in the PC.

Light stabilizers allow for a long time to maintain the original color of the product and its mechanical strength, providing protection from ultraviolet rays. Here you can point to derivatives of benzonitrizol, benzophenol, nickel and cobalt salts of substituted phenols, etc. It has been established experimentally that they are effective even at very low concentrations (0.1 –5.0%).

Antioxidants prevent thermo-oxidation and slow down destruction during extrusion and injection molding. Also, these additives are used to protect the polymer when processing in aggressive environments. The mixtures compounds of phenols and phosphides (0.1 – 2.0%) are the most popular here.

Plasticizers are low molecular weight organic substances that, when introduced into the composition at the preparation stage, reduce the interaction between adjacent macromolecules. At the same time, common plasticizers include esters of organic and inorganic acids, such as dibutyl phthalate, di-(2-ethylhexyl) phthalate, dibutyl sebacate, tricresyl and tributyl phosphate, refined products, (petroleum oil), coal and wood materials (cumarone indene resin, rosin, etc), epoxidized vegetable oils, etc. The temperature of decomposition of plasticizers should be higher than the processing temperature of PC (0.1 – 45.0%).

Combined additives (light stabilizer + antioxidant + dye, etc.) are widely used in the production of agricultural films. They allow to use the polymer for its intended purpose after recycling (0.1 – 45.0%).

Flame retardants are the most important additive limiting the flammability of PCM. Special flame retardants are used for the production of non-combustible and low-flammable films, sheets and molded products. They are divided into three large groups:

- chemically interacting with the polymer;
- intuminescent;
- additives mechanically miscible with polymer.

Flame retardants are most often halogen-containing and phosphorus-containing compounds, as well as metal hydroxides and other inorganic compounds (0.1 – 3.0%).

Antistatics can get rid of the static effect. For example, in the manufacturing of cases for audio and video equipment. Commonly used antistatic agents include alkylamines (up to 3.0%).

Slip agents are an internal lubricant in a polymer matrix that reduces melt viscosity, improves extrusion performance, makes the surface smoother, shinier and glossy, and reduces the coefficient of friction of products. Active ingredients are derivatives of higher fatty acids (depending on technology).

Anti-blocking additives prevent sticking, that is especially important in the production of polypropylene films. The packages made of a film with such additives are easily opened. Package walls do not stick together. Active ingredients are silicic acid or amide waxes (0.1 – 0.3%).

Antislips are additives that are similar in properties to antiblocking. They give the surfaces of the films roughness. They are used in the manufacturing of packaging, bags for packing bulk products and other products. One of the antislips is ultra high molecular weight polyethylene (0.1 – 0.3%).

Antiseptics prevent the biological destruction of PCM (up to 2.0%).

Nucleates affect the supramolecular structure of the polymer matrix, making it more fine-grained. At the same time, the

crystallization time is significantly reduced, the product is cooled faster. In addition, the effect of greater transparency of the films is created. Nucleates are used exclusively for polypropylene (0.1 – 0.3%).

Foaming additives are used to manufacture porous polypropylene and polyethylene products. The structure of the matrix is loosened, the extensions disappear with the introduction of additives. They are designed to produce decorative films used in the packaging of gift sets, as decorative ribbons, gaskets in cork and other purposes (depending on the requirements of technology).

Modifier in the production of stretch films. It gives them the effect of additional elasticity, and for some polymers, including polypropylene, impact resistance (no information).

This list is rapidly replenishing and expanding today, for example, so-called processing additives that improve the economy and technological parameters of processing activities (in each case individually) can be singled out separately.

Today, PC theory is an attempt to solve the problem of how the final physicochemical properties of finished products depend on the composition, chemical-technological and physical-technical parameters. The ultimate goal here is to obtain materials with desired properties. To achieve this goal it is necessary, setting the final properties of the designed material or product, to know the parameters of their manufacture.

General requirements for PC ingredients are as follows:

a) thermodynamic compatibility with the polymer matrix should be ensured, preferably before the formation of a molecular solution;

b) should not be volatile;

c) do not bleed out of the polymer matrix;

d) should be nontoxic, chemically and thermally resistant.

As an example (Table 1) [5], we consider the composition (in weight parts) of an electrically insulating PC from the Russian patent which is well illustrating the multifactorial nature of the identification problem and the versatility of PCM ingredients. And this is not the most difficult PC.

Therefore, for PCM ingredients, their effect on the polymer matrix is empirically determined. And to obtain a PC with the desired set of properties, both polymer matrices and individual ingredients are selected. The complexity of the problem is obvious, as is the complexity of the system.

It should be noted that for particular cases of the theoretical analysis of the properties of PCM, many CTS analysis methods are used: discriminant, factorial, regression, etc., to identify the control actions and calculate their consequences, while operating in numerical spaces [6 – 8]. But in the presence of nonnumeric, for example, linguistic parameters, such methods become of little use and require additional efforts to harmonize the source data and mathematical methods. Comparing and critically analyzing the existing mathematical approaches [2, 6 – 9] we can conclude that an important part of this problem is overcome by using FS for describing both numerical and nonnumeric parameters. The solution of the problem itself is carried out by methods of fuzzy mathematics. Important elements of the latter are increasingly being applied in the scientific analysis of the CTS [10 – 13].

Table 1: Composition of a Polymeric Composition (in Weight Parts)

Components	No.						
	1	2	3	4	5	6	7
PVC S-70*	*100*	*100*	*100*	*100*	*100*	*100*	*100*
Dioctyl phthalate	*40*	-	-	-	*70*	-	-
Diisononyl phthalate	-	*45*		*40*	-	*80*	-
Trioctyltrimeth yl metat	-	-	*50*	-	-	-	*50*
Isodecyldephen yl phosphate	-	-	-	*20*	-	-	*15*
Calcium carbonate	*40*	*60*	*80*	*100*	*120*	*140*	*160*
Tribasic lead sulfate	*3*	-	*5*	-	*4*	-	*6*
Dibasic lead phthalate	-	*4*	-	*3*	-	*5*	-
Antimony trioxide	*3*	*4*	*5*	*6*	*7*	*6*	*8*
Zinc oxide	*2*	*3*	*4*	*6*	*8*	*2*	*4*
Zinc borate	*2*	*5*	*6*	*3*	*3*	*8*	*4*
Diphenylolprop ane	*0.20*	*0.25*	*0.30*	*0.35*	*0.4*	*0.35*	*0.30*
Magnesium hydroxide	*20*	-	*40*	*25*	*60*	-	*80*
Aluminium hydroxide	-	*30*	*20*	*50*	-	*70*	-
Metallic silver in the form of nanoparticles	*0.001*	*0.005*	*0.01*	*0.015*	*0.02*	*0.025*	*0.03*

* PVC S-70 – polyvinyl chloride polymer matrix, Russian brand S-70.

In connection with the above, let us point out that the management of PCM properties is possible in various theoretical and further practical ways, such as:

1) pre-experimental heuristic identification,

2) the choice of the polymer type,

3) changes in production technology of polymer used as a matrix (polymerization, polycondensation, modification, crystallization, structuring),

4) regulation of the concentration and technical functions of the ingredients,

5) variation of composition and structure.

Among the newer modern experimental approaches to the management of PCM properties through the "composition – properties" relationship, for example, layer-by-layer deposition of polymers can be specified to control the conductivity of ophthalmic preparations through lenses [11], the regulation of technology to control the physicochemical properties of PCM [12 – 16], or increase, in the particular case, of the structural stability of such a biopolymer as DNA, when counter-ions are neutralized with the help of polyamines [15]. From what has been said, it is clear that controlling the properties of the polymer matrix, whether it be an olefin or DNA, is a separate scientific and technological task. Here, we have not found any works on the application of FS for the theoretical analysis of the polymer matrix — a full-scale experiment is fully used. So we have attributed points 4, 5 to the above list to our problems.

Here, the use of fuzzy analysis for the study of polymer systems and their particular case of PCM and comparison of mathematical procedures of fuzzy analysis [16, 17] made it possible to formulate the task of managing the properties of PCM and to obtain particular methods of its solution. Let us make the restrictions related to the subject area of this article and introduce it into the framework of the general task of studying CTS, while considering the PC as a system consisting of an undefined set and their number. Mathematical modeling and optimization of this CTS will be carried out under the conditions mentioned above, multicriteriality, incompleteness and vagueness of the initial information. This information can include the knowledge, experience and intuition of a specialist expert, which provides the basis for managing these systems — one of the main modern vectors of their development. We repeat that the identification of optimal and compromise technical solutions at the early stages of the formation of a plan for laboratory research and pilot product releases in world practice is known [18, 19], but it is still an urgent task focused both on the creation of new equipment and technology, and on the savings of intellectual and material resources. In principle, the problem of analyzing fuzzy data is relevant for analyzing the internal problems of the intelligent systems themselves, which are oriented toward interaction with both researchers and the consumer. As an example, one can cite a system for analyzing linguistic information for predicting the commercial attractiveness of goods [20] and other areas using problems with fuzzy data [21, 22].

The work of creating a fuzzy model is multidimensional and complex. It can be noted that in chemistry and technology, PCM tasks with fuzzy data are most of all:

- selection of compositions,
- evaluation of new materials properties,
- design of competitive technologies,
- consideration of possible environmental damage,
- safety accounting, etc.

Describing the problem of modeling CTS in this part, it is necessary to proceed from the fact that the very formulation of the problem of fuzzy modeling, computer implementation of calculations and their effectiveness can have, as discussed above, the nature of the set, which is largely determined by the formalization of the source data (parameters and conditions) and the adequacy of the test results obtained during the simulation. At the same time, depending on the level of fuzziness and the type of model obtained, the solutions may have different significance [22].

Let us add that modern research focused on processing fuzzy information uses the following route: "raw data – fuzzification – analysis – defuzzification – result". At the same time, they formulate variables, including linguistic ones, and, using conventional methods of fuzzing terms, go to FS and further, as for example, in [23] analyze the resulting data model in a mathematical package [24]. Also fuzzy components are explored using intelligent systems. Since the considered approach focuses on certain subject areas, the formalization of information is carried out in accordance with the requirements of these areas using a mathematical apparatus. This is often beyond the power of ordinary users of intelligent systems. In such cases, it is offered either detailed manuals, or to fully automate the process.

We performed the real work, which is the software of the information system (IS) summarizing the information available in this part and using our own experience in solving such fuzzy problems. Here it will be possible, firstly, to automate the initial processing of the raw data (to carry out the fuzzification process) and, secondly, to analyze the solutions for the information support of researcher work. The second part is focused on a wide range of chemical-technological tasks, such as: classification, identification, examination and management of the CTS. The specific analysis methods discussed below are focused on relative data and have no substantive substantiation; therefore, they can be applied in other fields of science and technology. We also point out that to solve the problems of our profile, we used the latest advances in the field of FS analysis [25 – 29], interval mathematics [30], and others [31], however, a universal method for processing fuzzy data, we have not found neither in Russian nor in foreign practice, although research in this direction is being actively conducted, as an analysis of scientific and technical literature of recent years shows.

This paper takes into account the complexity of the system and the PCM analysis and is aimed at developing a methodology for formalizing the source data in the form of fuzzy numbers, defining a library of mathematical models for selecting ingredients from a given set. Ultimately, information is provided for the quantitative support of the decision to choose PCM. Also, the typing of the source data will allow to unify and algorithmize the process of automated analysis of the results of the experiment.

2. Research objective

Now we will set a more specific task of forming the initial data, and then the mathematical method of calculation. We point out the conceptual fuzziness that exists and arises when describing the PCM:

- the technical functions of the ingredients are not presented;
- in the description of the ingredients of their concentration in the PC are given in the form of intervals;
- the level of action of the technical function of additives is characterized linguistically and is not related with their concentration in PCM;

- the possible relationship between the physicochemical parameters of the additives themselves and PCM and the positive or negative chemical interaction between them is not indicated.

Considering these fuzzinesses, which are more practical in practical technologies, the description of the data about the PCM by fuzzy numbers and operations with them by the methods of fuzzy mathematics will be adequate techniques. In general, the quality management of PCM will be carried out according to the scheme proposed in Figure 1 using fuzzy data analysis techniques.

Note that the main requirements for the PCM are formulated by consumers (customers) and process engineers, taking into account industrial capabilities and needs, on their part introduce additional requirements. If the former determine the range and tolerances of physicochemical, mechanical, ergonomic, environmental and other operational characteristics; the latter proceed from technological limitations, polymer matrix synthesis processes (polymerization, polycondensation, modification) and processing: (melting, mixing, casting, molding, stamping, extrusion, etc.) The formulation of such requirements is made by experts in the relevant fields, and, in many cases, is cited in reference data sources.

We assume that the basic consumer requirements already exist. It is necessary to construct the composition of the PCM that best meets these requirements and select ingredients with certain technological and technical functions.

The management of the design and analysis process of the PCM itself is carried out by assessment and selection of:

- polymer matrix,
- ingredients (including fillers),
- coordination of concentrations,
- justification of processing technology.

These actions need to be carried out in concert, since the nature of each polymer matrix determines a unique combination of possible ingredients and their content. You can get far from optimal low-potential low-quality PCM even choosing a polymer matrix with good physico-chemical characteristics and using an unsuccessful set of ingredients.

The following route is proposed for creating an original technical solution:

1. The choice of the polymer matrix is carried out on the basis of the analysis of the achieved level of technology by a variable method from the available kits, most often from the brand kits.
2. Determination of technical functions and concentrations of ingredients and their combinations is carried out through technical expertise.
3. Specific representatives are selected from the known ingredients, taking into account the available possibilities, and the compliance index for each of them is calculated.
4. PC is numerically evaluated based on the properties of the selected components by calculating the PC compliance index and, if necessary, proceeds to step 1.

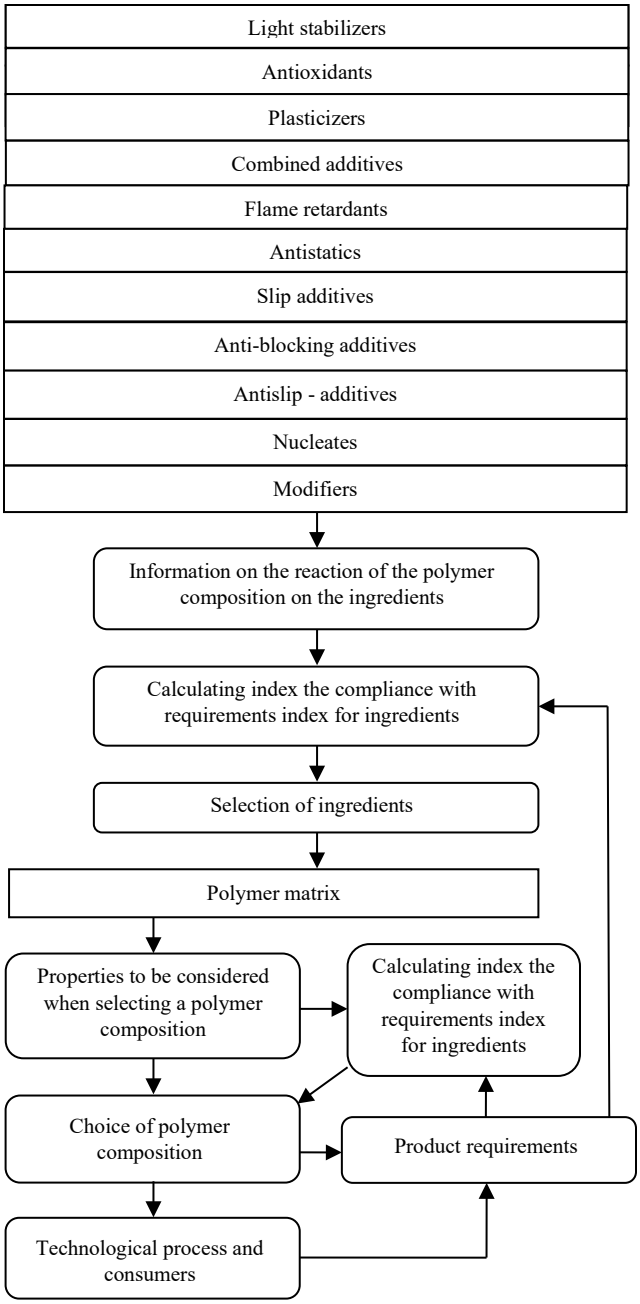


Figure 1: Control Scheme of the Physicochemical Properties of PCM and the Quality of Polymer Products

5. The decision is made on further actions: the choice of the polymer matrix, the specific composition of the PC, the individual ingredient, changes in the technological regime, etc.

Actions under item 1 are related to the experience of the technologist and researcher, the availability of information and material resources. They are obvious and, we believe, comments do not require. The actions under item 2 can be implemented either by designers or directly by process engineers — experts in the field of chemistry and PCM technology. It can also use reference books, databases and knowledge bases or knowledge systems created on the basis of generalization of subject information. In any case, the technology of these actions goes far beyond the scope of this article and is not discussed in detail here.

This article discusses in more detail the steps in paragraphs 3, 4 and 5 related with solving the problems of evaluating the technical functions of ingredients in a given set for designing a PC, in the context of the decision to choose its composition. A solution to the problem of optimal PC selection using a given polymer matrix and variability of additives is also proposed. In the simplest case, it is CTS type "one polymer matrix + one functional additive" or "polymer matrix + filler + functional additive". In the latter case, the "polymer matrix + filler" is considered inseparably. Then we will complicate the task.

We formulate the general task of determining compliance with the requirements of an abstract technical object and present the main points of its solution method, to assess the quality of the additive used to form a PC, and then the quality of PCM.

Let $Q = (Q_{ij})$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ be the matrix of input data describing the results of measurements or calculations of m characteristics of n representatives of a certain class of chemical systems $s_1, \ldots, s_n$, $Q_{ij} \in Q_j$, where Q_j is a set of values representing numbers, sets, words, etc.

Here and below, we will illustrate the reasoning with an example of the choice of a flame retardant for a polymer matrix presented in the form of a fibrous substrate (Table 2). At the same time, since this is not significant for the description of the mathematical methodology, the names of chemical compounds, as well as the dimension of physicochemical parameters will be omitted. The point is that it is necessary for those given in Table 2 data to evaluate the PC system being formed and in some sense choose the best option. The choice of flame retardant as an illustration of FS application in this case is related to its technical function (see above) and cannot be considered accidental or abstract. The fact is that in the practice of chemistry and chemical technology of high-molecular compounds, imparting noncombustibility of PCM or a significant decrease in combustibility is one of the most urgent tasks, along with increasing the durability of polymer products.

Table 2: Abstracted Characteristics of Industrial and Laboratory Flame Retardants [25]

Compound no., i	Q_{i1}	Q_{i2}	Q_{i3}
0 (Reference, q/δ)	*1.0/8*	*1.5/8.5*	*21/10*
1	*19*	*12.1*	*9*
2	*18*	*10.7*	*9*
3	*18*	*3.7*	*14*
4	*6 – 10*	*6 – 10*	*15*
5	*18*	*2*	*9*
6	*3*	*3*	*10.6*
7	*1.0 – 1.2*	*2.67*	*15*
8	*1.0 – 1.2*	*1.5 – 2*	*9 – 11*

It is clear that the method discussed below can be applied to any flame retardant, and in general, to an ingredient, both from the list discussed above and any other using any polymer matrix. The solution of the problem is divided into several stages, including: analysis of the initial data Q, construction of fuzzy numbers U, formalization of requirements in the form of fuzzy numbers, comparison of U with formal requirements, choice of PC, decision making.

3. Analysis of the source data

3.1. Method Description

The purpose of this stage is to bring different types of data to one type in the form of a segment. We use the operator proposed earlier [31], i.e. define the operator A, for this:

$$A(Q_{ij}) = [q_{ij1}; q_{ij2}],\ q_{ij1}, q_{ij2} \in \mathbb{R},\ q_{ij1} < q_{ij2},$$

where $\mathbb{R}$ — set of real numbers.

It is not possible to define this operation precisely, since in practice Q_{ij} can be represented in any form, but first you can describe the solution for the most typical forms by creating a library of data models, and later, if necessary, expand the library, which will allow you to adapt the proposed approach to almost any type of source data. Consider the most common cases.

1. Q_{ij} — real number.
2. Q_{ij} — set of real numbers.
3. Q_{ij} — linguistic value.

For all these cases, it is proposed to use the following operator.

$$A(x) = \begin{cases} A_1(x), \text{ если } x \in \mathbb{R} \times \mathbb{R}^+ \setminus \{0\}, \\ A_2(x), \text{ если } x \in \Re, \\ A_3(x), \text{ если } x \in t, \end{cases}$$

where

$$A_1: \mathbb{R}\times\mathbb{R}^+\setminus\{0\} \to \Im,\ \Im = \{[a; b] |\ a, b \in \mathbb{R},\ a < b\},$$

namely,

$$A_1(x, y) = [x - y; x + y],\ x \in \mathbb{R},\ y \in \mathbb{R}^+\setminus\{0\};$$

$$A_2: \Re \to \Im,\ \Re = \{X |\ X \subset \mathbb{R}\},$$

namely,

$$A_2(X) = [\inf X; \sup X],\ X \subset \mathbb{R};$$

$$A_3: t \to \Im_1,\ \Im_1 = \{[a; b] |\ [a; b] \subset [0; 1]\},$$

namely,

$$A_3(t_k) = [e_{k-1}; e_k].$$

Operator A_3 requires additional explanations on the set of terms t. Let be $Q_{ij} \in t = \{t_k |\ k = 1, \ldots, r\}$, and terms t_k are such that $t_1 \prec \ldots \prec t_r$, where $\prec$ is an ordering relation. We split the segment $E = [0; 1]$ into r parts in accordance with the terms t_k, while maintaining the same order, i.e.

$$E = \bigcup_{k=1}^{r} E_k ,$$

where $E_k = [e_{k-1}; e_k]$, $k = 1, \ldots, r - 1$, $E_r = [e_{r-1}; e_r]$, $0 = e_0 < e_1 < \ldots < e_r = 1$.

Illustrating with compound no. 1 from the example under consideration (Table 2), we obtain, for example, for parameter no. 1

$$Q_{11} = 19, \delta_{11} = 1,$$

$$q_{111} = 18, q_{112} = 20,$$

i.e.

$$A_1(19, 1) = [18; 20].$$

Further, similar to [25], fuzzy numbers are given for q values:

$$u_{ij}(x) = \exp\left(-\frac{\ln 2}{\delta_{ij}^2}(x - q_{ij})^2\right),$$

where $q_{ij} = (q_{ij1} + q_{ij2})/2$, $\delta_{ij} = (q_{ij2} - q_{ij1})/2$, $x \in \mathbb{R}$. Main provisions of fuzzification described in [25] are given in a more generalized form in clause 3.2.

As a result, a matrix of fuzzy numbers $U(x) = (u_{ij}(x))$, $i = 1, \ldots, n, j = 1, \ldots, m$ (Table 3) is obtained.

Further it is necessary to formalize the requirements for representatives of the class of systems. In the context of our example it is necessary to formalize the operational requirements for PCM, for which can serve properties defined in statics and in the dynamics, for example degradation, such as durability and biodegradability. The latter is a time challenge. Characteristics must be expressed quantitatively and / or qualitatively. For further actions, we proceed in the same way as [31], namely, let us set a hypothetical fire retardant s_0, possessing such characteristics that fully meet the operational requirements (Table 2, fire retardant 0). At the same time, the characteristic q_{0j} reflects the required value, and δ_{0j} is the tolerance value, within which the value of the corresponding characteristic satisfies the requirements.

Table 3: Values of the Elements of the Matrix $U(x)$ Calculated from the Table 2

No. compound, i	q_{i1}/δ_{i1}	q_{i2}/δ_{i2}	q_{i3}/δ_{i3}
0 (Reference, q/δ)	*1.0/8*	*1.5/8.5*	*21/10*
1	*19/1*	*12.1/0.1*	*9/1*
2	*18/1*	*10.7/0.1*	*9/1*
3	*18/1*	*3.7/0.1*	*14/1*
4	*8/2*	*8/2*	*15/1*
5	*18/1*	*2/0.1*	*9/1*
6	*3/1*	*3/0.1*	*10.6/0.1*
7	*1.1/0.1*	*2.67/0.01*	*15/1*
8	*1.1/0.1*	*1.75/0.25*	*10/1*

3.2. Computational experiment

To evaluate systems from a given class, we use the method described in [25], which consists of the following.

Let $S = \{s_i | i = 1, \ldots, n\}$ is a set consisting of n objects. For each object s_i, m characteristics $Q_{ij}, j = 1, \ldots, m$ are defined. Depending on the conditions of use, the set of considered characteristics may expand and contract, up to the introduction of chemical, biological, ergonomic and other properties. We select the reference object s_0 and its characteristics Q_{0j}, and s_0 can be either an element of the set S or not, it depends on the specific conditions of the problem and the preferences of the research or expert opinion. The characteristics of Q_{0j} are selected in such a way that s_0 is optimal in terms of the requirements for the quality of the object.

It is required to rank the elements of the set S according to the given m characteristics for compliance with the standard.

For each characteristic Q_{ij} we construct a FS, $\hat{Q}_{ij}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. To do this, we define the variables x_j with the range of values of G_j. Let the values of the variable x_j, satisfying the characteristic Q_{ij}, make up the segment X_{ij}. Then

$$X_{ij} = A(Q_{ij}).$$

We select the membership function μ_{ij} for the criterion Q_{ij}. Based on the construction of the set X_{ij} we find that in the q_{ij}, function has a maximum point, within the set X_{ij} the membership function takes values greater than 0.5, and outside X_{ij} it is less than:

$$\mu_{ij}: G_j \to [0; 1],$$

$$\mu_{ij}(q_{ij}) = 1,$$

$$\mu_{ij}(x_j) \geq 0.5 \Leftrightarrow x_j \in X_{ij}.$$

As a membership function, you can, for example, use the following functions:

$$\mu_{ij}(x_j) = \exp\left(-\frac{\ln 2}{\delta_{ij}^2}(x_j - q_{ij})^2\right), i = 0, \ldots, n, j = 1, \ldots, m.$$

As a result, we get FSs

$$\hat{Q}_{ij} = \{x_j | \mu_{ij}(x_j)\}, i = 0, \ldots, n, j = 1, \ldots, m.$$

We'll calculate the degree of equality ν_{ij} of the corresponding FS [26] to determine to what extent the characteristic of the object si is close to the characteristic of the reference object s_0:

$$\nu_{ij} = \max_{G_j} \min(\mu_{ij}(x_j), \mu_{0j}(x_j)), —$$

and the value 1 will correspond to absolute equality, and 0 to absolute inequality.

Further using the aggregate function (as such, we can take, for example, weighted voting), we'll obtain an integral estimate ν_i of the correspondence of the set of object characteristics s_i of the set of object characteristics s_0.

We'll calculate the degree of equality ν_{ij} of the corresponding FSs [32], using this method in our experiment, in order to determine to what extent the characteristics of flame retardant no. i is close to the characteristic reference no. 0:

$$\nu_{ij} = \max_{\mathbb{R}} \min(u_{ij}(x_j), u_{0j}(x_j)).$$

For the membership functions of the reduced form [33], we have obtained that the maximin value is reached at the point:

$$\nu_{ij} = u_{0j}(x_{ij}^*),$$

where $x_{ij}^* = \left(q_{ij}\delta_{0j} + q_{0j}\delta_{ij}\right)/\left(\delta_{0j} + \delta_{ij}\right)$, $i = 1, \ldots, n, j = 1, \ldots, m$.

For example,

$$x_{11}^* = \left(q_{11}\delta_{01} + q_{01}\delta_{11}\right)/\left(\delta_{01} + \delta_{11}\right) =$$

$$= \left(19 \cdot 8 + 1 \cdot 1\right)/\left(8 + 1\right) = 17,$$

$$\nu_{11} = u_{01}(x_{11}^*) = \exp\left(-\frac{\ln 2}{\delta_{01}^2}(x_{11}^* - q_{01})^2\right) =$$

$$= \exp\left(-\frac{\ln 2}{14^2}(16.867 - 1)^2\right) \approx 0.0625,$$

similarly, $\nu_{12} \approx 0.349$, $\nu_{13} \approx 0.438$.

Having conducted a weighted vote, we'll obtain an integral assessment ν_i of the correspondence of the set of object characteristics s_i to the set of object characteristics s_0:

$$\nu_i = \sum_{j=1}^{m} \alpha_j \nu_{ij},$$

where $\alpha_j \geq 0$, $\sum_{j=1}^{m} \alpha_j = 1$. The weights α_j are selected in accordance with the level of importance of each of the characteristics and are determined by the expert in the subject area so that the PC fully meets consumer requirements, technological process, etc. It is accepted here $\alpha_1 = \alpha_2 = 0.33$, $\alpha_3 = = 0.34$. The results of calculations based on the data Table 2 are presented in Table 4.

Table 4: Evaluation of Fire Retardants to Compliance with Operational Requirements

Fire retardants no., *i*	ν_i
1	*0.28*
2	*0.33*
3	*0.60*
4	*0.76*
5	*0.51*
6	*0.81*
7	*0.93*
8	*0.83*

Thus, the no. 7 fire retardant is preferable from the point of view of the presented requirements.

The proposed computational model can form the basis for automated data analysis when choosing the optimal CTS, including with the participation of experts. The approximate route of operation of such an IS is shown in Figure 2. Here, the requirements for the evaluated objects (PCM) are set directly in the IS interface when formulating the initial task, allowing specialists to quickly adjust its formulation in the context of the research.

When generating generic CTS evaluation tasks, it is possible based on known experimental data to form a library (database) of expert assessments (along with the data model library), which will relieve the user of the IS from having to obtain or collect expert assessments and fully automate the evaluation process of a specific CTS.

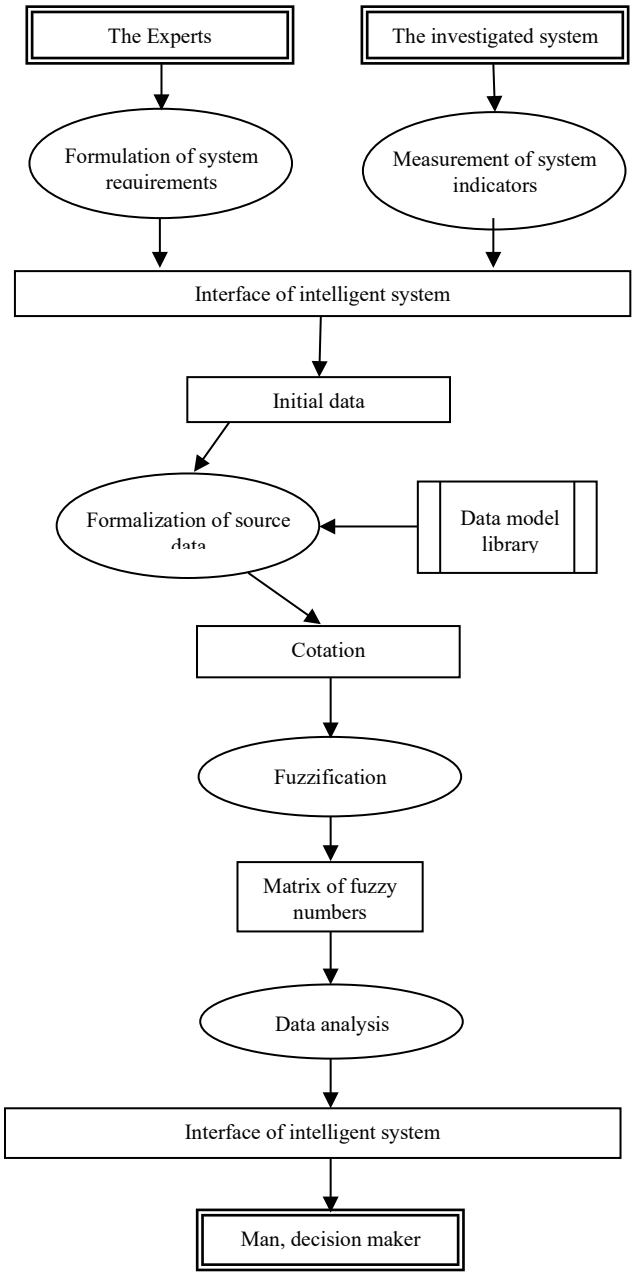


Figure 2: Scheme of the Information Flow in the IS

4. The optimal choice of a single additive for the polymer composition

4.1. Method Description

Let p_k, $k = 1, \ldots, l$ — be polymeric matrices, s_{ik}, $i = 1, \ldots, n$ — the additives for p_k. As mentioned above, by Q_{0j} we denote the value of the j-th property, which the PC should possess.

After modifying the polymer p_k by adding s_{ik} its properties become equal Q_{ikj}, $i = 1, \ldots, n$, $k = 1, \ldots, l$, $j = 1, \ldots, m$.

Further we get fuzzy numbers

$$u_{ikj} \text{ based on } X_{ikj} = A(Q_{ikj}),$$
$$u_{0j} \text{ based on } X_{0j} = A(Q_{0j}), \quad (1)$$

where $i = 1, \ldots, n$, $k = 1, \ldots, l$, $j = 1, \ldots, m$.

To select one additive and one polymer matrix we calculate

$$\nu_{ikj} = u_{0j}(x^*_{ikj}), \quad (2)$$

where $x^*_{ikj} = \left(q_{ikj}\delta_{0j} + q_{0j}\delta_{ikj}\right) / \left(\delta_{0j} + \delta_{ikj}\right)$, $i = 1, \ldots, n$, $k = 1, \ldots, l$, $j = 1, \ldots, m$ and

$$\nu_{ik} = \sum_{j=1}^{m} \alpha_j \nu_{ikj} . \quad (3)$$

In the end, choose the additive $s_{i_0 k_0}$ and polymer matrix p_{k_0}, where $(i_0, k_0) = \arg \max_{\substack{(i,k),\\ i=1,\ldots,n,\\ k=1,\ldots,l}} \nu_{ik}$.

4.2. Computational experiment

We present a study of PC with alternative polymer matrices of polyurethane elastomer and polycaproamide and use a set of fire retardants presented in Table 5, for an extended demonstration of the capabilities of the fuzzy data analysis method. The number of flame retardants for simplicity of notation is the same as in the Table 2, but this is not essential for the method and, in general, the amount of additives for each polymer matrix can be different. In addition, the reference value, that is, additive number 0, is left the same as in Table 2. This is essential for the considered mathematical method of PC analysis.

Table 5: Background Characteristics of Fire Retardants for Elastomer Matrix [34]

Compound no., i	Q_{i1}	Q_{i2}	Q_{i3}
0 (Reference, q/δ)	*1.0/8*	*1.5/8.5*	*18/10*
1	*12.7*	*17.1*	*20*
2	*8.5*	*8.5*	*18*
3	*20.6*	*20.7*	*20*
4	*18.7*	*18.7*	*17*
5	*6.2*	*6.2*	*16*
6	*18.0*	*22.1*	*17*
7	*28.1*	*27.0*	*20*
8	*4.4*	*4.4*	*10*

Let p_1 be polycaproamide, and p_2 — polyurethane elastomer, data on additives to which are given in Table 2 and Table 5 respectively. After the fuzzification of the initial data using (1), we obtain the fuzzy numbers u_{i1j} (Table 3) and u_{i2j} (Table 6) for p_1 and p_2, respectively.

Table 6: Parameters of Elements of the Matrix $U(x)$ Calculated on the Data of Table 5

Compound no., i	q_{i1}/δ_{i1}	q_{i2}/δ_{i2}	q_{i3}/δ_{i3}
0 (Reference, q/δ)	*1.0/8*	*1.5/8.5*	*18/10*
1	*12.7/0.1*	*17.1/0.1*	*20/1*
2	*8.5/0.1*	*8.5/0.1*	*18/1*
3	*20.6/0.1*	*20.7/0.1*	*20/1*
4	*18.7/0.1*	*18.7/0.1*	*17/1*
5	*6.2/0.1*	*6.2/0.1*	*16/1*
6	*18.0/0.1*	*22.1/0.1*	*17/1*
7	*28.1/0.1*	*27.0/0.1*	*20/1*
8	*4.4/0.1*	*4.4/0.1*	*10/1*

Applying to the data Table 3 ($k = 1$) and Table 6 ($k = 2$) formula (2) we obtain the values ν_{ikj}, $i = 1, \ldots, 8$, $k = 1, 2$, $j = 1, 2, 3$. Next, using (3), we obtain ν_{ik}, $i = 1, \ldots, 8$, $k = 1, 2$, are given in Table 4 for $k = 1$ and in Table 7 for $k = 2$.

Table 7: Evaluation of Elastomer Fire Retardants in Compliance with Operational Requirements

Fire retardants no., i	ν_{i2}
1	*0.45*
2	*0.71*
3	*0.35*
4	*0.34*
5	*0.81*
6	*0.33*
7	*0.34*
8	*0.77*

Among the data Table 4 and Table 7 find the maximum value of *0.93*, which corresponds to the addition of s_{71} and the polymer matrix p_1.

Thus, a specific recommendation was received on the flame retardant, which is part of the PC. Extensive work is needed to formulate a recipe based on these recommendations. This, for example, the definition of other components or their part in the PC, which allows you to take the first step in developing a new PCM in an automated way. Here we note that the proposed analytical method also has the possibility of choosing additional components for PC design. This is covered below.

Above (see Table 1) we have already indicated that thanks to the growth and expansion of consumer requirements, the complexity of polymer systems are increasing simbatically. Therefore, identification tasks are also becoming more complex.

Consider the complicated task of choosing several ingredients (additives) for PC design.

5. Selection of several additives for PC

5.1. Method description

Let it be necessary to choose one polymer matrix from the set p_k, $k = 1, \ldots, l$ and $r \le n$ additives for it from s_{ik}, $i = 1, \ldots, n$.

To do this, we define r properties $Q_{j_1}, \ldots, Q_{j_r}$, to improve which the choice of r additives is directed.

For each polymer matrix p_k, we define a set of additives $s_{i_s k}$ ($i_s \in \{1, \ldots, n\}$, $s = 1, \ldots, r$, $k = 1, \ldots, l$) such that the additive $s_{i_s k}$ provides the best property value Q_{j_s} among the additives $s_{1k}, \ldots, s_{nk}$. However, some numbers may coincide with each other if the same additive provides the best values for several properties at once. In this case, for the modification, not r, but a correspondingly smaller amount of additives is used, but the result will correspond to the set task — the same r properties are improved.

Next, we obtain ν_{ik}, $i = 1, \ldots, n$, $k = 1, \ldots, l$ as it was described in Section 4 and define

$$\nu_k = \sum_{s=1}^{r} \alpha_{j_s} \nu_{i_s k} . \quad (4)$$

Determine the number

$$k_0 = \arg \max_{k=1,\ldots,l} \nu_k$$

and choose a modification of the polymer matrix p_{k_0} a set of additives $s_{i_1 k_0}, \ldots, s_{i_r k_0}$.

5.2. Computational experiment

To demonstrate the capabilities of the method presented here, we will use the same data (Table 2 and 5) and the same notation as above. Since the formulas for calculating the same, the values of ν_{ik} are also given in Table 4 and 7.

Table 8: Values of the Degree of Compliance of the Characteristics of Additives to Consumer Requirements to PCM

Compound no., i	$k = 1$			$k = 2$		
	ν_{i11}	ν_{i12}	ν_{i13}	ν_{i21}	ν_{i22}	ν_{i23}
1	*0.0625*	*0.349*	*0.438*	*0.235*	*0.102*	*0.994*
2	*0.0843*	*0.452*	*0.438*	*0.552*	*0,.632*	*0.950*
3	*0.0843*	*0.956*	*0.755*	*0.0173*	*0.0316*	*0.994*
4	*0.712*	*0.767*	*0.814*	*0.0365*	*0.0625*	*0.912*
5	*0.0843*	*0.998*	*0.438*	*0.752*	*0.813*	*0.867*
6	*0.966*	*0.979*	*0.480*	*0.0472*	*0.0187*	*0.912*
7	*1.0*	*0.987*	*0.814*	*0.000427*	*0.00226*	*0.994*
8	*1.0*	*0.999*	*0.5*	*0.885*	*0.924*	*0.5*

Let $r = 2$, and $j_1 = 1$, $j_2 = 2$. To determine i_1 and i_2, ν_{ikj} values are needed that were not of immediate interest previously and were omitted from the examples of calculations. Now we bring these data to Table 8

To find i_1 for $k = 1$, we find the largest among the ν_{i11} values. This is *1*. This value is achieved for two supplements. We take $i_1 = 7$. Similarly, we get $i_2 = 8$. Acting in the same way, for $k = 2$ we define $i_1 = 8$, $i_2 = 8$.

Next, we calculate by (4)

$$\nu_1 = \alpha_1 \nu_{71} + \alpha_2 \nu_{81} = 0.33 \cdot 0.93 + 0.33 \cdot 0.83 \approx 0.58,$$

$$\nu_2 = \alpha_1 \nu_{82} + \alpha_2 \nu_{82} = 0.33 \cdot 0.77 + 0.33 \cdot 0.77 \approx 0.51.$$

The value $\nu_1 \approx 0.58$ indicates that for $k = 1$, the PC has a higher degree of compliance with the requirements. As a result, we choose a PC based on textile fiber ($k = 1$) with additives no. 7 and 8 (Table 2). Thus, a specific recommendation on the main additional component of the PC is also received here. The formation of a complete recipe based on these recommendations has the peculiarity that the proposed composition is already a multi-component system, but still has an incomplete character and requires more design and experimental work to implement the full "task-practice" algorithm. This, as well as noted above, can be the definition of additional components, their concentration in PC, technological and operating conditions. Calculations should be considered as a quantitative justification for the decision to develop a new PC. With the development of these approaches it is possible to create an automated system.

6. Conclusion

Summarizing the above, we note that often the CTS design problem arising at the initial stages is the task of choosing the best or, in a certain sense, optimal variant from the set of admissible ones. They must meet the specified requirements on the "composition – property – application" platform. When you try to get a solution in terms of incomplete and fuzzy information about the object of design, there is a high probability of incorrect or inaccurate multiple solutions. Therefore, the use of a clear strategy can minimize this probability, also reducing the cost of design and practical implementation of a technical solution.

The approach proposed in this article to mathematical modeling using fuzzy information in the field of chemical technology is essentially the mathematical support of a profiled IS. The formation of a database of models within the system allows us to bring the linguistic resources of the user interface to a new level of interaction using language environments more similar to natural languages and setting the input data in a form that is quite familiar to the researcher.

We should also note that only one of the most important aspects of using fuzzy simulation of PCM is shown here. It is aimed at using information in the form of fuzzy numbers when choosing PC components from a uniform set. As can be seen from the above results, the analysis of experimental data is well algorithmized and can be implemented as an IS of a chemist-researcher and a chemist-technologist. The above mathematical calculations are illustrated by the example of PCM design in general methodical terms, so they can be transferred to other natural and engineering systems in which multicomponentity is the main feature.

The development of the proposed computational methods creates theoretical foundations for the formalization and automation of the process of quality management of technical systems. In this case, it is possible to attract highly qualified experts only once to customize methods for solving specific tasks. The necessary step in this case is the formation of a library of settings and the solution of basic ("zero") tasks. This will further allow you to work offline.

At the same time, the expert has several roles in stock:

- to identify the links between informal entities of the subject domain and formal entities of the theory of FS;
- the formation of an optimal set of requirements that a quality product must meet;
- applied to the PCM — selection of the PC composition according to the selected requirements.

This information will further allow to formalize the initial data of the subject area in terms of the theory of FS and use them in calculations, when upon receipt of an order from the consumer a set of requirements for PCM is automatically generated, and then estimates of formed requirements for specific candidates for PC components are calculated. Then a decision is made on the technical implementation of the composition, which ensures high quality of PCM. The analysis given here is given as a possible prospect for the development of the applied use of fuzzy mathematics tools in engineering. The stated essence of the method shows its role and value in solving the problem of quality management of PCM and other technical systems.

At the same time, we selectively point out the nearest additional applied work-analogues and areas in which methods for analyzing fuzzy information are being developed. These are economics [35], management systems [27], analysis of leasing projects [36], traffic optimization [37], occupational studies [38]. As a prospect, it can be viewed as technical compositions of medicinal and cosmetic forms, building and drilling mixtures and solutions, alloys, and some colloidal systems. As a nontrivial and nontechnical one, it is possible to propose to consider personnel recruitment tasks, for example, for diplomatic or other complex, for example, artistic work, from a certain number of applicants with a quantitative assessment of personal data and interview results. The basis for this is [39, 40]. True, the fully successful outcome of the future work of candidates cannot be guaranteed — one thing is a technique, the other is people-specialists. The risk factor, as in polymer and other systems, remains, but it is much higher.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The reported research was funded by Russian Foundation for Basic Research and the government of Volgograd region, grant № 18-48-340011 "Quality management of polymer products based on the optimization of formulation of composite materials"

References

[1] I. V. Germashev, T. P. Mashihina, V. E. Derbisher, "Analysis of Fuzzy Data for Polymer Production Quality Management" in 2018 International Russian Automation Conference, RusAutoCon, Sochi Russia, 2018. https://doi.org/10.1109/RUSAUTOCON.2018.8501653

[2] J.-o. Kim, C. W. Mueller, Factor Analysis: Statistical Methods and Practical Issues, SAGE Publications, 1978. https://dx.doi.org/10.4135/9781412984256.

[3] Y. S. Vygodsky, T. V. Volkova, O. N. Zabegaeva, Z. Y. Chistyakova, V. A. Shanditsev, M. I. Buzin, Y. V. Zubavichus, O.V. Sinitsyna, G. G. Nikiforova, A. P. Krasnov et al., "Synthesis and characteristics of the composites based on poly(caproamide) and multiwalled carbon nanotubes" Polym. Sci. Ser. C+, 51(1), 63–73, 2009. https://doi.org/10.1134/S1811238209010111

[4] P. D. Shestakov, A. V. Galtsev, A. S. Tarasov, "Hardening Polymer Composites with Carbon Nanotubes" International Scientific Research Journal, 6(6), 79–81, 2012.

[5] M. G. Svistunov, V. E. Golovin, K. K. Koshelev, G. A. Kosheleva, Y. N. Nesterov, O. K. Barashkov, E`lektroizolyatsionnaya polimernaya kompozitsiya, Patent RU 2636383, 2017.

[6] W. R. Klecka, Discriminant Analysis, SAGE Publications, 1982. https://dx.doi.org/10.4135/9781412983938.

[7] M. S. Aldenderfer, R. K. Blashfield, Cluster Analysis, SAGE Publications, 1984. https://dx.doi.org/10.4135/9781412983648.

[8] A. Faure, Perception et Reconnaissance des Formes, Nombreux schémas N&B, 1985.

[9] A. G. Blokhnin, "Fuzzy-set algebra" J. Comput. Sys. Sc. Int+, 37(5), 747–754, 1998.

[10] M. G. Voskoglou, I. Y. Subbotin, "Application of the triangular fuzzy model to assessment of analogical reasoning skills" American Journal of Applied Mathematics and Statistics, 3(1), 1–6, 2015. https://doi.org/10.12691/ajams-3-1-1.

[11] D. Silva, L. F. V. Pinto, D. Bozukova, L. F. Santos, A. P. Serro, B. Saramago, "Chitosan/alginate based multilayers to control drug release from ophthalmic lens" Colloid Surface B, 147, 81–89, 2016. https://doi.org/10.1016/j.colsurfb.2016.07.047.

[12] T. Ikeda, H. Tamura, T. Sakurai, S. Seki, "Control of optical and electrical properties of nanosheets by the chemical structure of the turning point in a foldable polymer" Nanoscale, 8(30), 14673–14681, 2016. https://doi.org/10.1039/c6nr01066k.

[13] A. Kobayashi, T. Nakaza, T. Hirano, S. Kitagawa, H. Ohtani, "Variation in the chromatographic, material, and chemical characteristics of methacrylate-based polymer monoliths during photoinitiated low-temperature polymerization" J. Sep. Sci., 39(13), 2459–2465, 2016. https://doi.org/10.1002/jssc.201600344.

[14] B. Strachota, J. Hodan, L. Matejka, "Poly(N-isopropylacrylamide)-clay hydrogels: Control of mechanical properties and structure by the initiating conditions of polymerization" Eur. Polym. J., 77, 1–15, 2016. https://doi.org/10.1016/j.eurpolymj.2016.02.011.

[15] L. Y. T. Chou, F. Song, W. C. W. Chan, "Engineering the structure and properties of DNA-nanoparticle superstructures using polyvalent counterions" J. Am. Chem. Soc., 138(13), 4565–4572, 2016. https://doi.org/10.1021/jacs.6b00751.

[16] I. V. Germashev, V. E. Derbisher, T. M. Kokorina, I. A. Murashkina, "Assessing the quality of knitted fabrics using the theory of inexact sets" Izvestiya Vysshikh Uchebnykh Zavedenii, Seriya Teknologiya Tekstil'noi Promyshlennosti, (4), 7–10, 1999 (article in Russian with an abstract in English).

[17] V. E. Derbisher, I. V. Germashev, G. G. Bodrova, "Fuzzy-set-based quantitative estimates of the efficiency of thermo- and photostabilizing additives in polymeric compositions" Polym. Sci. Ser. A+, 39(6), 630–633, 1997.

[18] B. A. Khidhir, W. Al-Oqaiel, P. M. Kareem, "Prediction models by response surface methodology for turning operation" American Journal of Modeling and Optimization, 3(1), 1–6, 2015. https://doi.org/10.12691/ajmo-3-1-1.

[19] M. R. S. Emami, "Fuzzy logic applications in chemical processes" The J. of Math. and Comput. Sci., 1(4), 339–348, 2010. https://doi.org/10.22436/jmcs.001.04.11.

[20] S. Krishnamoorthy, "Linguistic features for review helpfulness prediction" Expert Syst. Appl., 42(7), 3751–3759, 2015. https://doi.org/10.1016/j.eswa.2014.12.044.

[21] K. Nasrollahzadeh, M. M. Basiri, "Prediction of shear strength of FRP reinforced concrete beams using fuzzy inference system" Expert Syst. Appl., 41(4), 1006–1020, 2014. https://doi.org/10.1016/j.eswa.2013.07.045.

[22] A. Salski, F. Bartels, "A fuzzy approach to land evaluation" IASME Trans., 5(2), 774–780, 2005.

[23] M. I. Kizhisseri, M. M. A. Mohamed, "Fuzzy-based wastewater quality indices for pollution classification: a case study in the United Arab Emirates" Environment Systems and Decisions, 36(1), 62–71, 2016. https://doi.org/10.1007/s10669-015-9579-9.

[24] P. H. Chandra, S. M. S. T. Kalavathy, A. M. I. Jayaseeli, J. P. Karoline, "Mechanism of fuzzy ARMS on chemical reaction" in V. Snášel, A. Abraham, P. Krömer, M. Pant, A. Muda (eds) Innovations in Bio-Inspired Computing

and Applications. Advances in Intelligent Systems and Computing, 424, 43–53, 2016. https://dx.doi.org/10.1007/978-3-319-28031-8_4.
[25] I. V. Germashev, V. E. Derbisher, T. F. Morozenko, S. A. Orlova, "Otsenka kachestva tehnicheskih ob``ektov s ispol`zovaniem nechyotkih mnozhestv" Industrial laboratory. Diagnostics of materials, 67(1), 65–68, 2001.
[26] I. V. Germashev, V. E. Derbisher, "Fuzzy Optimization of Polymer Compositions" Theor. Found. Chem. Eng., 35(4), 418–421, 2001.
[27] A. A. Uskov, I. A. Kiselev, "Comparing fuzzy arithmetic with quaternion arithmetic and applying the former to control system analysis" Large-scale Systems Control, (48), 59–70, 2014 (article in Russian with an abstract in English).
[28] A. I. Orlov, E. V. Lutsenko, Sistemnaya Nechyotkaya Interval`naya Matematika, KubGAU, 2014.
[29] A. I. Orlov, Nechislovaya Statistika, MZ-Press, 2004.
[30] E. Egrioglu, C. H. Aladag, M. A. Basaran, U. Yolcu, V. R. Uslu, "A new approach based on the optimization of the length of intervals in fuzzy time series" J. Intell. Fuzzy Syst., 22(1), 15–19, 2011.
[31] I. V. Germashev, V. E. Derbisher, E. V. Derbisher, E. A. Markushevskaya, "The Fuzzy Data Analysis for the Chemical-Technological Systems Optimization" Information Technologies, 24(3), 153–160, 2018 (article in Russian with an abstract in English).
[32] V. P. Meshalkin, E`kspertny`e Sistemy` v Himicheskoy Tehnologii, Chemistry, 1995.
[33] I. V. Germashev, V. E. Derbisher, "Properties of unimodal membership functions in operations with fuzzy sets" Russian Mathematics, 51(3), 72–75, 2007. https://doi.org/10.3103/S1066369X07030115.
[34] I. V. Germashev, V. E. Derbisher, S. A. Orlova, "Evaluation of activity of the fireproofing compounds in elastomer compositions by means of fuzzy sets" Kauchuk i Rezina, (6), 15–17, 2001 (article in Russian with an abstract in English).
[35] Y. I. Davydov, "Primenenie teorii nechyotkih mnozhestv k e`konomicheskim zadacham" News of St. Petersburg State Agrarian University, (26), 261–266, 2012.
[36] Y. I. Davydov, "Teoriya nechyotkih mnozhestv v otsenke e`konomicheskoy e`ffektivnosti lizingovy`h proektov v usloviyax neopredelennosti i riska" News of the International Academy of Agrarian Education, (14-2), 32–39, 2012.
[37] V. A. Romanenko, "Fuzzy optimizing of transfer air transportation system parameters" Bulletin of the Samara State Technical University, Series: Technical Sciences, (4), 73–80, 2012 (article in Russian with an abstract in English).
[38] G. A. Kryzhanovsky, V. V. Kupin, "Analiz vazhnosti professional`ny`h zadach na osnove nechetkih chislovy`h e`kspertny`h otsenok" Bulletin of St. Petersburg State University of Civil Aviation, (2), 43–50, 2012.
[39] S. V. Skorokhod, "Otbor personala v usloviyah nechyotkoy informatsii na osnove e`kspertny`h otsenok" News of SFU. Engineering Science, (9), 125–130, 2008.
[40] S. V. Skorokhod, "Primenenie nechyotkih chisel dlya otsenki kvalifikatsii personala" News of TRTU, (3), 214–216, 2005.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 100-105 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Multimedia Application Development With Islamic Critical Reflection Through 3-2-1 Technique for Novice Teacher Internship Program

Evi Fatimatur Rusydiyah[1,*], Rakhmawati Rakhmawati[1], Eni Purwati[1], Moh. Hafiyusholeh[2], Ahmad Hanif Asyhar[2]

[1]*UIN Sunan Ampel Surabaya, Faculty of Education and Teacher Training, Indonesia*

[2]*UIN Sunan Ampel Surabaya, Mathematics Department, Indonesia*

ARTICLE INFO

Article history:
Received: 20 April, 2019
Accepted: 30 April, 2019
Online: 24 May, 2019

Keywords:
Macromedia Flash
Islamic Critical Reflection
Research and Development

ABSTRACT

The use of technology in the learning process has become the most significant point in helping teachers reaching the objectives of the learning. This study focuses on the use of multimedia by novice teachers in comprehending the internship program's guidance. The Islamic Critical Reflection model is a new and innovative model in the field of education which development is a result of integration between Islamic value and science. This study aims to answer problems related to the internship supervision which occurs among novice teachers. The multimedia application in the form of Flash Macromedia is developed using Islamic Critical Reflection model for internship program of novice teachers at the Tarbiyah and Teacher Training Faculty UIN Sunan Ampel Surabaya. The research design of this study is Research and Development. It includes the process of collecting information, designing a product, validating the design, improving the design, piloting the product, re-revising the product, trying out the product, revising further and creating products for large numbers. The results of this study indicate that the products produced by Macromedia Flash application integrated with Islamic Critical Reflection Model for Tarbiyah and Teacher Training Faculty of UIN Sunan Ampel Surabaya's internship program are feasible to use. The results are from the experts' analysis and users' trials. Experts involved in the validation of this product are Islamic Education specialists, application experts, educational technology experts, and curriculum expert.

1. Introduction

UIN Sunan Ampel Surabaya has a vision and mission that is hoped to be reached through the milestones between the period of 2019-2045. The milestones are described in the UIN Sunan Ampel Surabaya's Development Master Plan 2019-2045 [1], [2]. One of the biggest duty of UIN Sunan Ampel Surabaya and also of other higher education institutions managed by the Ministry of Religion is to integrate knowledge and religion as the differentiator between the institutions with those managed by the Ministry of Education.

Higher education nowadays has more challenging responsibility in preparing its students to be ready for the 4.0 industrial revolution. This preparation can be managed by an internship program for novice teachers. It is because this module is an integration between theory and practice done at schools. This program aims to prepare students to be professional teachers.

* Evi Fatimatur Rusydiyah, Jl. A. Yani 117 Surabaya, +62 8121635409 & evifatimatur@uinsby.ac.id

Within this context, UIN Sunan Ampel has done many adaptation programs. One of the programs is designed to train novice teachers in mastering teaching skill that is intact and integrated. Therefore, after graduating, novice teachers who have taken the internship module are hoped to be ready for teaching [3], having a positive identity as a teacher, and acquiring more knowledge in the field of teaching [3]. It is possible because the internship program has several elements that are involved: supervisor, mentor teacher, and novice teacher.

The challenge is that the internship program that is running at the moment is considered to be not effective and not efficient yet in its guidance system. The problem that is emerged is the unstructured guidance program within each element. The evaluation and reflection in the internship program have only been partially conducted without a thorough involvement of various weakness points. The result is, there is no comprehensive information toward the development of novice teachers during the internship program. This condition then resulted in less competent novice teachers. Whereas, this internship program is designed to

be the prime program the faculty can offer, and also designed to be the educational core program.

To resolve the problem, Tarbiyah and Teacher Training faculty strives to develop a supervision system through critical reflection approach in the internship program. The technique implemented is the 3-2-1 technique with Islamic values as the base of its integration. Therefore, the process will facilitate teaching and learning, and it would be central to the teacher's development. The implementation of the 3-2-1 technique is done within the discussion process between the supervisor, mentor teacher and novice teacher. The three elements provide three positive inputs of what was found during the internship. After that, two negative inputs are presented and followed by a solution [4]. The Islamic Critical Reflection approach is used during supervision. It is based on Surah Al-Hujurat verse 6-8.

In this research, the concept is made in the form of the application using Flash Macromedia. Included in it, is the first supervision until the eighth. In other words, this research aims to develop a multimedia application with Islamic Critical Reflection for the Novice Teacher Internship Program.

In this research, multimedia as a tutorial application in facilitating teacher mentors, lecturers, and students in carrying out practical activities in field experience, so that they can understand the steps of supervision they will do for each. For teacher mentors, they will understand what they must provide in each supervision. For lecturers, they can understand what must be done when they are supervised. Likewise, for students, they will understand what they should get every time the lecturer supervises. Therefore this multimedia is not used for the learning process in the classroom, but this multimedia is used as a tutorial application in making supervising steps clear.

With the application, it is hoped that the novice teacher's professionalism is improving in the way that novice teachers can critically reflect and integrate it with Islamic values and pedagogical knowledge. "Critical reflection involves concern about the broader issue of teaching and learning*"* [5] is the notion that Murray and Kujundiz hold. They explained that [6] "critical reflection involves a process of analyzing, reconsidering, and questioning experiences within a broad context of issues such as analyzing assumption about teaching, raising awareness about the context of teaching. They imagine alternative ways of thinking".

The concept is visualized in the form of flash Macromedia. The application is developed through the Technology Acceptance Model (TAM). TAM model modification is done by Venkantesh [7] by adding trust variable with the title: Trust enhances Technology Acceptable Model which evaluating the relationship between TAM variable and trust. TAM other modification is Trust and Risk in Technology Acceptance Model (TRITAM) that uses trust and risk variable named TAM variable [8]. This model is one of the models developed from the theory of using a system of technology. It is considered to be the most affected in explaining personal acceptance towards the benefit and information technology used as a system [9].

One research that is related to the novice teacher program is Toman et al., [10]. It develops reflective thinking method and determines the novice teacher reflective thinking level. In 2007, Mathew discussed how mentor teacher creates opportunities for the novice teachers to developed reflective practice during the internship at school [11]. Gonen [12] in the research that was aimed to evaluate the experience that caused some changes after reflection is done. Moreover, it was also aimed to figure out what experience gained by the novice teacher after receiving guidance from the mentor teacher [12].

The importance of this research is for both the mentor teacher and the novice teacher understand the supervision process thoroughly, especially in implementing the 3-2-1 reflective principal as a part of novice teacher professional development. Novice teacher will be more confident with the guidance process that used Islamic Critical Reflection. Mentor teacher will also work professionally in guiding novice teacher.

2. Design

This research is aimed to develop an application in the form of Macromedia flash to be used in the internship program. This multimedia application is used to make an understanding of the duties of lecturers, mentor teacher,, and novice teacher when conducting internship program as a tutorial in supervising. Therefore this application is not used in class as media teaching-learning but is used in the supervision process. According to Borg and Gall [13], this research is categorized as Research and Development (R&D) research.

Steps that are generated from the process are understood to be the R&D cycle. They are consisted of studying research findings, developing the product, testing the product, and revising it to improve its weaknesses.

The R & D design used in this study is as follows.

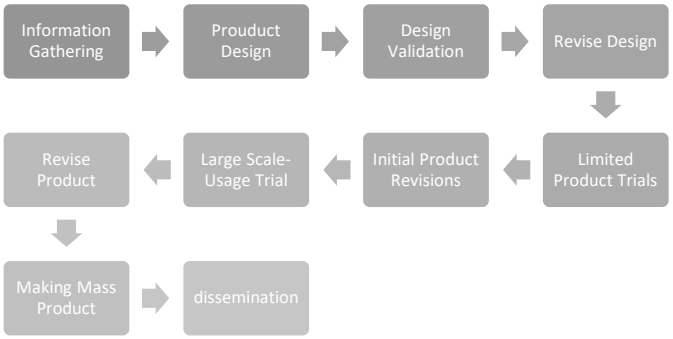


Figure 1. R & D flow

The R & D steps as referred to in the figure are information gathering, product design, design validation, revise design, limited product trials, initial product revisions, large-scale usage trials, revise product, making mass product, and dissemination [13]. These steps are a guide for researchers to conduct this research.

The test subjects in this R & D are stakeholders involved in implementing the practice of the field experience. The

stakeholders are 50 partner schools of the Tarbiyah Faculty UIN Sunan Ampel Surabaya which are occupied by field experience, 50 supervisors, and 150 mentor teachers.

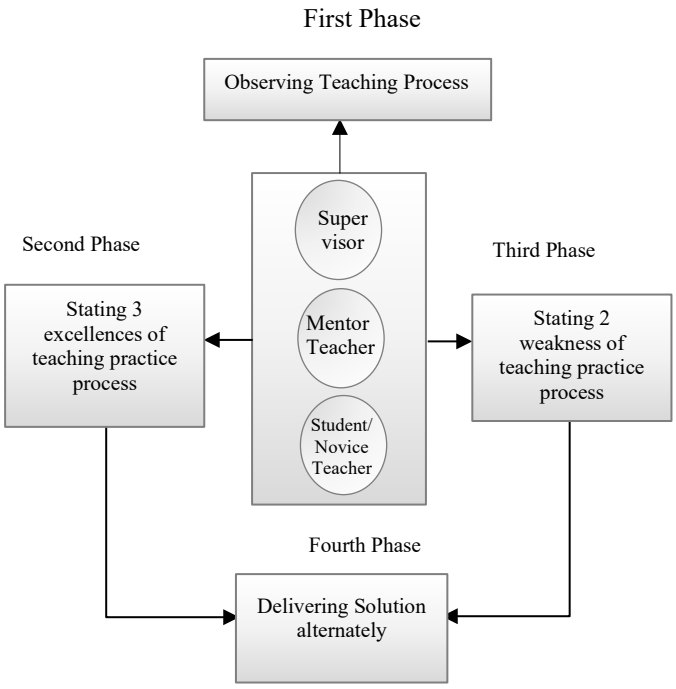


Figure 2: Islamic Critical Reflection Model

The product produced is based on the Islamic Critical Reflection model. This is a distinction as UIN Sunan Ampel integrated its learning process to Islamic principles and research [14], [10], [12]. The steps in Islamic Critical Reflection are observation, advantages, deficiency, and solution. The scheme can be seen in Figure 2.

The instrument which is used to measure the product is the Technology Acceptance Model (TAM). It is to explain personal acceptance towards the benefit and the use of information technology as a system [9]. After the application is made, the validity test was conducted. The experts to validate were ones who are expert in Islamic Education, Technology Education, and Curriculum. Operational suitability of the product was conducted in ten different primary schools and secondary school in Surabaya and Sidoarjo, East Java. The questionnaire was the main instrument for this stage.

3. Finding and Discussion

As in the R & D design above some of the findings in this study are as follows:

3.1. Gather Information

Information about the implementation of internship program in schools was conducted on Sunday, May 13, 2018 through a Focus Group Discussion (FGD) which was attended by 47 good tutors who were in Ibtidaiyah Madrasah, Madrasah Tsanawiyah / Middle School, Islamic Senior High Schools/ Vocational Schools in the region Surabaya and Sidoarjo and 41 Supervisors at the Sunan Ampel State Islamic University (UINSA). The FGD results show the need for understanding the concept of how the guidance model can improve the quality of UINSA internship's students who are qualified, the need for easy access that can be reached by students, civil servant teachers and Field Supervisors about guidelines for implementing effective and efficient internship program.

3.2. Stage of Product Design

On June 5, a Forum Group Discussion (FGD) was held which was attended by experts, namely educational technology experts, multimedia experts, Islamic education experts, and curriculum design experts from development products including a model of student assistance with Pamong Teachers and Field Advisors with 3-2-1 (triad meeting) which is visualized in Flash Media Media multimedia. On June 10, an FGD was held which was attended by experts again by presenting users. This FGD designed a product that will be formed in the internship application development model using Islamic studies on how to provide critical assistance based on an Islamic approach.

At this stage it was found, the practice of field training needs to be done with a reflective approach in communicating to convey information. This 3-2-1 approach is an implementation of Islamic teachings, that motivating students must be done politely with techniques to promote positive aspects rather than negative aspects. Likewise, the technique 3-2-1 strengthens the process of training students in finding solutions based on an analysis of their strengths and weaknesses [6]. Therefore approach 3-2-1 is an approach that leads to the analysis of the theory of critical reflection.

The 3-2-1 approach is also supported by the pyramid theory of human needs from Malow [15]. Humans in their lives have different needs, those needs from the high to the low. The lowest needs are the need to acknowledge oneself, then the need for compassion, love and love, then the need for security, and the highest needs are physiological needs. The higher the education of a person the higher the level of recognition desired [16]

Figure 3: Preliminary Display of the Pre Service Teacher Interference Model Application through the Islamic Critical Reflection Approach

3.3. Stage of Validation Design

The validators involved in this study were 1) Islamic religious education experts, multimedia experts, learning technology

experts, and curriculum experts. The results of the validator are as follows.

Islamic Education Expert

The application was developed through several steps. The first one is the observation step based on the basic Islamic teaching, which is Surah Al Maidah verse 17 and surah Al-Anam verse 46 about observing the process as one of Allah's power. In observing the novice teacher, mentor teacher could see how a novice teacher is going through a process in becoming a professional teacher and how a novice teacher is going through a process to be better.

The second and third step is to give a positive and negative suggestion. The second and the third step is the feedback from the supervisor, mentor teacher, and novice teacher. They all work together to suggest positive points. The inspiration for this step is Surah Al Baqarah verse 219, Surah Al Haj verse 28, Surah Al Isra' verse 14 and Surah Al Ankabut verse 3. Based on the verses, it can be concluded that as a supervisor, a mentor teacher, and a novice teacher, all the three elements have to deliver the truth, which is: conveying positive points to understand the deeper meaning, to contemplate, to evaluate the positive and the negative sides. Contemplation is designed by giving three positive characters and two negative experience.

The fourth step is the solution stage. It is based on Surah Al Baqarah verse 153, Surah Al Baqarah verse 109, Surah Al Ahzab, verse 70-71. All of the surah mentioned gives a clue that supervisor, mentor teacher and novice teacher have to show positive attitude such as to talk politely.

To ease the user in operating the application, guidance and summary result were provided thoroughly as it is seen in Figure 4.

Rangkuman

Praktik Pengalaman Lapangan II (PPL) di Fakultas Tarbiyah dan Keguruan UIN Sunan Ampel Surabaya adalah suatu rangkaian kegiatan untuk menerapkan berbagai teori kependidikan yang

Figure 4: Application Guidelines And An Application Summary

The Macromedia flash application consists of application guidelines, internship material, supervision stages, stages on supervision using the 3-2-1 technique. The third is to show the positive sides of the novice teacher's teaching. The second is to show the weaknesses of the novice teacher's teaching, and the first one is the solution. The 3-2-1 technique was used in a discussion and is conducted when the supervisor supervises the novice teacher. The 3-2-1 technique will be used not only by the novice teacher but also as a supervisor.

The result of the research is suitable with Maslow concept about human need pyramid. It is said that human need recognition. The recognition needed is toward self-actualization, affection, safety, and psychological needs [15]. Those needs are basic human needs. It also supported by several research: Schulte dan Marthann [17], Medcalf et all[18], Emad [19] dan Sari dan Dwiarti [20]. These research found out that motivation has a significant effect on work and study achievement.

The result of this 3-2-1 reflection technique is hoped to shape the novice teacher's self-confident. The teaching and learning will result in positive, and novice teacher can find their professionalism through reflecting every meeting. Therefore, they will be a professional teacher.

Application Analysis

The application that has been made was measured for its suitability using TAM indicator which consists of Perceived of Usefulness with its sub-indicators as follow: *1) Make Job Easier; 2) Work More 3)* Increase Productivity *4)* Effectiveness 5) Improve job performance *(6)* Useful. The second indicator is Perceived ease of use with its sub-indicator as follow: *1)* Easy to learn *2)* Controllable *3)* Clear and understandable *(4)* Flexible*, 5)* Easy to become skillful *(6)* easy to use [21]. This research also used an application that was developed based on the human-machine interaction [22] indicator by Davis [23] and Hendrawati [24]. It has several indicators, such as *learnability, efficiency, memorability, errors, and satisfaction.*

Through the application, it is gained several positive aspects as follow: the application could give guidance for a supervisor so that the supervisor could use it easily. It could also aid the supervisor's understanding in guiding a novice teacher. It helps the supervisor to improve novice teachers' professional competencies. This application is designed to present steps in supervising internship programs using the 3-2-1 technique effectively. This, in return, beneficial for faculty, supervisors and lecturers, novice teachers, school and stakeholders to develop teachers' professionalism.

Based on the application analysis, it is found out that the application design using buttons that are familiar for users is compatible with these Operating System: Windows, Linux, and Android. This development supports Koral and Emel [25], also Akar and Güzin [26] who suggest that technology acceptance relates to content comprehension which becomes the focus. It is also stated by Konak et al. [27] who stated that technology acceptance affects collaborative work. The same notion explained by Davis [21] and Hendrawati [24] who figured that technology acceptance had affected intrinsic motivation to occur.

As the previous research, information technology system acceptance could be explained within group work. The bigger the group who receive the information technology system, the bigger its practice would change. The group would also tend to use or try something new that could be used or developed in an information technology system

3.4. Education of Technology analysis

Education technology is a knowledge which relates to theory and practice on how to design, develop, implement, organize, and measure the education process to reach the educational aim. One

of its important components is a reflection. Reflection is a stage in an educational process so that the goal is reached. Therefore, reflection is used as one of the stages in this application so that it would affect novice teacher positively.

This condition is the same as several research which stated that every novice teacher has a reflective thinking skill [10]. Novice teacher has contributed positively towards the teaching skill that is developed to think reflectively [11]. Reflective needs to be applied when conducting an internship [12].

The same notion goes to video animation. The message is delivered effectively using video. The research that was conducted by Duygu Sonmez and Meral Hakverdi-Can [28] explained that video and application could affect users' comprehension in teaching reflectively.

The result of this study support the research by Mohd dan Elmagzoub A [29], Komalasari, et al. [30], Djamas, et al. [31], and Richter research's [32]. These researchers found out that learning through application brings a positive effect to the learning result. In this research, it was also found that the mentor teacher understood it more on how to do supervision job during the internship. Likewise the results of research conducted by DeVore et al. [33], Linton [34], Hamdi, et al. [35] Syurigin and Krasnova [36], Lailiyah et al. [37] and Darmawan [38].

3.5. Curriculum Analysis

During the internship program, eight supervisions were conducted. The internship was for two months on holiday between July and September every year [3]. The first supervision is an introduction, observation, school introduction, and sit in with mentor teacher as a complete model. The three activities were conducted in the first week within the supervision activity.

The second supervision is a practical observation which was guidance teaching (with the composition of mentor teacher took up 75% and novice teacher took up 25%) and conference (discussion between mentor teacher, supervisor and novice teacher) on a learning process that has been conducted and lesson plan preparation. Other points that were discussed were: practical guidance teaching (mentor teacher took up 50%, and novice teacher took up 50%), and conference (discussion between mentor teacher, supervisor, and novice teacher) about the learning process that has been conducted and lesson plan preparation. The second supervision was done in the second week.

The third supervision was a practical observation on guidance teaching (mentor teacher took up 25%, and novice teacher took up 75%) and a conference (discussion between mentor teacher, supervisor, and novice teacher) on the learning process that has been conducted and lesson plan preparation. This supervision was conducted in the third week of the internship program.

The fourth supervision is a practical observation on independent teaching and a conference (discussion with a mentor teacher, supervisor and novice teacher) on the learning process that has been conducted and lesson plan preparation. This was done in the fourth week.

The fifth supervision is on the fifth week. The supervision was a practical observation on independent teaching and a conference (discussion between mentor teacher, supervisor, and novice teacher) on the learning process that has been conducted and lesson plan preparation.

The sixth supervision is a practical observation on independent teaching and a conference (discussion between mentor teacher, supervisor, and novice teacher) on the learning process that has been conducted and lesson plan preparation. The seventh supervision is a practical observation on independent teaching and a conference (discussion between mentor teacher, supervisor and novice teacher) on the learning process that has been conducted and lesson plan preparation. This was conducted during the sixth week during the internship program. The eighth supervision was an observation on presentation preparation about the internship result and the finishing of the report. This was done in the sixth week.

The activities above can be visualized in a Macromedia flash application. This application was used in order to aid internship guidance comprehension in the means of attractive visual for users. Not only that, the internship guidance scheme as part of the curriculum becomes more organized and systematic. This is because each supervision has a clear aim and target and therefore, could improve the professionalism of novice teacher [16]. This process could be said as a professionalism improvement process by providing review activity in the reflection during the conference. This is also understood as the implementation of democracy in education [39].

The basis of this development and research is on the analysis of curriculum expert. The result was the internship curriculum design that will influence professionalism achievement of a novice teacher. This is as what Capka [40] has found that the curriculum design could influence student professionalism in the field of accountancy. Another research was by Jasperson [41] which stated that the internship curriculum with reflection principles could improve students' self-development. The same result was concluded by Johnson [42] who found that the internship process contributes to student's carrier development after graduation.

4. Conclusion

Internship model application development for novice teacher using the Islamic Critical Reflection was conducted based on four experts. The Islamic religion education expert stated that this application represents Islamic teaching principles as the basis of educational theory development. Application expert stated that this application meets the easiness and readability aspect of TAM (Technology Acceptance Model). Education technology expert stated that educational aspect through reflective teaching is used for feedback to improve the professionalism of novice teacher. It is found in this application. Curriculum expert stated that the cycle of 2-month supervision in this application is a good way in improving novice teacher's professionalism Conflict of Interest.

Acknowledgment

Our gratitude goes to UIN Sunan Ampel Surabaya for giving a research grant in 2018.

References

[1] *RIP UIN Sunan Ampel Surabaya Tahun 2019-2045,*

[2] "Peraturan Menteri Agama Republik Indonesia Nomor 56 Tahun 2015 Tentang Statuta Universitas Islam Negeri Sunan Ampel Surabaya." p. 6.
[3] Tim Penyusun, Buku Panduan PPL I Fakultas Tarbiyah dan Keguruan Universitas Islam Negeri Sunan AMpel. 2017.
[4] A. Fred and J. P. A. M. Kessels, "Linking Theory and Practice : Changing the Pedagogy of Teacher Education."
[5] T. Gale, "Preparing Professionals : student teachers and their supervisors at work Asia-Pacific Journal of Teacher," no. November 2018.
[6] M. Murray & Kujundzic, N, *critical Reflection : A text book for critical thinking. quebec*. Canada: McGill-queens University Press, 2005.
[7] Venkatesh, Morris, and Davis, "User Acceptance of Information Technology: Toward a Unified View," *MIS Q.*, vol. 27, no. 3, p. 425, 2017.
[8] H. K. Lui and R. Jamieson, "eTransformation TRiTAM : A Model for Integrating Trust and Risk Perceptions in Business-to-Consumer Electronic Commerce," *Methodology*, no. September 2001, pp. 349–364, 2003.
[9] Jogiyanto, *Sistem Informasi Keperilakuan*. Yogyakarta: ANDI Offset, 2008.
[10] T. Ufuk, "Analysis Of Pre-Service Science Teachers ' Views About The Methods Which," no. October, pp. 179–189, 2014.
[11] J. Peechattu, "Reflective practices: a means to teacher development," pp. 1–6, 2017.
[12] S. I. K. Gonen, "A Study on Reflective Reciprocal Peer Coaching for Pre-service Teachers : Change in Reflectivity," vol. 4, no. 7, pp. 221–235, 2016.
[13] J. P. G. Gall Walter R. Borg, Educational research : an introduction / Meredith D. 1942.
[14] M. Harvey, "Editorial - Reflection for Learning in Higher Education," vol. 13, no. 2, 2016.
[15] A. H. Maslow, "Toward a Psychology of Being. A Psychology Classic." Start Publishing LLC, 2012.
[16] E. Rusydiyah, "Rehearsal Model as the Practicum Model in Preparing Professional Pre-service English Teachers," in *International Conference on English Language Teaching (ICONELT 2017)*, 2017.
[17] M. Schulte, "Adult Learning Degree and Career Pathways: Allusions to Maslow's Hierarchy of Needs," *J. Contin. High. Educ.*, vol. 66, no. 1, pp. 62–64, 2018.
[18] N. A. Medcalf, T. J. Hoffman, and C. Boatwright, "Children's dreams viewed through the prism of Maslow's hierarchy of needs," *Early Child Dev. Care*, vol. 183, no. 9, pp. 1324–1338, 2013.
[19] G. Emad, "Self-fulfillment development among the Arab sector in Israel," *J. Educ. Train. Stud.*, vol. 5, no. 5, pp. 81–92, 2017.
[20] E. Sari and R. Dwiarti, "Pendekatan Hierarki Abraham Maslow pada prestasi kerja karyawan PT. Madubaru (PG Madukismo) Yogyakarta," *J. Perilaku dan Strategy. Bisnis*, vol. 6, no. 1 February, pp. 58–77, 2018.
[21] J. Fred D. Davis, "A Technology Acceptance Model For Empirically Testing New End-User Information Systems: Theory And Results," 1985.
[22] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "Extrinsic and Intrinsic Motivation to Use Computers in the Workplace' FRED D. DAVIS~," 1992.
[23] J. Fred D. Davis, "A Technology Acceptance Model For Empirically Testing New End-User Information Systems: Theory And Results," 1985.
[24] T. Hendrawati, "Analisis Penerimaan Sistem Informasi Integrated Library System (Inlis)," *Visi Pustaka*, vol. 15, no. 3, 2013.
[25] E. Koral Gümüsoglu and E. Akay, "Measuring Technology Acceptance Level of Teachers by Using Unified Theory of Acceptance and Use of Technology, Online Submission," *Int. J. Lang. Educ. Teaching. v5*, no. 4, pp. P378-394, 2017.
[26] M. Akar and S. Guzin, "A Structural Model for Relationship between Web Pedagogic Content Knowledge and Technology Acceptance of Preservice Teachers.," *Malaysian Online J. Educ. Technol.*, vol. 7, no. 1, pp. 1–14, 2019.
[27] A. Konak, S. Kulturel-Konak, M. Nasereddin, and M. R. Bartolacci, "Impact of Collaborative Work on Technology Acceptance: A Case Study from Virtual Computing.," *J. Inf. Technol. Educ.*, vol. 16, no. 1, 2017.
[28] D. Sonmez, "Videos as an Instructional Tool in Pre-service Science Teacher Education Meral Hakverdi-Can **," *Eurasian J. Educ. Res.*, vol. 46, no. 46, pp. 141–158, 2012.
[29] M. Babiker and A. Elmagzoub, "For Effective Use of Multimedia in Education, Teachers Must Develop Their Own Educational Multimedia Applications.," *Turkish Online J. Educ. Technol.*, vol. 14, no. 4, pp. 62–68, 2015.
[30] K. Komalasari and D. Saripudin, "Value-Based Interactive Multimedia Development through Integrated Practice for the Formation of Students' Character.," *Turkish Online J. Educ. Technol.*, vol. 16, no. 4, pp. 179–186, 2017.
[31] D. Djamas and V. Tinedi, "Development of Interactive Multimedia Learning Materials for Improving Critical Thinking Skills," *Int. J. Inf. Commun. Technol. Educ.*, vol. 14, no. 4, pp. 66–84, 2018.
[32] J. Richter, K. Scheiter, and A. Eitel, "Signaling text-picture relations in multimedia learning: A comprehensive meta-analysis," *Educ. Res. Rev.*, vol. 17, pp. 19–36, 2016.
[33] S. DeVore, E. Marshman, and C. Singh, "Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics," *Phys. Rev. Phys. Educ. Res.*, vol. 13, no. 1, p. 10127, 2017.
[34] J. N. Linton, "Institutional Factors for Supporting Electronic Learning Communities.," *Online Learn.*, vol. 21, no. 1, pp. 238–256, 2017.
[35] A. Z. Hamdi, A. H. Asyhar, Y. Farida, N. Ulinnuha, D. C. R. Novitasari, and A. Zaenal, "Sentiment Analysis of Regional Head Candidate's Electability from the National Mass Media Perspective Using the Text Mining Algorithm."
[36] V. Y. Shurygin and L. A. Krasnova, "Electronic Learning Courses as a Means to Activate Students' Independent Work in Studying Physics.," *Int. J. Environ. Sci. Educ.*, vol. 11, no. 8, pp. 1743–1751, 2016.
[37] S. Lailiyah, T. Nusantara, C. Sa'Dijah, E. B. Irawan, Kusaeri, and A. H. Asyhar, "Structuring students' analogical reasoning in solving algebra problem," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 296, no. 1, 2018.
[38] D. Darmawan, H. Kartawinata, and W. Astorina, "Development of Web-Based Electronic Learning System (WELS) in Improving the Effectiveness of the Study at Vocational High School" Dharma Nusantara"..," *JCS*, vol. 14, no. 4, pp. 562–573, 2018.
[39] C. M. Reigeluth, Instructional Design theories and models. 1999.
[40] J. Capka and C. Foltin, "Policy Implications of a Proposed Framework to Improve the Accessibility and Effectiveness of Internships in Accounting.," *eJEP eJournal Educ. Policy*, 2017.
[41] J. O. Jasperson, "How to Integrate Student Internships into Legal Studies Research and Curriculum: A Case Study.," *Univers. J. Educ. Res.*, vol. 5, no. 9, pp. 1504–1509, 2017.
[42] K. W. Johnson, "A Case Study Exploration of Internships in Undergraduate Business Education," 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 106-116 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

An Approach to Vulnerabilities, Threats and Risk in Voting Systems for Popular Elections in Latin America

Segundo Moisés Toapanta Toapanta*[1], Iván Fernando Marriott Saá[1], Félix Gustavo Mendoza Quimi[1], Luis Enrique Mafla Gallegos[2]

[1]*Department Computer Science, Universidad Politécnica Salesiana (UPS), Guayaquil, Ecuador*

[2]*Faculty of Engineering Systems, Escuela Politécnica Nacional del Ecuador (EPN), Quito, Ecuador*

ARTICLE INFO

Article history:
Received: 28 March, 2019
Accepted: 17 May, 2019
Online: 24 May, 2019

Keywords:
Elections
Latin America
Voting
E-Voting

ABSTRACT

The problems in information security regarding vulnerabilities, threats and risks in voting systems for popular election in Latin America and the world persist; because in most of the countries of the world there is no maturity in democracy and defined policies; the problems of confidentiality, integrity and authenticity in the electoral processes can be evidenced. The objective is to perform the analysis to identify the threats, risks and weaknesses in electoral systems in Latin American countries and determine which of the systems used by different countries may be appropriate, to be considered as an alternative. The deductive method and exploratory research has been used to perform the analysis of the articles and information regarding electoral processes. It resulted in a description in statistical tables of the threats and weaknesses that must be examined to implement a system for electoral voting; considering the culture, technological availability and social conditions of each country. It was concluded that to mitigate the potential risks of the information, it is necessary to identify the weaknesses in the electronic voting system to improve the integrity and security of the electoral process; made in the last three presidential elections in Brazil, Colombia, Ecuador.

1. Introduction

Numerous countries from Latin America incorporate several referendum systems for political elections. When exposed to the different cultures of countries, they can pose several risks and threats to the voting process.

Since the 90's, Latin America commenced a slow but steady process of adapting to new technologies and making use of them to enter the trend of adopting E-Voting. In 1996 and 1998 respectively, Brazil and Venezuela were the first nation to interact with this new system [1].

The desire involved with enforcing automation in voting systems is to optimize this process as well as the distribution of the outcome [2]. It aims an optimal way to try to minimize the election time as well as the waiting period of the outcome as much as possible. This has as a sole goal supplying security while decreasing the potential for election tampering at the polls and providing fairness and integrity to the electoral process.

Some voting system processes require massive data handling and processing. Therefore the implementation of computerized procedures has obtained a significant advantage in terms of speed, security, and integrity [3].

Although not all countries in Latin America force or punish people who do not wish to exercise their right to vote, all citizens should be aware of the value of their vote in a democratic system. If they do not show concern for the person who will govern their country, it will be easier for them to be victims of abuse of power.

Why is it essential to perform an analysis in Latin America electoral systems in order to have free and fair elections?

Currently, the voting systems used presents different vulnerabilities, threats, and inherent risks. If not counteracted, these will have a serious negative impact on the voting process.

This research is organized as follows. In the second section, the materials and methods are showed. In the third section, the findings of the research are exposed. In the fourth section, the

*Segundo Moisés Toapanta Toapanta, & Email: stoapanta@ups.edu.ec

results are inspected. Based on these discussions future work is discussed in the fifth section. Finally, in the sixth section the conclusions of this research are presented, as well as some final remarks.

The ultimate goal of this research is to deliver an analysis focused on the distinctive properties in election systems introduced in Latin American countries. Furthermore, determine which of these systems is the most appropriate to be regarded as an efficient and viable alternative.

The deductive method is used to evaluate relevant information from articles pertaining to this research. The articles and books reviewed in relation to the subject are summarized in the Appendix.

The results reached are as follows:

1. A diagram that lists the qualities which are directly linked to non-functional and functional requirements.
2. A diagram showing the inherent vulnerabilities of the different type of voting systems, also caused by levels of malfeasance, fraud, and electronic illiteracy.
3. A table that identifies the various levels of criticality divided by a color code and a complementary table which summarizes the different risks that negatively impacts the voting systems.
4. Based on the historical information of the last three elections in Brazil, Colombia and Ecuador, data have been compiled and tabulated in tables that allow us to analyze the reality of the electoral processes that are being carried out at present.

It is finally concluded that once the weaknesses in Voting Systems have been alleviated and/or eradicated, from an objective perspective, E-Voting becomes the least risky and most efficient to incorporate in Latin America.

2. Materials and Methods

2.1. Materials

Latin America is made up of a large number of countries. For this study are taken into account those who can cover the greatest differences between populations, area, and implemented voting system as seen in Table 1, being these: Brazil, Colombia and Ecuador.

Brazil is one of the largest countries with the biggest population in Latin America and one of the pioneers in successfully migrating from traditional elections to electronic voting. Brazil has streamline the voting process and instituting the electronic voting system as the one used throughout the country since 1996 [4]. There are many countries that have decided to implement some registered pilot systems such as the case of Ecuador, which in 2004 decided to carry out E-Voting tests in some of its provinces, resulting in the country deciding to use the automated paper-based voting; contrary to Colombia, despite being one of the pioneers in 1992 when conducting an E-Voting pilot, has steadfastly refused to change keeping the traditional voting system [5].

Table 1: Differences between the Latin America representative countries.

Country	Population	Area	Voting System
Brazil	208494900	8,516 millions km²	Direct-recording Electronic
Colombia	49292000	1.142 millions km²	Voting Ballot
Ecuador	17096789	283,560 km²	Automated paper-based voting

There are several types of popular vote electoral system but those used in the selected Latin American countries can be classified mainly into two types.

- *Voting ballot system*

Many countries in Latin America use the ballot system referred from now on as a traditional election. Ballot system gives to their citizens the opportunity to be part of the selection of the political entity that will represent their country [6]. In the traditional election system, the regulatory authorities are responsible for providing voters with ballots and ballot boxes so they can exercise their right to cast a ballot. Consequently, the counting of the votes is carried out manually. Nevertheless, to announce the results, it must be taken into account that the winning candidate must win the absolute majority of valid votes that is more than 50% of the voters. If none of the voters exceeds this expectation, the election process must be repeated.

As a result, elections are made using a ballot and an urn. This particular system needs the direct involvement of a vast number of individuals belonging to a community that in the overwhelming majority of situations were not properly trained.

- *Electronic Voting System*

This type of system is also known as E-Voting. It was used for the first time in Latin America for the elections of president, state governors and legislators in 2000 in Brazil. Since 1996, different municipalities of the country started with the installation of electronic ballot boxes.

This has various mechanisms that store the votes where the elector or civil society entities can verify the security and the perfect functioning of the system. It depends exclusively on the reliability of the software used by the device, so the use of this system has been extended to Argentina and Mexico where the corresponding pilot tests have been carried out using the same technology [7].

The E-Voting electoral system has been divided into three different categories:

1. Automated paper-based voting

Among the three categories in which the E-Voting voting system has been divided, the automated reading of paper-based

voting is the most inaccurate since both voting and counting are done by hand as in traditional elections. The information collected is transferred to paper and with special characters that are processable. Each of the results of the count is entered into machines designed specifically for the processing of this data, which can take between days and weeks.

2. Direct recording electronic electoral system (DRE)

In this type of voting systems, machines created specifically for this purpose are required, also known as electronic ballot boxes. They function in a similar way than a computer since the voting is done manually, but the machine offers simplicity in voting through either touch screens or buttons for selection. After the elections, the stored data are exported by means of some removable memory device and/or with a printed copy so that the voter and the entities belonging to civil society can check the security and integrity of the vote.

3. DRE on public network

The implementation of various technological advances has allowed some countries to opt for the use of a public network to offer citizens a more effective way to get involved in the voting process. Once the voting is done, the data will be encrypted to avoid manipulation by external agents. Then they can be broadcasted as individual results or as a final batch at the end of the elections. The computer will provide the results.

According to [3, 8, 9] any of the electoral systems, whether traditional or E-Voting must have the following actors to function correctly:

Voter: Person who is entitled to vote. In many countries, citizens who are eligible to vote are those who are literate over the age of 18 and under the age of retirement in their country in some countries, this group is obligated to vote. People who are between the ages of 16 and 18 are allowed to exercise their right to vote, not being obligatory for retired or illiterate people.

Registration Authority: Authority responsible for ensuring the voting rights of voters. They provide real information of citizens who are duly registered and present the characteristics that consider them eligible to effectuate their right to vote.

Counting Authority: In Latin America, it is an obligation of the state to have various electoral institutions. They have the purpose of ensuring the planning and fulfillment of the elections, and controlling the registration of the different applicants to candidates and their respective associations during political operations from the moment it is decided to carry out an electoral process until the results are disseminated. In some cases, they are also in charge of providing the infrastructure to carry out the elections. These are usually the educational centers located near the residence of the voters to facilitate access at the time of the referendum.

Additionally, electoral systems regardless of the type should include five general phases:

Registration: It is the responsibility of the state that each country has a civil registry where all citizens are enrolled without any exception. This system allows you to collect the all kinds of information to determine if citizens have all the requirements that allow them to exercise their right to vote.

Authentication: Each electoral institution designates various authorities that comply with the verification of identifications of all those who attend to the designated election site. If the citizen presents the designated documents and complies with all the necessary regulations, then the responsible authority has the obligation to register that the voter has participated in the corresponding election.

Voting: At the time of voting, it must be taken into account that there are certain citizens who have a preference at the time of voting, such as people over 60, ill people, pregnant women, people with disabilities, and women carrying children. In some countries, the law obliges people to vote. If a voter who is forced to vote does not show up, the voter will be in debt with the Electoral Justice.

Vote counting: Once the election period is over, the authorities in charge restrict the passage of people from outside the process and proceed to carry out the corresponding accounting of the votes. Depending on the country where the election takes place, the vote count can vary from hours to days, which generates concern among the citizens since the results can generate controversy when the information is filtered.

Results Delivery: Once the vote count has been carried out by the designated persons, the authorities belonging to the electoral institutions have the duty to verify if the results obtained in the previous process have coherence in order to formally disclose the results using the chosen means of communication.

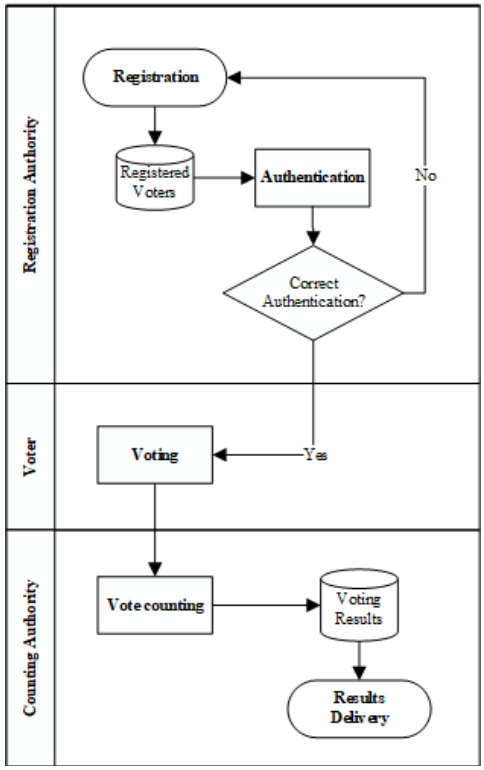


Figure 1: Electoral system Process Flow.

Figure 1 displays a general electoral system process flow regardless of type. It includes three different actors and five general phases, as discussed above.

Furthermore, according to [10] both voting systems consist of different types of rules which may differ depending on the country that performs the electoral process. There are mainly two types of electoral systems that define the voting rules independently of the voting system employed, the majority in which the one with the most votes wins and the proportional ones

in which it is sought to ensure that the charges are distributed proportionally in relation to the number of votes received by competitors. There is a third type known as a mixed system that is the combination of certain characteristics and rules of the previous types.

In Latin America, the proportional election system is currently used. This is subdivided into a list and a single transferable vote and the mixed election system, which may be dependent or independent.

- **Majoritarian Systems**

As the name implies, are those in which the candidate or party with the greatest number of votes is chosen as the winner. This type of systems have the distinction of having two types of rules that can be applied to the same situation, such as the single member districts in which the winning candidate is the one with the most votes even though it does not exceed 50% of total votes.

This does not happen in the two-shift system that performs what is known as the second round when no candidate crosses the 50% barrier. When there is more than one seat available for political posts, multi-member models known as plurinominal district are used, where the voters can grant their vote to more than one candidate, taking care not to exceed the maximum number of seats. This type of model is used for the election of presidents, mayors, governors and senators.

- **Proportional Representation**

Unlike majority systems, the principle of the proportional model is to ensure that minor parties also achieve representation. These parties do not need to receive the majority of votes to be elected as indicated by the single transferable vote where the voter can vote for more than one party ranking their choices according to their preferences.

When counting, the first option of each voter is taken into account and the less voted parties votes are successively redistributed according to the order of choice. These proportional systems also have the voting list where each party presents a candidate list to the voter, depending on the country and the type of election that is being made. The voter is allowed to choose the party or a specific candidate from this list.

- **Mixed Systems**

Some countries such as Bolivia, Venezuela and Mexico have adopted mixed systems. They use proportional and majority models at the same time to elect representatives. The two models can be independent, if they exist in parallel, or dependent, if the voting rule of one system influences another. It is common that in mixed systems, voters must vote twice. The first uses a majority system as single member districts and the second implements a proportional representation system.

2.2. Methods

The Deductive Method has been used for this research. It has been possible to determine the threats and vulnerabilities that negatively affect electoral systems.

Voting Systems Characteristics

As stated in [11–13], the defining features that an election system must comply to be effective are the following:

1. Anonymity of the voter

The vote of a voter must remain secret. Everyone can know the voters and the candidates, but only the voters can know their votes. This means that no historical record of the voters' elections should be kept. In E-Voting systems, a digital record of the vote is used. This is an electronic record generated by the same ballot box that counts votes and fulfills the votes storing function without losing the characteristic of anonymity that the traditional voting system has.

If external agents wanted to know the identity of a particular voter, they should know the system in its minimum details for the extraction of the digital record. In order to visualize the original order of the votes' incoming, a computerized system records the time of access to the electronic ballot box. In order to carry out voter identification, it is only necessary to monitor the voting order of the voters and compare them with the entry record.

In most electronic voting processes, voter's credentials are usually encrypted, but total security does not exist. Originally, the servers of those who administer the voting can decrypt the original information. Therefore, in these cases the anonymity of the vote is not fully guaranteed.

2. Detection of falsification of votes

The system needs to be able to tell the difference between a false and a real vote. The ballot paper is audited manually in traditional referendums while in E-Voting is the system that needs to be able to differentiate.

According to experts, it is possible to produce the exchange of votes only with a movement generated by the operator of the system which would produce a falsification of votes. The destruction of hard drives of electronic voting is also a particular reason for doubt. As claimed by the specialist in IT (Information Technology) Märt Põder, electronic elections are a semi-unprotected wall of a city [14].

3. Detection of duplication of votes

In the traditional system, voters can only vote once. It is necessary to keep a historical record of the people who voted, in order to keep track of the total number of votes and there are not duplicated votes.

However, if the electronic voting process does not have sufficient security guaranteed. In the case of DRE on public network, there is the possibility that different votes will be received from a single IP address, thus altering the results of the elections. There may also be the case of a family of several members, who share internet access, how can it be identified if the IP address belongs to one or more people?

4. Integrity of the ballot paper

Once the vote record is either saved, as a record in a database or placed in the voting booth, it cannot be changed. The extraction of this data is done through a mobile medium. This medium contains the data of the whole voting process. Once the whole process is finished, the ballot box prints the result of that electoral section.

In this way, the result of the election becomes transparent and of public knowledge as soon as voting is completed. This procedure occurs simultaneously in all the electoral sections.

Some countries use the electronic voting system using ballot boxes that do not possess Internet connection or any means of data transmission, being the only cable connected to this the power cord. Moreover, if necessary, it may only be connected to an auxiliary battery for example, if power is missing.

Imagine a safe in which a large sum of money is saved. To find out if the content has changed, simply count the money and compared with the balance stored on paper in your wallet. If the values are not equal, it means that the balance or contents of the safe have changed. The electronic ballot box would be a safe in which the balance and the money are stored in the same place. In case someone can open the safe, it would be possible to withdraw the money and at the same time, change the balance written on the paper to reflect the new value. In this way, it would be difficult to detect that the safe was compromised.

5. Voting availability

One of the indispensable principles that should guide elections with electronic votes is the absolute availability which ensures that elections cannot be postponed or temporarily interrupted by unavailability of the ballot boxes. They must be stable and secure enough so that the entire voting process takes place during the date and time set by the government. Therefore, the voter is able to exercise his right to vote within the date and time established by the institution in charge.

6. Accuracy in the count

In countries with a particularly dense population that use the traditional system, the voting and counting process takes several weeks increasing the uncertainty in the voters. The electronic voting systems allow the population to know the results in a shorter time. This can generate more confidence because the votes cannot be altered, duplicated or eliminated without this action going unnoticed.

The electronic voting system will not allow invalid votes to be counted nor included in the results. In DRE systems, the process of registering and counting votes is free of human errors, and just as in the traditional voting system, the total of votes must be equal to the total of voters who attended to vote.

7. Security in the system

The electoral system needs to be resistant against failures and attacks. Corruption attempts by authorities, voters, and hardware or software failures should not affect the system.

Inevitably, exhaustive discussions have been raised regarding the current consequences that the use of computer technology can have for the security and integrity of the election process.

Moreover, malicious software can alter the electronic printing on the screen altering the voter eagerness to cast a vote. In the same way, an intruder can manipulate the outcome of the results. People who oppose or challenge E-Voting give several reasons supposed flaws which might influence the accuracy of the electoral process.

Nevertheless, several technical answers ensure that these exploitable vulnerabilities are entirely fictitious. Depending on the used system, traditional voting may be more susceptible to alterations than an efficient E-Voting system.

When the E-Voting systems are subject to external security audits, cases of adulteration of votes and registers are frequent. As no system is free from attacks, public administrations are not exempt from the fact that their institutional IT security can be attacked or violated. Thus, it is essential the developing of a voting system capable of resisting attacks and failures, as well as not being affected by corruption attempts by internal and external agents such as the authorities, voters or even programmers.

8. Simplicity of use

One of the most important characteristics when implementing the electronic voting system is to facilitate citizens the possibility of participating in democratic processes, encouraging them not to abstain and increasing the total records. In addition, the system must be simple enough for anyone to use it. Its simplicity is also an advantage for those citizens with limited or non-existent technical skills, disabled voters and citizens residing in a city other than their own to vote.

Voting Systems Vulnerabilities

In [11, 13–17] several vulnerabilities within the Voting systems are identified, being the most important:

1. Social engineering

It is a term that describes attacks that are based on deceiving a person so they willingly hand over their private data without realizing it. For example in the real world, criminals by searching your garbage can steal your personal information, thus finding sensitive data and carrying out transactions with them.

Figuring out the name of your pet, school, or any other information that may be used as a security question in the accounts of the Internet, your account may be compromised. These are attacks identified as low level, but they are still methodical and successful attacks.

2. Digital Division

The world in which we live is separated by the barrier of technology. This affects not only the underdeveloped countries or the rural class, but also the developed countries. In the Internet era, this is called a digital division. Despite the efforts that are being made to equip societies and citizens with digital technologies, the truth is that there are profound inequalities at international, national, regional and local levels. Digital divisions influence the process of electronic voting. Households with Internet and those who do not possess it, faster connection than others are clear examples of digital divisions. This will directly affect the capacity of access the system.

3. Security problems related to the client and the server

It can never be assumed that the user's computer is completely secure. Due to the architecture of current systems, it is possible for external agents to use various mechanisms to infiltrate users' systems in the form of viruses that allow them to spy on the ballots, prevent voters from voting, or directly modify the votes.

However, there are countries that offer a solution for this type of altercations. They offer special open source and secure voting systems so the state is responsible for the electronic machine software at the operating system level. These provides voting booths or special machines that have a much lower probability of doing something wrong due to viruses or the inappropriate use of computers.

4. Security problems related to the connection

The connection refers to the link between the client who votes and the server which is responsible for counting the votes.

One of the main problems when using the electronic voting system is the distrust in the reliability of the system. Many people believe that it is safer to carry out the voting in a traditional way since there is a connection between the voter and the authority responsible for votes counting instead of using a machine that collects and analyzes the votes quickly and efficiently. To grant the corresponding results in the least amount of time possible. For DRE on public network, this link must be trusted and secure for as long as the records are being delivered to the server.

5. Vote purchase

There are many disadvantages in any type of voting. The security of electronic voting is of the most criticized, not for only being vulnerable to internal attacks but also for the possibility that external agents collects votes by any means. They do it either directly, in a way of coercion or threats or indirectly imposing the delusion that the candidate is good. This usually involves giving salaries, gifts or assignments for specific services to potential voters, such as distributing notes, and hanging posters.

Currently there is no real protection for the problem of buying votes in any of the electoral systems because it is much easier to organize than discover it.

6. Disenfranchisement

In agreement with the law in certain countries, the voting rights for people or a specific group of individuals is restricted. This in some cases can occur without legal basis due to incoordination or mismanagement, deliberately or involuntarily. It can be given in the case of people who are deprived of liberty and are not disqualified to vote. The decision to restrict this right is due to the fact that people who are deprived of their liberty have committed a crime that violates the constitutional laws that govern a country and therefore they are not allowed to participate in processes that would compromise society as such.

7. Handling of the Voter's data

It consists in manipulating the information of the voter to harm it. This information may be demographic or historical data which can be used by political entities to postulate content that favor the political campaign in their favor. The E-Voting is done through machines using software that records vote information that could be manipulated. For example, when we pay a ticket at the Bank, we can save the voucher or receipt in a PDF file. If the same fee is subsequently charged again, the voucher allows proving that the debt has been paid. Would you trust a bank that after paying a ticket the money simply disappeared from your balance without the destination of that money being at least registered in your bank record? Most likely the answer would be no.

Similarly, the electoral system used in Brazil has exactly that problem so it has decided to implement anonymous vouchers that allow the user to ensure that their vote has been made with satisfaction for the comparison of the values recorded in the electronic memory.

8. Ballot Tracking

When conducting an electoral process, one of the guidelines is to ensure that the materials provided by the authorities are protected and in good condition. They are designated to various groups of Special Forces such as military, navy, police and others to monitor them by preventing other individuals from doing scrupulous acts that can compromise the elections.

This event of ballot tracking is done when the designated sites by the electoral institutions do not have the necessary equipment to perform the final vote counting that will indicate the results.

9. Ballot Stuffing

The filling of ballots is one of the vulnerabilities of the traditional voting system. Many citizens are determined to commit electoral fraud so that their preferred candidate can win. This happens when they do not have control of what happens in the ballot box since this individual could be carrying their own ballots or the person in charge of administering the ballots could adulterate them and pass them off as votes of absentees. Depending on the country, the penalty varies between months up to 10 years.

10. Digital Illiteracy

Around millions of people in Latin America lose their connection in an increasingly digital world. Mainly, people who live in the rural regions of their country and people over 55. This lack of knowledge in the digital age does not allow them to enjoy a variety of benefits without the need to get assistance from a person who has the capabilities to help.

In recent years, several government institutions offer services via the Internet to streamline all types of processes. Therefore, if the person cannot function quickly and efficiently in the digital world, he/she is marginalized from the rest of their social environment.

3. Results

In the present work, it can be deduced that any of the two voting systems presents characteristics and vulnerabilities that must be considered when trying to implement a system that satisfies all the requirements that are needed for a satisfactory electoral process.

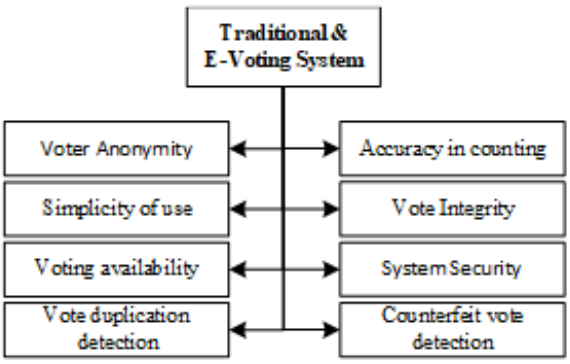


Figure 2: Characteristics of voting systems.

Figure 2 contains a diagram with eight characteristics of both voting systems, and these must be present when using any of them. If one of these fails, the results of the election could be compromised.

Likewise, Figure 3 displays a diagram of the weaknesses affecting the two varying types of electoral systems. Both systems

have different vulnerabilities. Some of these vulnerabilities are shared among them.

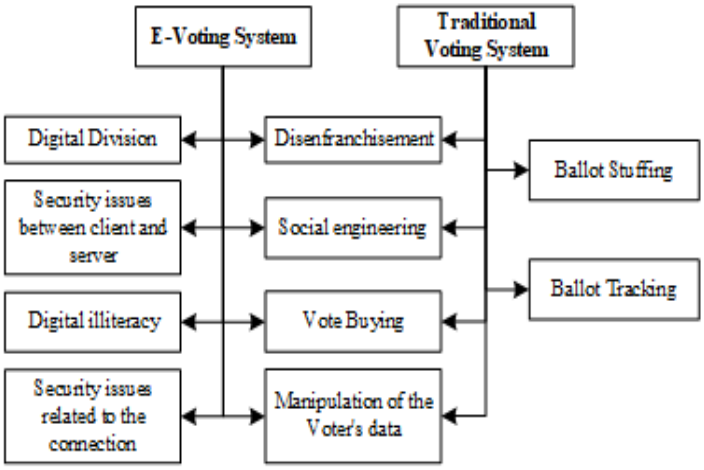


Figure 3: Vulnerabilities affecting the voting systems.

Therefore, it is discernible that half of the vulnerabilities present in E-Voting are related to the appropriate management of technology while the other vulnerabilities presented in both systems are related to society. This leads to understanding that the traditional voting system is more exposed to human error and/or social interference.

Table 2 defines the different levels of criticality of voting systems. Four levels have been considered which are separated by color code.

Subsequently in Table 3, the various risks affecting the systems are detailed and grouped by type of risks and levels of criticality. As can be seen, the E-Voting system presents two types of risk that would have a great impact on society while the traditional voting system does not have that kind of impact. Social risk influences each of the risks present in both types of voting system. This means, society is the one that decides whether the chosen voting system is going to be successful or not. Both systems present the same amount of moderate risk. As a result, the three risks that do not compromise the system belong to the E-Voting system and are not subject to technological errors.

Table 2: Criticality levels

Impact	Description	Color Code
Very Severe	The risk is high, directly affects the voting system and can change the outcome of the vote.	Red
Serious	The risk is moderate and can alter the outcome of the election if it is not detected and corrected on time.	Orange
Moderate	The risk is minor and does not directly affect the voting system.	Yellow
Mild	The system has not been jeopardize.	Green

The voting systems, when fulfilling these characteristics, can mitigate their inherent vulnerabilities thus eliminating the hazard and dangers presented when implemented.

Pursuant to the data collected and exposed by governments in historical reports [18 - 28] in Table 4A and Table 4B, it can be observed the citizen participation in the 3 selected countries. There were several electoral processes in the election of Presidents and Vice-presidents made every 4 years. In 10 out of 12 elections, a second round was necessary for the election of its president.

In conducting the research, work was concluded that there are big gaps over the time to take, collect, organize, and disseminate the results depending on the type of voting system used in each country.

Table 5 data shows a comparison between each of the countries. For example, between Brazil and Colombia, Colombia represents only 23.30% of registered voters whereas in Brazil only 49.64% registered voters showed up to exercise their right to vote.

Comparing the countries of Ecuador and Colombia, they present a difference of 30.85% in citizen participation. Although according to Table 1, Colombia has a greater population than Ecuador. Less than half of the registered individuals in Colombia vote. The collection and dissemination of results may take between 9 to 10 hours less than in Ecuador.

Table 3: Risks affecting voting systems.

Risk	System	Type of Risk	Impact
Security issues related to the connection	E-Voting	Technological/ Social	Very Severe
Related security issues with the client and the server	E-Voting	Technological/ Social	Very Severe
Social engineering	E-Voting / Traditional	Social	Serious
Ballot Tracking	Traditional	Social	Serious
Ballot Stuffing	Traditional	Social	Serious
Manipulation of the Voter's data	E-Voting / Traditional	Technological/ Social	Serious
Vote Buying	E-Voting / Traditional	Social	Moderate
Disenfranchisement	E-Voting / Traditional	Social	Mild
Digital Division	E-Voting	Technological/ Social	Mild
Digital illiteracy	E-Voting	Technological/ Social	Mild

Table 4A: Results of the three last presidential elections in Brazil, Colombia and Ecuador (First Round).

Presidential Elections	BRAZIL			COLOMBIA			ECUADOR		
	2010	2014	2018	2010	2014	2018	2009	2013	2017
Registered voters	135804084	142822046	147299471	29997574	33023716	36227267	10529765	11675441	12816698
Valid votes	101590153	104023802	107050749	14573593	12160881	19336134	6897912	8602603	9442495
Blank votes	3479340	4420489	3106937	261530	672782	338581	534149	179230	286069
Null votes	6124254	6678592	7206968	170874	351739	242002	496687	684027	736743
Total Voters	111193747	115122883	117364654	15005997	13185402	19916717	7928748	9465860	10465307

Table 4B: Results of the three last presidential elections in Brazil, Colombia and Ecuador (Second Round)

Presidential Elections	BRAZIL			COLOMBIA			ECUADOR
	2010	2014	2018	2010	2014	2018	2017
Registered voters	135804084	142822046	147299471	29997574	33023716	36227267	12816698
Valid votes	99463917	105542273	104838753	13061192	15341383	19247062	9895407
Blank votes	2452597	1921819	2486593	482003	619396	807924	69436
Null votes	4689428	5219787	8608105	198003	403405	265857	670731
Total Voters	106605942	112683879	115933451	13741198	16364184	20320843	10635574

However, it should be taken into account that both Brazil and Ecuador use electronic voting systems. Being DRE and automated paper-based voting respectively, where according to Table 5, Ecuador has a 5.62% more participation in relation to Brazil.

Table 5: Average of the collected data.

Presidential Elections	Brazil	Colombia	Ecuador
	Average		
Registered voters	141975200	33082852	11673968
First Round			
Valid votes	104221568	15356869	8314337
Blank votes	366822	424298	333149
Null votes	6669938	254872	639152
Total Voters	114560428	16036039	9286638
Second Round			
Valid votes	103281648	15883212	9895407
Blank votes	2287003	636441	69436
Null votes	6172440	289088	670731
Total Voters	111741091	16808742	10635574
Average Turnout	79.71%	49.64%	85.33%

Table 6: E-Voting pilots in Latin America.

COUNTRY	Pilot Year	E-Voting
Colombia	1992	DRE
Brazil	1996	DRE
Peru	1996	DRE
Venezuela	1998	DRE
Paraguay	2001	DRE
Venezuela	2002	DRE
Costa Rica	2002	DRE
Argentina	2003	DRE
Ecuador	2004	Paper Based
Dominican Republic	2006	DRE
Mexico	2006	DRE
Argentina	2011	DRE
Argentina	2013	DRE
Ecuador	2014	DRE
Panama	2014	DRE
Argentina	2015	DRE
Peru	2015	DRE
Peru	2018	DRE
Colombia	2018	DRE

4. Discussion

Since 1992, Latin America has decided to be part of technological change. This is why some countries undertook the adaptation process to migrate from the traditional voting system to E-Voting.

According to the data provided by the different counting authorities [29–32], Table 6 shows that around 11 countries decided to implement and register pilot tests from 1992 to 2018. Only Brazil and Venezuela decided the implementation of E-Voting in their current government.

In the present investigation, it was decided to use Brazil as a model representative of the Latin American countries using the electronic voting system in its entirety. Venezuela has been excluded, due to its current situation which has not allowed the collection of electoral data from reliable and truthful sources. It should also be noted that Argentina has made about 4 documented pilot tests, but it has not yet been able to migrate satisfactorily towards electronic voting.

Also in the analysis that has been carried out in the present investigation, the characteristics and vulnerabilities of the

traditional system and the E-voting, indicated in Figure 2 and Figure 3 respectively, have been established. They take into consideration that all the characteristics are necessary for the correct functionality of both systems. Voting Integrity and System Security have been identified as the most fundamental characteristics in voting systems. The main reason for this is that the integrity of the vote ensures that the election is conducted in an anonymously, safely and transparent manner. On the other hand, the security of the system assures that the entire electoral process is free of faults, so safeguarding its effectiveness.

Table 7: Last presidential election time details.

Country	Election Date	Begin	End	Vote Tallying	Tallying time
Brazil	28/10/2018	8:00	17:00	28/10/2018 19:13	2:13:00
Colombia	17/6/2018	8:00	16:00	17/6/2018 19:51	3:51:45
Ecuador	2/4/2017	7:00	17:00	3/4/2017 6:06	13:06:00

However, it must be taken into account that both Brazil and Ecuador uses E-Voting systems. Additionally, Ecuador represents only 8.22% of registered voters of Brazil, where according to Table 5 only the 85.33% show up to vote. Nevertheless, the differences of the various types of E-Voting are reflected in the last two phases. These are the vote counting and the results delivery. According to Table 7 [33–35], we can consider a time factor of 6.13 which equals to 11 hours. In addition, by continuing to use the traditional voting system, Colombia takes 1.38 more hours than Brazil in vote counting depending on the type of system and the authority in charge.

5. Future work

In Ecuador, Article 292 of the Organic Electoral Law of Democracy [36] determines that if people who have the obligation to vote do not show up to vote, they will be sanctioned with 10% of their basic salary. On the other hand, in Brazil article 7 of the electoral code [37] dictates that the citizen must justify his fault before a judge where he will decide the percentage of the fine between 3% and 10% of the minimum wage.

Therefore, it is recommended to conduct a study analyzing the probability that the participation of citizens in choosing their leader is related to the reliability of the voting system or the penalty fee for not exercising their right to vote.

The gradual increase in corruption in Latin America [38] is one of the most relevant facts that is not been considered from a technical point of view, but it still should be considered. In Figure 4, the country corruption indices by country can be observed. Venezuela leads the list with 87%, while Argentina is last with an index of 41%.

As it can be seen in Table 6, Latin America has not conducted a pilot test for the E-Voting System DRE in public network. It is necessary that future research focus on the development of new methodologies and technologies so that it can be correctly implemented.

6. Conclusions

The countries of Latin America that have decided to use both partially and fully E-Voting system have come to the conclusion that once vulnerabilities have been mitigated it becomes the least risky and most efficient to implement. It is a safe, transparent and auditable process.

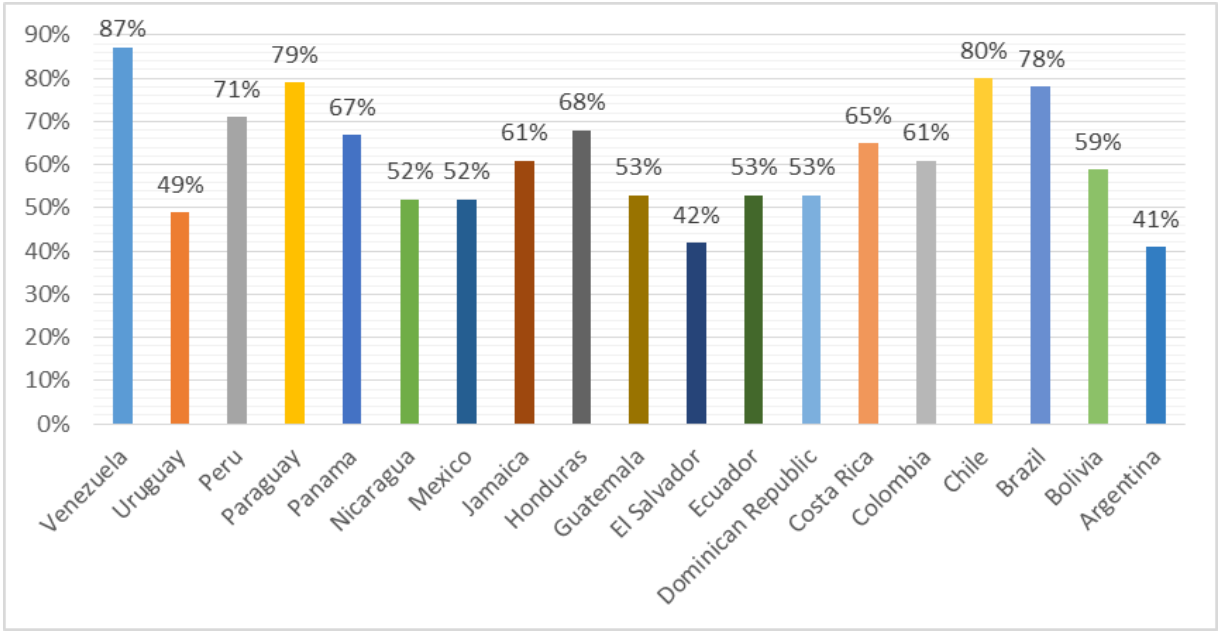


Figure 4: Corruption Indices by Country in Latin America.

Taking into account, voting systems consist of five phases as shown in Figure 1. In the automated paper-based voting, four of these phases are directly linked to the margin of human error while in the other two types of E-Voting the whole process is electronic. It can be discerned that DRE and DRE on public network reduces voting time which leads to the reduction of operational costs. Finally, it was concluded that to alleviate the potential risks of voting systems is necessary to identify and neutralize the weaknesses in the E-Voting system. This can be considered as an alternative approach to increase the security of the process.

Hence, when implementing this type of system, it is essential to consider the security requirements. With this in mind, without the appropriate measures E-Voting can be a real challenge.

7. Acknowledgment

The authors thanks to Universidad Politécnica Salesiana del Ecuador, to the research group of the Guayaquil Headquarters "Computing, Security and Information Technology for a Globalized World" (CSITGW) created according to resolution 142-06-2017-07-19 and Secretaría de Educación Superior Ciencia, Tecnología e Innovación (Senescyt).

8. Appendix

Appendix: Articles and books reviewed in relation to the subject.

Title	Publication Year
E-voting en Colombia: Avances y desafíos en la implementación.	2019
A Scheme for Three-way Secure and Verifiable E-Voting.	2019
The Electoral Success of the Left in Latin America: Is there any room for Spatial Models of Voting ?	2019
Electronic voting	2018
The Good , the Bad , and the Ugly : Two Decades of E-Voting in Brazil	2018
Towards a Secure Online E-voting Protocol Based on Palmprint Features.	2018
Secured and transparent voting system using biometrics.	2018
Election fraud and privacy related issues: Addressing electoral integrity.	2017
Principles of comparative politics	2017
Smart voting	2017
Electronic Voting.	2017
Towards security modeling of E-voting systems	2016
Digital divide impact on e-voting adoption in middle eastern country.	2016
Identification of Non-Functional Requirements for Electronic Voting Systems: A Systematic Mapping.	2015
From piloting to roll-out: Voting experience and trust in the first full e-election in Argentina.	2014
Trust in elections, vote buying, and turnout in Latin America.	2013

References

[1] D. F. Aranha and U. Estadual, "The Good , the Bad , and the Ugly : Two Decades of E-Voting in Brazil," no. December 2018, 2019.

[2] A. Kiayias, "Electronic voting," *Handb. Financ. Cryptogr. Secur.*, no. October 2018, pp. 59–82, 2010. https://doi.org/10.1201/9781420059823.

[3] C. D. De Faveri, A. Moreira, J. Araújo, and V. Amaral, "Towards security modeling of E-voting systems," *Proc. - 2016 IEEE 24th Int. Requir. Eng. Conf. Work. REW 2016*, pp. 145–154, 2017. https://doi.org/10.1109/REW.2016.37.

[4] J. Pomares, I. Levin, R. Alvarez, G. Mirau, and T. Ovejero, "From piloting to roll-out: Voting experience and trust in the first full e-election in Argentina," *2014 6th Int. Conf. Electron. Voting Verif. Vote, EVOTE 2014 - IEEE Proc.*, 2015. https://doi.org/10.1109/EVOTE.2014.7001136.

[5] F. P. Pardo, "E-voting in Colombia: Progress and challenges in its implementation [E-voting en Colombia: Avances y desafíos en la implementación]," *Rev. Derecho del Estado*, no. 42, pp. 211–248, 2019. https://doi.org/10.18601/01229893.n42.08.

[6] S. E. E. Profile, "The Electoral Success of the Left in Latin America : Is there any room for Spatial The Electoral Success of the Left in Latin America : Is there any room for Spatial Models of Voting ? *," no. April, 2019.

[7] M. Carreras and Y. Irepoğlu, "Trust in elections, vote buying, and turnout in Latin America," *Elect. Stud.*, vol. 32, no. 4, pp. 609–619, 2013. https://doi.org/10.1016/j.electstud.2013.07.012.

[8] M. Nassar, Q. Malluhi, and T. Khan, "A Scheme for Three-way Secure and Verifiable E-Voting," *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA*, vol. 2018–Novem, pp. 1–6, 2019. https://doi.org/10.1109/AICCSA.2018.8612810.

[9] A. Meraoumia, H. Bendjenna, M. Amroune, and Y. Dris, "Towards a Secure Online E-voting Protocol Based on Palmprint Features," *Proc. - PAIS 2018 Int. Conf. Pattern Anal. Intell. Syst.*, pp. 1–6, 2018. https://doi.org/10.1109/PAIS.2018.8598520.

[10] W. R. Clark, M. Golder, and S. N. Golder, *Principles of comparative politics*, 3rd Revised Edition. Washington, United States: SAGE Publications Inc, 2017.

[11] S. Sepulveda, M. Bustamante, and A. Cravero, "Identification of Non-Functional Requirements for Electronic Voting Systems: A Systematic Mapping," *IEEE Lat. Am. Trans.*, vol. 13, no. 5, pp. 1577–1583, 2015. https://doi.org/10.1109/TLA.2015.7112018.

[12] M. Lubis, M. Kartiwi, and S. Zulhuda, "Election fraud and privacy related issues: Addressing electoral integrity," *2016 Int. Conf. Informatics Comput. ICIC 2016*, no. Icic, pp. 227–232, 2017. https://doi.org/10.1109/IAC.2016.7905720.

[13] C. J. Lakshmi and S. Kalpana, "Secured and transparent voting system using biometrics," *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 343–350, 2018. https://doi.org/10.1109/ICISC.2018.8399092.

[14] V. Lepassalu, "IT-EKSPERDID ÜLISTATUD E-HÄÄLETUSEST: Valijate hääli saab võltsida ühe näpuliigutusega, süsteem on kui katkine linnamüür!," 2017. [Online]. Available: http://www.pealinn.ee/tagid/koik/it-eksperdid-ulistatud-e-haaletusest-valijate-haali-saab-voltsida-n191310. [Accessed: 25-Apr-2019].

[15] J. Epstein, "Electronic Voting," vol. 10615, 2017. https://doi.org/10.1007/978-3-319-68687-5.

[16] W. Tambaram, "Smart voting," pp. 143–147, 2017.

[17] M. K. Alomari, "Digital divide impact on e-voting adoption in middle eastern country," *2016 11th Int. Conf. Internet Technol. Secur. Trans. ICITST 2016*, pp. 409–412, 2017. https://doi.org/10.1109/ICITST.2016.7856741.

[18] Consejo Nacional Electoral, "ELECCIONES PRESIDENCIALES 2006 - ELECCIONES PRESIDENCIALES 2009," 2009. [Online]. Available: cne.gob.ec/es/estadisticas/publicaciones/category/1285-elecciones-presidenciales-del-ecuador-1948-2017?download=2012:rafael-correa-delgado. [Accessed: 24-Apr-2019].

[19] Consejo Nacional Electoral, "ELECCIONES PRESIDENCIALES 2017," 2017. [Online]. Available: cne.gob.ec/es/estadisticas/publicaciones/category/1285-elecciones-presidenciales-del-ecuador-1948-2017?download=2013:lenin-moreno-garces. [Accessed: 24-Apr-2019].

[20] Tribunal Superior Eleitoral, "Divulgação de Resultados de Eleições," 2018. [Online]. Available: divulga.tse.jus.br/oficial/index.html. [Accessed: 25-Apr-2019].

[21] Registraduría Nacional del Estado Civil, "República de Colombia Elección Presidente y Vicepresidente - 30 de mayo de 2010," 2010. [Online]. Available:

https://www.registraduria.gov.co/elecciones_anteriores/2010PR/escrutinio.php. [Accessed: 24-Apr-2019].
[22] Registraduría Nacional del Estado Civil, "República de Colombia Elección Presidente y Vicepresidente - Segunda Vuelta - 20 de junio de 2010," 2010. [Online]. Available: https://www.registraduria.gov.co/elecciones_anteriores/2010PR2/escrutinio.php#. [Accessed: 24-Apr-2019].
[23] Registraduría Nacional del Estado Civil, "Escrutinio Presidente 1ra Vuelta," 2014. [Online]. Available: https://elecciones.registraduria.gov.co/esc_pre1v_2014/.
[24] Registraduría Nacional del Estado Civil, "Escrutinio Presidente 2da Vuelta," 2014. [Online]. Available: https://elecciones.registraduria.gov.co/esc_pre2v_2014/.
[25] Registraduría Nacional del Estado Civil, "ELECCIÓN DE PRESIDENTE 2018," 2018. [Online]. Available: http://elecciones1.registraduria.gov.co/esc_pre_1v_2018/. [Accessed: 24-May-2019].
[26] Registraduría Nacional del Estado Civil, "ELECCIÓN DE PRESIDENTE SEGUNDA VUELTA 2018," 2018. [Online]. Available: http://elecciones1.registraduria.gov.co/esc_pre_2v_2018/. [Accessed: 24-Apr-2019].
[27] Tribunal Superior Eleitoral, "Estatísticas e Resultados da Eleição," 2010. [Online]. Available: http://www.tse.jus.br/hotsites/estatistica2010/est_resultados.html. [Accessed: 25-Apr-2019].
[28] Tribunal Superior Eleitoral, "Informações e sobre as eleições 2014 Dados Estatísticos," resultados presidentes, 2014.
[29] INFOBAE, "Salta vota para elegir concejales, diputados y senadores provinciales," 2013. [Online]. Available: https://www.infobae.com/2013/11/10/1522618-salta-vota-elegir-concejales-diputados-y-senadores-provinciales/. [Accessed: 26-Apr-2019].
[30] El Comercio, "Se ejerció el voto electrónico en la cárcel de Cuenca," 2014. [Online]. Available: https://www.elcomercio.com/actualidad/politica/se-ejercio-voto-electronico-carcel.html. [Accessed: 25-Apr-2019].
[31] El Comercio, "Mi Perú usa el voto electrónico para elegir a su primer alcalde," 2015. [Online]. Available: https://elcomercio.pe/lima/mi-peru-voto-electronico-elegir-primer-alcalde-247817.
[32] EL UNIVERSAL, "Confirma INE que no habrá voto por Internet en 2018," 2017. [Online]. Available: https://www.eluniversal.com.mx/articulo/nacion/politica/2017/03/15/confirma-ine-que-no-habra-voto-por-internet-en-2018. [Accessed: 25-Apr-2019].
[33] El Mercurio, "Revive el 'En tiempo real' de la elección que terminó con Bolsonaro presidente de Brasil.," 2018. [Online]. Available: https://www.emol.com/noticias/Internacional/2018/10/28/925529/Eleccion-presidencial-en-Brasil.html%0A. [Accessed: 25-Apr-2019].
[34] EL TIEMPO, "Iván Duque es el nuevo Presidente: reviva aquí la jornada." [Online]. Available: https://www.eltiempo.com/elecciones-colombia-2018/presidenciales/resultados-elecciones-presidenciales-2018-colombia-segunda-vuelta-231720. [Accessed: 26-Apr-2019].
[35] EL UNIVERSO, "CNE declara ganador de presidencia de Ecuador a Lenín Moreno, con 99,65 % de votos escrutados.," 2017. [Online]. Available: https://www.eluniverso.com/noticias/2017/04/04/nota/6123701/9965-votos-cne-proclama-ganador-presidencia-ecuador-lenin-moreno. [Accessed: 25-Apr-2019].
[36] ASAMBLEA NACIONAL, "LEY ORGANICA ELECTORAL, CODIGO DE LA DEMOCRACIA," 2009. [Online]. Available: www.lexis.com.ec/wp-content/uploads/2018/07/LI-LEY-ORGANICA-ELECTORAL-CODIGO-DE-LA-DEMOCRACIA.pdf.
[37] Tribunal Superior Eleitoral, "Código Eleitoral - Lei nº 4.737, de 15 de julho de 1965," 1965. [Online]. Available: http://www.tse.jus.br/legislacao/codigo-eleitoral/codigo-eleitoral-1/codigo-eleitoral-lei-nb0-4.737-de-15-de-julho-de-1965. [Accessed: 27-Apr-2019].
[38] C. Pring, *Las personas y la corrupción: América Latina y El Caribe*. 2017. https://doi.org/978-3-96076-062-7.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 117-125 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Cognitive Systems for Monitoring: Architectural View

Alexander Vodyaho[1], Evgeniy Postnikov[2], Alexander Ekalo[2], Vasiliy Osipov[3], Nataly Zhukova[3], Michael Chervontsev[*,2]

[1]*Saint Petersburg Electrotechnical University, Saint-Petersburg, 197022, Russia*

[2]*Research and Engineering Center of Saint Petersburg Electro-Technical University, Saint-Petersburg, 197022, Russia*

[3]*Saint Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, Saint-Petersburg, 197022, Russia*

ARTICLE INFO

Article history:
Received: 22 January, 2019
Accepted: 08 May, 2019
Online: 24 May, 2019

Keywords:
Cognitive Systems,
Cognitive Monitoring System,
Architectural Cognitive Approach
Design of Monitoring Systems

ABSTRACT

The modern stage of information technology development is characterized by the acute need to use cognitive technologies for the solution of tasks of practice, in particular, in the sphere of real world objects monitoring and management. Practical usage of cognitive systems of monitoring is significantly limited to two factors now: operation at the level of knowledge leads to loss of speed, and the high complexity of software solutions leads to increase in cost of development. In order to solve these problems it is suggested to use architectural cognitive approach to design of systems of monitoring. In the article the concept of cognitive monitoring is defined. A new approach to creation of cognitive systems of monitoring which functioning is based on use of model of a target object and the model of monitoring system presented in terms of knowledge are proposed. The idea of this approach is generation of the loaded architecture according to these models. The generalized structure of cognitive system of monitoring is given, the concept of the cognitive monitoring machine, which basic elements are the subsystem of creation of models, a subsystem of transformation of models, a subsystem of processing of models, a generator of architecture, a generator of reports and reactions, a repository and a set of adapters is entered. The structure of the platform focused on realization of this approach is described. The example of cognitive monitoring system is given.

1. Introduction

Cognitive information technologies begin to get into everyday life actively. More and more widely such smart devices as smart phones, the smart TV, the smart house, the smart systems of monitoring of a condition of the patient, etc. are used. Fast extension of scope of the Internet of things (IoT) leads to appearance of the heterogeneous network structures, huge by the size, including very large number of the various elements using a large number of various stacks of protocols. At the same time the huge volume of structured and unstructured data is created, and the structure of such systems permanently changes. Efficiency of the control of such huge information systems (IS) with permanently changing structure of interconnections and device types significantly depends on the evaluable information about the IS state, i.e. the solution of the task of monitoring. Existing systems for monitoring are not always applicable in the considered conditions. Essentially new paradigms of monitoring are strongly required. They can be built on the base of cognitivity concept.

The term cognition comes from Latin cognito that means I think. In particular, in the Cambridge dictionary the term cognition is defined as rational cerebration. In case of such determination emphasis is placed on how a human studies, remembers and argues, but not on discrete facts. The term cognitivity came from philosophy, and then began to be used also by psychologists. In particular, cognitive psychology studies how a human think. The concept of cognitivity is very closely coupled with a concept of artificial intelligence (AI). In essence, the AI IS allow understand a situation, to study it and based on it to make decisions on those actions which are required to be undertaken. The classical AI systems can solve those problems which can be accurately formulated and programmed. From this point of view the cognitive systems (CS) are capable to realize processing just as it is done by a human, so CS can be considered as extension of traditional AI IS.

*Corresponding Author: Michael Chervontsev, Email: chervontseff@mail.ru

One of the main goals of CS development is minimizing the semantic gap between the human and machine i.e. computer acts as a human. In order to reach this effect a computer must be able to work with knowledge (to mine knowledge, to process knowledge). Modern software platforms as a rule do not suggest effective support for knowledge processing and there are limited number of analyst, architects and programmers who can solve problems in terms of knowledge. So, we have a gap between cognitive architectures and modern software development platforms. Another problem with CS is that their implementation requires logical inference, knowledge processing, etc. This leads to serious problems with the speed of processing in CS and as a result limits the scope of CS usage.

The article suggests a new approach to cognitive monitoring. It is an architectural approach that assumes usage of target system models which can be used for generation scripts for cognitive behavior realization. Models can also be used for generation of "loadable architectures". Loadable architecture is peaces of code to be loaded into processors of distributed cognitive monitoring systems. In the second section of the article, an analysis of modern CS, the levels of their cognitivity and approaches to CS systems implementation are analyzed. In the third section the solved problem is formulated. In the fourth section basic principles of cognitive monitoring systems are considered. Generalized structure and the models of the systems are described in the fifths and the sixths sections. The architectural approach to cognitive monitoring systems development is presented in the section seven. In the last section an example of practical use of the proposed approach for systems of operational management of networks of cable television is given.

2. Modern state of Cognitive computing (CC)

On the idea [1, 2] cognitive systems shall differ essentially from traditional IS, including the AI IS. They shall not be programmed rigidly, and shall study like the human in the course of functioning and communication with people, i.e. over time they shall function more and more effectively [3]. According to this paradigm the CS shall be based not on the algorithms and rules which are a priori put in them, but shall be based preferentially on training [2]. It is expected that usage of cognitive technologies can expand abilities of IS and allow them solve problems which are traditionally solved by human. It can be such problems as planning, reasoning, training, operation with incomplete and doubtful data, etc. [4].

CS can have the following properties:

- Ability to self-learning. CS shall minimally use traditional "hard" programming. They shall acquire permanently knowledge from the environment. At the same time CS shall create and check permanently hypotheses, at the same time learning goes preferentially in the mode "without teacher".
- Ability to adaptation. CS shall have ability to adapt both to specific features of the problems to be solved (content), and to the changing environment (context).
- Dynamism. The dynamism can be defined as ability to work in the real time mode.
- Interactivity – ability to interact with different stakeholders for the purpose of obtaining knowledge from them.
- Ability to process unstructured data, in particular, the data provided in a natural language.
- Ability to scaling. Ability to add auxiliary resources in order to improve performance.
- Ability to manage own structure. Assumes a possibility of self-diagnostics, reconfiguring in case of failure of separate elements, solving problems of performance optimization, security management, etc. [6].

To estimate the level of cognitivity of IS one can use discrete or continuous measures of approximation (maturity model) which, in particular, can be defined in terms of levels of a maturity of the technologies. These levels define what cognitive opportunities and in what order can be added in case of creation of an IS, forming some road map which defines transition from traditional IS to cognitive IS. It is possible to define 2 alternative approaches to delimitation of a cognitive computing: to define them in terms of ability to solve problems just as it is done by a human or in terms of methods of the decision of tasks.

The first approach assumes comparing of behavior of the machine and the human in case of the decision of target tasks. For example, if an expert is not able to define with confidence, who solves a problem: a human or a computer, then it is possible to speak about cognitive computing. It is possible to define, how intensively "human" approaches to problems solving is used. For example, for service oriented IS the level of cognitivity can be defined by means of calculation of what part of services can be considered as cognitive. The problem consists that how the human solves the problem is not always known. It should be noted that for practice such approach is of no use.

The second approach according to which cognitive computing is defined by the used architectural concepts is of bigger practical interest. When using this approach, cognitive computing can be defined as a family of architecture solutions which use models of the environment and model of the system, and these models shall be built in terms of knowledge. Taking into account the definition of the term software architecture [5] where architecture includes the IS development, then for cognitive IS this definition is to be expanded and formulated as a family of architecture which in the process of operation and (or) in the process of development use the models of the environment and model of the IS provided in terms of knowledge. If cognitivity is implemented only at a development stage, then existence of tools for transformation of cognitive solutions for class IS in not cognitive solutions for a concrete solution is required. Such approach can be useful in cases when use of cognitive approaches in the run time mode is not possible because of strict requirements on speed and (or) on available resources.

Implementation of CS is based on existing approaches to creation of the AI systems. In [7] 4 alternative approaches to creation of the AI systems are defined: IS which think as a human, IS which think rationally (reasonably), IS which operate as a human, IS which operate rationally (reasonably). It is necessary to mark that concepts to think or to operate as a human and to operate rationally are different things.

The existing CS use all 4 approaches. In particular, in order to understand how a human solve problems, it would be useful to analyze what types of mistakes are made by a human and by a computer when they solve similar problems. If they make similar mistakes, then they thing in the same way [8]. It is obvious that this approach is more useful to physiologists, but not IT specialists as for last it is important that the IS be able to solve problems with the required characteristics of quality. The more important is that when a human thinks, he or she operates with knowledge. Therefore from creation of the CS as a system which operates with the models in terms of knowledge is more useful for IT practice. The detailed and fresh analysis of the state of the art of CS relating to 2018, one can find in [8].

The CS commonly integrates many technologies and platforms thus it can be considered as integration technology. So CS is closely coupled with MAS. MAS is a multidimensional concept. It can be considered, at least, from 3 points of view: i) from the point of view of the theory, MAS can be considered as an approach to creation of systems of the distributed AI, ii) from the point of view of the application programmer, the MAS can be considered as separate architectural style or as a combination of several architectural styles [9], iii) from the point of view of the system programmer, the MAS is a hardware-software platform.

3. The problem of cognitive monitoring systems creation

One can define monitoring as a process of acquisition data about the current status of a target system (TS), data processing and (or) transformations for presenting needed information to stakeholders in a proper form, particularly to persons who makes decisions (DMP). The process of monitoring can be conceded as a business process which can be created both in statics and in runtime.

The monitoring system (MS) can be defined as systems which realize monitoring process. MS can be realized both as stand alone IS or IS subsystem. In the latter case we have self monitoring. Results of monitoring are used for solving problems of target system (TS) management (diagnostics, self-repair, optimization of performance, reconfiguration, etc. [10], [11]).

One can conceder MS and TS as a single system, or a TS can be considered as a context in which MS operates. The choice of approach depends on specifics of the problem to be solved. If the monitoring subsystem cannot send control information to TS, then it is better to conceder TS as an environment. If MS can manage TS, then we have monitoring control system (MCS). As a rule, TS has connections with other objects and IS. The set, consisting from target object and related objects forms the TS. Both TS and MS may include a program component (business logic) and the infrastructure component (IC). IC is sensors, the executive mechanisms, servers and network equipment, etc. The generalized structure of MS is shown in fig. 1.

TS can be defined as S, EM, EQ, BP, M, D , where: S – set of sensors, EM – a set of the executive mechanisms, EO – a set of elements of the equipment, BP – a set of the business processes which realize business logic, M - a set of signals (messages) which are used for information interchange, L - records in log-files about events like beginning and the end business process, D - sources of signals (data) which are generated in the context of specific business process. Message has a format: BP, T where BP – the identifier of business process, T – time tag, each time tag has a format N, V, T , where N is a parameter name, V – a parameter value, T – a time stamp.

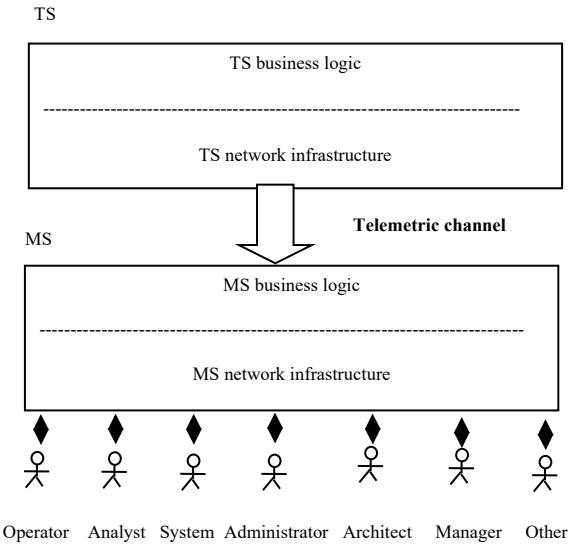


Fig. 1. MS structure

At the top level MS can be classified by three parameters: structure of TS (centralized, distributed), structure of MS (centralized, distributed) and processing method (postponed, on line).

MS can be presented in the form shown in fig. 2.

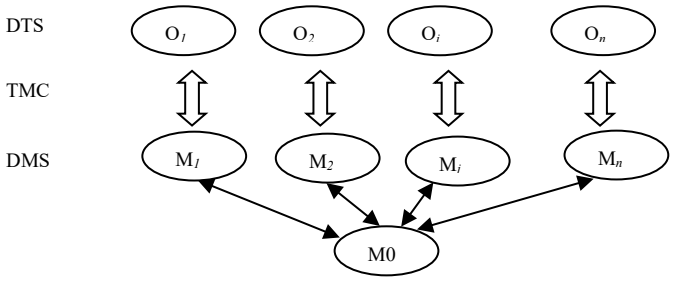


Fig 2. MCS structure

This variant corresponds to a case when TS is a set of distributed objects (DTS). It can be, in particular, mobile objects. MCS also is realized as distributed system (DMS). In fig. 2 M are modules, a tree topology of MS is shown, but also other topology can be used. DTS and DMS can be connected by means of several channels of telemetry (TMC).

It is possible to define the following main problems solved by MS:

- The analysis of telemetry in the postponed mode. On an input of MS the event stream in a format time-name-value arrives. It is necessary to define time point or an event since which deviations are observed and to trace history of development of the situation.
- On line processing of telemetry and TS control. On an input of MS the event stream in a format time-name-value arrives. It is required to check, if the message comes in a proper time and parameters are in norm it means that

situation develops in a proper way. If the events show that the situation has deviations from the sample one, then procedure of correction is to be started. It can be a procedure of change of settings or reconfiguration by means of sending message of name value type.

- Network infrastructure of MS management. If MS is realized as distributed system, the task of monitoring and control of MS own network infrastructure appears. The task is similar to the previous task.
- Debugging of a process of operation with telemetry. Assumes carrying out different experiments at a stage of debugging of procedures of monitoring of real MS.
- Simulation of operation of MS. Carrying out different experiments at a stage of primary debugging of MS.

For solving the enumerated problems monitoring systems should be able to solve such "human" tasks as find and usage of the most informative sources, reasonable distribution of existing resources, observed objects modeling, etc. Also MS should flexibly respond to changes in internal and external monitoring conditions. For this cognitive monitoring systems are required. To build such systems appropriate models and architectures should be developed.

4. Basic principles of cognitive monitoring systems

The Cognitive Monitoring Systems (CMS) can be defined as the monitoring systems realized on the basis of the principles of cognitive computing. It should be noted that most often usage of CMS makes IS cognitive. It, in particular, concerns the systems constructed on the basis of the concept of IoT. Cognitivity is not inherent feature of IoT, but CS are often built on IoT platform. It is possible to speak about 3 aspects of implementation of the cognitive monitoring (CM) concept. In this case it is necessary to answer 3 questions:

- How cognitive processing is organized?
- How the functionality is distributed between the CMS subsystems?
- How the network media is organized?

The answer to the first question consists in the description of what mechanisms of cognitive processing are used in the case of creation of MCS and by means of what mechanisms they are implemented. The answer to the second question consists in describing the CMS structure and platform. A CMS can be realized as a system with the client-server architecture, as a distributed fog system, on the base of multi-agent platform, etc. It should be noted that sensors can be cognitive also. The answer to the third question assumes the description of how the monitoring data communication media is organized. In view of the fact that modern TS more and more often represent as a set of distributed mobile objects, usage of a communication media based on ideas of Cognitive Internet of Things (CIoT) quite evident.

The classification of possible variants of CSM organization is given in fig. 4. The approaches can be classified by 3 main features: from the point of view of used structural and architectural concepts and from the point of view of used models. From the point of view of architectural solutions, it is possible to select 4 basic alternative approaches: use of classical solutions, use of vertical solutions, use of horizontal solutions and use of hybrid solutions. Traditional solutions assume, for example, use of solutions on the basis of SNMP and cognitive processing can be executed on the server. Vertical solutions are multi-layer solutions in which cognitive processing is implemented at several levels. Examples of such approach is fog computing. Horizontal solutions assume one level organization. In this case cognitive processing is implemented by a group of the processors working at one level. As a rule, it is P2P systems. Service oriented architectures and multi-agent solutions can be conceded as examples of horizontal approach. Hybrid solutions are a combination of horizontal and vertical approach. MAS working on the several levels can be conceded as an example of this approach.

CMS assumes use of 2 main types of models: TS model and own model. If the dynamic (compiled) architecture is used, then it is possible not to use models in loadable modules. In CMS only separate model can be used, for example, only TS model. The models can be either static, or dynamic. Static models are generated in the process of the CMS development. Dynamic models can be created and (or) changed in run time.

In the process of CMS development different architectural concepts can be used. In the simplest case it can be the fixed architecture. It can be either an architecture designed manually or the architecture generated from architectural meta models according to an approach which will be described below. When using the virtual environment use of the loaded (on demand) architectures from architecture library is also possible. The most effective and difficult solution is use automatically generated dynamic architecture which can adapt both to content, and to context.

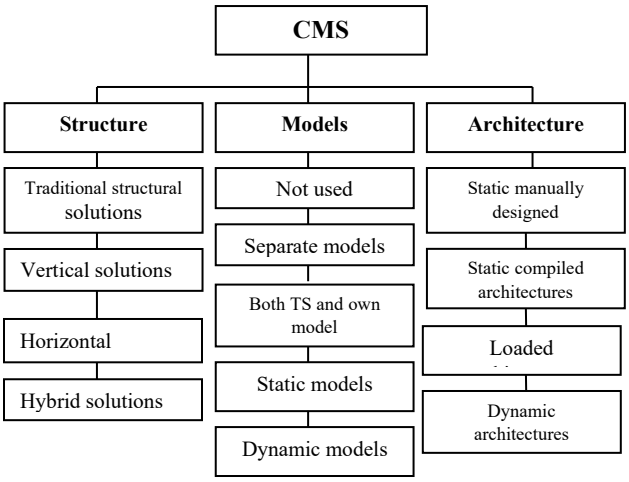


Fig. 3. Classification of approaches to CMS implementation

5. Generalized structure of cognitive monitoring systems

The generalized structure of CMS is shown in fig. 4. The TS can have the built-in sensors which can realize separate elements of cognitive processing, as a rule at the level of signals. The core element of CMS is cognitive machine which is responsible for building required models, their maintenance in actual state, transformation of models and realize decision-making on models. Communication between the cognitive machine and the field of

cognitive sensors can be realized with the help of a cognitive network [8].

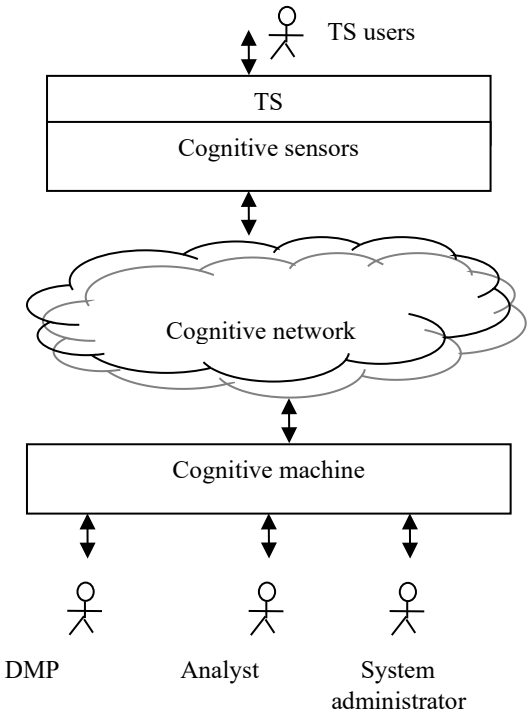


Fig. 4. Generalized structure of CMS

The basis of CMS is the cognitive machine of monitoring (CMM) which can be considered as an architectural framework as the platform or as CMS toolkit depending upon the point of view. CMM consists of a set of services of different level which can be used for creation of concrete CMS. The conceptual structure of CMM is shown in fig. 5. CMM includes 7 subsystems: Model Mining Module (MMM), Model Transformation Module (MTM), Model Processor (MP), Architectures Generator, Generator of Report and Responses, Repository and a set of Adapters.

MMM is a module which is responsible for creation and maintenance of models in actual state. MTM is a module which is responsible for model transformations. PM is a module which is responsible for processing of models, in particular realizes reasoning on models, the generator of architecture can generate loadable modules and is used in CMS with dynamic architecture. The repository is used as storage for models, data, knowledge, scripts, etc. The report generator and responses performs several functions. If a CMS operates in MS mode this module presents information about TS status to DMP. If the CMS operates in MCS mode, then it sends controlling impacts to TS. In the case of use of the dynamic loaded architecture, this module is responsible also for loading of architectural modules in distributed environment of CMS. Adapters are used for information and knowledge search in external sources.

CMM operates in a following way. Data from sensors come to the CMM machine. If the model is unknown or requires specification and (or) verification, then for this purpose MMM used. The standard mode of operation assumes permanent checking of models on correctness and reorganization of models in case of arrival of messages about events. All models are stored in the repository. In CMM main decisions are made as a result of logical inference on knowledge base. For this purpose it is required to mine knowledge from an input data stream and external sources. This procedure realizes MMM which integrate a number of lower level services. MMM can mine knowledge from data and events. Processing in MMM can be realized according to well-known JDL model [12]. MP realizes functions of making different operations with models, by the most part by means of SPARQL [13] requests.

Such solutions are relevant first of all in the cases when TS and (or) a connective network is built as IoT. If the structure of the serviced TS is known and stable then loadable modules can be stored in repository libraries. The use of the loaded (on demand) architecture [14] is relevant, first of all, for fog structures when resources of modules of the intermediate level are restricted.

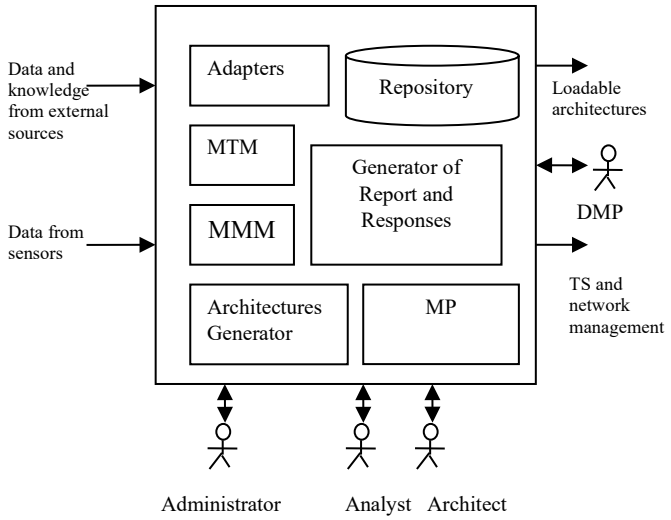


Fig. 5. Cognitive machine of monitoring structure

6. Models and models transformation in cognitive monitoring systems

As it was stated above, cognitive computing and, respectively, cognitive monitoring, assumes the widest use of models, including the models provided in terms of knowledge. Nowadays creation of models in the case of design of the IS undoubtedly is a bottleneck and in the most cases continues to be executed manually by analysts. This activity needs both big expenses and does not exclude errors. If the structure of the IS to be modeled permanently changes, then the task of creation of adequate models becomes even more difficult. Thus, the problem of automatic formation of models is rather actual one.

In relation to CMS models shall provide the solution of the following main problems: estimation of a current status of TS, prediction of TS behavior, to estimation of CMS own condition, prediction own status, synthesis of architectural models from the architectural description. For CMS models in terms of knowledge are of prime interest. First of all, it is ontological models.

Nowadays different methods of receiving models from data and information on events are known. In our context at least 3 technologies are of interest: extraction of knowledge from data (data mining) [15], creation of models of processes from log-files in the form of Petri nets (process mining) [16], creation of automata, in particular, multilevel automata models [17]. Today it

is separate loosely interconnected technologies. It is advisable to consider these technologies as a single technology, which can be called Model Mining (MM). MM can be conceded as umbrella technology and can be determined as MM = { Mod, Met, Tmet, I, R } where Mod – models, Met – methods of a mining of models from data and messages about events, Tmet – methods of transformation of models, I – instruments of models mining of, R - an implementation of models. The MM can be considered as integration technology. Usefulness from introduction of this concept is that it allows: i) to expand a concept of Model Driven Engineering (MDE), ii) gives the chance to look at a problem of automatic synthesis of models integrally, iii) allow understand what elements which are necessary for realization of systems of automatic model synthesis are absent, iv) allow accumulate knowledge in the form of models (model knowledge).

According to the MDE approach to describe a system a stack of models is created. The key concepts of the approach are the following: model (M), meta model, (MM), model wirering (MW) and model transformation (MT). Representation of a set of models in the form of a stack needs to determine the procedures of transition from model of the top level to models of the bottom level and vise verse. Process of model transformation is a process which defines how it is possible to receive target model from one or several initial models. Models can be transformed either horizontally, or vertically. In the first case the model is transformed to model which belongs to the same level. In the second case the transformed model is model of higher or lower level. The special type of models - VM are used for binding of models of different types. The basic goal of these models consists in establishment of compliance between separate elements of models (low level binding) [18].

In relation to CSM both horizontal, and vertical transformations of models are of practical interest. Not always it is possible to realize transformations procedures automatically and to receive needed model. For example, a service allows receive results of a mining in the form of the multi-level finite state automata, but MP can work with ontology, and for generation of executable code UML description is needed.

7. Architectural approach to cognitive monitoring systems development

In the process of CMS development an architect faces at least with two problems.

The first problem is that very often MS must operate in real time. In relation to CMS the main problem is that implementation of intellectual elements of behavior requires manipulation with knowledge. Manipulations with knowledge, assumes implementations of a logical inference, i.e. execution of a large number of machine operations which cannot always be executed in parallel mode that makes difficult to use CMS not only in hard real-time systems, but also in online IS with moderate requirements to speed of operation. Besides, it is not always possible to predict time required for realization of a logical inference precisely. It is possible to solve this problem due to increase of performance of the used servers and (or) usage of parallel algorithms. All these approaches, at least, when using the modern platforms, give effect only in the case of very moderate requirements to real time.

The second problem is that CMS realization assumes implementation of mechanisms of knowledge processing which are most often provided in the form of ontology. It is necessary to mark that development of such type of IS requires from the analyst and the programmer rather high qualification and experience or it is necessary to invite knowledge engineers, i.e. the process of CMS development becomes rather expensive.

For solving these problems architectural cognitive approach to CMS development is suggested. The idea of this approach consists in development of the domain oriented platform for CMS which includes the architectural framework (AF) and tools for development on its basis of concrete CMS. AF includes a set of services. The generalized structure of the platform is shown in fig. 6.

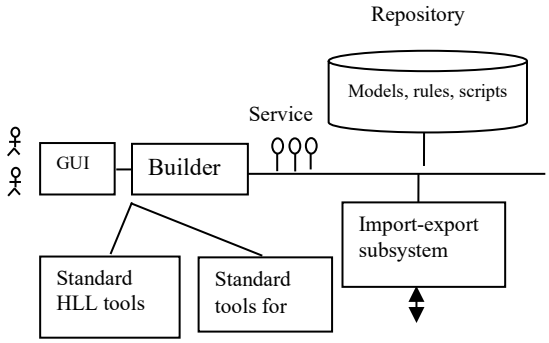


Fig. 6. Cognitive machine of monitoring structure

The platform represents a set of tools for architectural development of CMS. The platform consists of the following elements: a repository for models, rules, scripts and data storage, a data, information and knowledge import export subsystem, a set of services. Besides, the platform encapsulates toolkit for CMS architectural development – builder, with the help of which one can build models and business processes of monitoring information processing. Besides it standard high level languages (HLL) design tools and object-oriented modeling (OOM) can be used. The basis of the platform are services, such as: services of extraction rules from data (data mining), services of automatic creation of models (model mining), services of models transformation, services of generation of modules of the loaded architecture, models finding services, engines services, services of access to data and knowledge models access, etc.

The platform can support model transformation, i.e. generation of automata or neuron network solutions from cognitive models. For example, when we have strict requirements on time, use of automata may help to solve problem. In the case of context is changed then new automate can be generated and loaded as module. For loading of modules it is possible to use separate temporal slots or it can be done when model mining subsystem defines the model is to be corrected. In this case one can speak about the adaptive (agile) architecture in a sense [19].

In general, the described above approach to use of models can be defined as Cognitive Supported Architecture (CSA). Perspectives of practical application of this approach are defined by accessibility of services of transformation which allow use instead of heavy models in the form of knowledge other more light models such as automata machines, tables, decision trees etc which

allow receive higher speed of computations. Such approach allows use the platform as a modeling and development tool. The offered approach is, first of all, the integration approach directed to integration of already available technologies and solutions. Now the platform is in a development stage.

8. Use case. CMS of resources on networks of cable television

Let us conceder as an example of practical use of the described earlier approach for system of operational management of networks of cable television.

Generalized structure. The Systems of a Cable Digital Television (SCDT) are one the traditional directions of telecommunication technologies. The modern SCDT are sophisticated distributed IS with hundreds of thousands of subscribers, which include the client side and sever side equipment. The server side equipment is installed in Data Acquisition and Processing Centers (DAPC) of operators. Server side equipment is, as a rule, powerful servers clusters which realize procedures of client side equipment monitoring and measurements. The network is divided to segments. Each segment has its own segment server (SS). The receiver of a digital television (the TV-tuner, the Set-top box, STB) which provides basic and expanded functionality of the TV set is installed on the side of the subscriber.

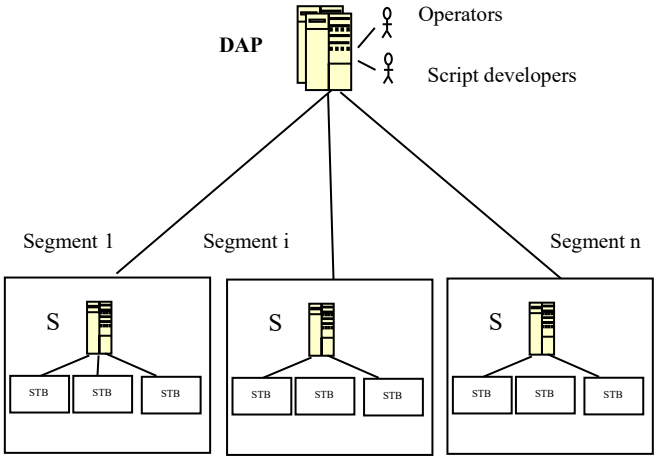


Fig. 7. Structure of SCDT

For management of SCDT the systems of Operation Management System (OMS) which standard structure is shown in fig. 7 are traditionally used. Command centers realize monitoring of STB of a network for the purpose of determination of their current status and provide technical support. Besides, the solving this task, OMS realizes functions of information distribution and management of computational resources.

OMS functionality. Standard OMS solves 2 groups of tasks: the tasks of technical support of the SCDT and the task of statistics receiving for the benefit of analytical and marketing departments; the continuous monitoring the status of services, operational informing the operator in case failures, registration of originating errors, their identification, localization; assessment of possible ways of elimination, operational debugging, detection of origins of problem situations, the analysis of a situation in the network in general or in separate subnets, detection of dependences between originating errors on different components, dependences of origin of errors and network condition, actions of the user, the forecast of origin of errors, early diagnostics and preventing of appearance of errors, etc. The main OMS task is provision to operators of information on a status of SCDT, detection and localization of the failures.

The standard requirements to OMS. The main requirements to the modern and perspective OMS are about a possibility of operation with big data, flexible logic of operation (a possibility of adaptation in real time to an environment status, actions of users), the minimum requirements to technical characteristics of STB and parameters of a network, high rates of readiness. At the same time OMS shall be updated quickly taking into account the new realized logic; new needs of analysts for solving analytical tasks.

Typical problems of OMS development and usage. In the process of OMS operation the high dynamism of the environment takes place. Work loading of a network and devices is permanently changing depending on behavior of users and operability of technical means. In this situation it is very difficult to forecast the SCDT future status. As a result, the situation when additional loading from a monitoring system leads to system failure is possible. In the course of OMS functioning there is always a danger of origin of effects of "avalanches" when in case of origin of a malfunction there is the avalanche increase in traffic between the local server and STB caused by repeated attempts of receiving and sending data.

OMS developers often face the following problems: i) the long and sophisticated cycle of software debugging; the cycle of testing and delivery of the new version of a software takes about 6 months that makes impossible to change program code often because of complexity of process and high cost of release of new versions; ii) the logic which is to be realized by the STB is defined by many factors such as network parameters, structure of a network, specific tasks of the concrete operator; as a result, development of many modifications of the systems is required, at the same time there are the limited number of systems of each type is used.

Possible approaches to solving the problem of SCDT monitoring. The task of SCDT monitoring can be considered as the task of establishment of causes and effect relationships on a set of events where the reason is a failure, and the result is an event message which can be either message from STB to the server, or creation of a monitoring artifact created on a server side. The problem is that the structure of SCDT is permanently changing because of appearance of subscribers, appearance of new equipment, etc.

Appearance of the new STB types which, on the one hand, have new functional capabilities and on the other hand, they have new architectural features such as a modified set of statuses, the new message formats, and it is necessary to develop big number of new drivers. Really, in case of appearance of the new STB type it is necessary to develop new diagnostic scripts, thus, a new role appears – a scripts developer that, naturally, leads to increase in the total cost of ownership (TCO).

From the point of view of the SCDT owner is extremely desirable to automate the procedure of scripts development, or, at

least, radically simplify it. One of possible approaches to this problem solution is use the domain specific languages (DSL) [20] for script development. This approach allows to simplify the process but not to automate it.

Ontological approach. The main idea of ontological approach is that the diagnostic script is created on the basis of ontological descriptions of elements, in particular, of STB. When ontological approach is used then when it is necessary to add new element type to SCDT the script developer makes its ontological description and adds it to working ontology. The ontological description is used for generation of automata model.

Ontological approach can be considered as a superstructure over other approaches. It, first of all, concerns automata approach. The ontology in this case can be conceded as an instrument of dynamic structures description. Ontological approach can be also used together with model approach. In this case for scripts generation instead of object models ontological models are used. It is obvious that the systems which use build-in ontology cannot show high performance but for this case it is not very important. Use of ontological approach, on the one hand, simplifies operation of the analyst because it allows build ontological descriptions of elements on the basis of existing,

but on the other hand, requires from the analyst to have the skills of ontologies usage.

Diagnostics system operation algorithm. The generalized algorithm of system operation of monitoring includes 4 steps:

- Step 1. Algorithm execution is initiated in case of origin of an erratic situation in STB
- Step 2. Parameters which values are required for bug fixing in operation of a receiver are defined.
- Step 3. Sending to STB of the requests for obtaining parameters with the help of standard commands is carried out. A command execution result is information messages with parameter values or log-file. The analysis of the received values is made; if necessary, additional requests are made.
- Step 4. On the base of received data the way of elimination of failure is defined and implemented.

In article [21] the formalized algorithm of diagnostics and elimination of erratic situations in STB is described. The mathematical apparatus for solving the problem of automatic scripts synthesis, models and indices of efficiency are suggested. The stage of determination of parameters of a status of STB which are necessary for identification, localization and elimination of an erratic situation is in details considered. The algorithm of automatic synthesis of programs of additional parameters acquisition from receivers is offered. The algorithm is constructed on the basis of conditional finite state operational automata. For solving the problem of diagnostics the models which describe receivers, the environment in which they function, actions of the users are used.

The set of models includes the following models: function models which describe the functions provided to users, model of program components which describe the software of a receiver, models of parameters of components which determine parameters of a status of software components. The set of models includes also the following main models: i) the model of platform-independent parameters, ii) models of platform-dependent parameters which are defined by STB model type, iii) context-sensitive models (model of the environment and methods of use of STB) which define a context of formation of algorithms of monitoring, iv) model of a data communication network, v) behavior model of the user, vi)status model of STB and their components, vii) status models of STB describe a receiver status for the given context according to the volume of the supported user functionality, viii) model of erratic situations. All listed above models in details are described in [20].

Author's experience in development and use of the described system shows that despite limited cognitive opportunities, use of the cognitive approach based on application of models gives a certain positive effect from the point of view of minimization of TCO. System testing shows that it allows create rather effective diagnostic scripts.

It is necessary to mark that described above OMS realizes rather low level of a maturity from the point of view of cognitivity, in particular now self learning mechanisms are not yet realized. Authors try to do the best to realize this facility in the next system version.

Conclusion

It is possible to claim that today cognitive technologies reach the sufficient level of a maturity for use at design of real systems. Not least it concerns the CS of monitoring. In many cases existence of a cognitive subsystem of monitoring, allows to conceder the IS as a CS.

But there are a number of moments which prevents a wide use of CS. The first problem is that for creation CS and respectively cognitive monitoring needs it is necessary to invite high skill developers who can work with knowledge based IS. For many small companies very often it is too expensive. Suggested framework can be used by architects who have limited experience in the field of knowledge based IS development. The second problem under consideration is a problem of CS operation in real time mode. For solving this problem it is suggested to use loadable generated architectures.

The further direction of research will be pointed on integration of suggested architectural framework with existing architectural framework and further investigations in the field of automatic architectures generation.

References:

[1] Kelly, J.E.: Computing, cognition, and the future of knowing. IBM Research. Oct 2015 http://www.research.ibm.com/software/IBMResearch/multimedia/Computing_Cognition_WhitePaper.pdf
[2] Kelley, J.: Smart Machines: IBM's Watson and the Era of Cognitive Computing. Columbia Business School Publishing. Sep 2013 *Kelly III, Dr. John (2015).* "Computing, cognition and the future of knowing" *(PDF). IBM Research: Cognitive Computing. IBM Corporation*. Retrieved February 9, 2016
[3] Balani, N.Cognitive IoT (2015) http://navveenbalani.com/
[4] Schatsky, D., Muraskin, C., Gurumurthy, R.: Cognitive technologies: the real opportunities for business. Deloitte review, no. 16, pp. 115–129 (2015)
[5] International Standard ISO/IEC/IEEE 42010 Systems and software engineering —Architecture description http://www.iso.org

[6] How is cognitive computing different from big data and NLP? http:/www.coseer.com
[7] Russell, S, Norvig, P. Artificial Intelligence: A Modern Approach, 3rd Edition, 2010
[8] Sangaiah A. K., Thangavelu A., Sundaram V. M. S. Cognitive Computing for Big Data Systems Over IoT. Frameworks, Tools and Applications. Cham, Switzerland. Springer International Publishing AG -2018. — 375 p.
[9] Bass, L., Software Architecture in Practice. 3rd ed. / L Bass., P.Clements, R. Kazman. — Upper Saddle River, NJ.: Addison-Wesley. 2013. — 661 p.
[10] Okhtilev M. Yu., Sokolov B.V., Yusupov R.M. Intelligent technologies of status monitoring and control of the structure of complex technical objects. M. Nauka.-2005. — 291 p. (In Russian)
[11] Beliaev S.A., Vasiliev A.V., Kudriakov S.A. The monitoring of information trends system's architecture based on the free software // Programmnye produkty i sistemy (SOFTWARE&SYSTEMS) №4, 2016, pp.85-88 (In Russian)
[12] Blasch E., Bosse E, Lambert D. High-Level Information Fusion Management and System Design, Artech House Publishers, Norwood, MA. 2012
[13] Gasevic D., Djuric D., Devedzic V. Model Driven Architecture and Ontology Development. Springer-Verlag, 2006.
[14] Sommerville, I., Software Engineering. — Boston, Massachusetts, Addison-Wesley, 2011. -773 p.
[15] Zaki M., MeiraW., Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press, 2014
[16] van der Aalst W. Process Mining. Data Science in Action. 2nd ed., Berlin Heidelberg. Springer-Verlag. —2016
[17] Osipov V. Yu. Automatic Synthesis of Action Programs for Intelligent Robots // Program. Comput. Software. 42 (3), 2016. Pp. 155 – 160.
[18] Zivin B. E., Jouault J, and Valduriez P., On the Need for Megamodels. https://scinapse.io/papers/195085068
[19] Babar M. A., Brown A. W., Mistrik I. Agile Software Architecture. Waltham, MA: Elsevier Inc. 2014. -- 392 p.
[20] Kelly S. Tolvanen J. Domain-Specific Modeling: Enabling Full Code Generation. — John Wiley & Sons, 2008. — 340 p.
[21] Vodyaho A.I., Mustafin N.G., Zhukov N.A. Ontological approach to creation of monitoring systems of resources on networks of cable television, Izvestiya SPbGETU “LETI” 2017, No. 2, pp. 29-38 (In Russian)

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 126-137 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Early Detection of Alzheimer's Using Digital Image Processing Through Iridology, An Alternative Method

Fernando Hernández[1], Roberto Vega[1], Freddy Tapia[1,*], Derlin Morocho[2], Walter Fuertes[1]

[1]*Computer Science Department, Universidad de las Fuerzas Armadas ESPE, 170501, Ecuador*

[2]*Electrical and Electronics Department, Universidad de las Fuerzas Armadas ESPE, 170501, Ecuador*

ARTICLE INFO

Article history:
Received: 08 January, 2019
Accepted: 08 May, 2019
Online: 26 May, 2019

Keywords:
Alzheimer's
Iridology
Digital Image Processing
Machine Learning
Newton's Interpolation
Alternative medicine

ABSTRACT

Over the years, alternative medicine has carried out numerous systematic studies related to the ophthalmological field, specifically the eye. Studies that propose as support of analysis to the Iridology, which studies the alterations of the iris in correspondence with the organs of the human body. Thus, it presents the opportunity to investigate neuronal diseases related to certain alterations that occur in the iris such as Alzheimer's, in this way generate an alternative method for the early detection of Alzheimer's based on iridology and based on the digital processing of images. Addressing this problem, a systematic literature review was carried out to evaluate the characteristics of the iris, which allow to establish criteria to determine if a person could have a problem associated with Alzheimer's. With all this, the prototype is evaluated with an image data of N individuals who have and do not have such a clinical picture. In addition, within the evaluation is considered the extraction of several parameters that allow detecting anomalies in the iris, at the same time the best range of approximation in the pixelated will be analyzed, for the location of the brain area in the iris applying Newton's interpolation, which will allow the modeling of a function that meets the best approximation criteria. The criteria for evaluation are presented in terms of precision, sensitivity and predictive capacity, with which the behavior of the characteristics and the impact that is generated on the performance of the proposed system can be understood. Also for a technical support the algorithm was exposed to an evaluation of results estimation and this process generated an adequate radius for the correct extraction of the segment of the iris, where the whole study will be carried out to determine and validate whether the person has Alzheimer's or not.

1. Introduction

The present work is an extension of the work originally presented at the CISTI'2018 - 13th Iberian Conference on Information Systems and Technologies, held on June 13 and 16, 2018, in Caceres, Spain [1].

The objective of expanding this work was to address in a more detailed way the alternative method proposed to detect Alzheimer's as a neuronal disorder in the early stage, which affect society today. The starting point for this type of pathological analysis is to determine patterns associated with the iris of the eye, through the use of digital image processing to obtain and issue criteria related to alternative diagnosis. It is worth mentioning that the study is based on iridology that is based on alternative medicine as a support for its scientific validation.

The statistics show that Alzheimer's covers 60% of all neuronal diseases, which has awakened an alert and a great medical challenge around the world, in public health issues. Alzheimer's, being considered one of the most complex neurodegenerative disorders, studies several factors that influence the possible diagnosis, which in itself is already complicated. Because the symptoms are notorious and occur in very advanced stages (age and progression of the disease) and severe disease, making the treatment useless because it is irreversible [2].

Alternative medicine uses techniques based on iridology, as a strategy for the study of pathologies based on the iris of the eye, that is, the colored membrane of the eye [3]. The iridology relates

*Freddy Tapia, Sangolquí-Ecuador, +593998926648, fmtapia@espe.edu.ec

the iris to the organs of the body, and links these alterations based on the analysis of characteristics among which we highlight: color, texture, nerve rings, gaps, inflammations, among others. The result obtained is a diagnosis of non-scientific diseases [4].

The iris on its side is a flat and ring-shaped membrane, it contains pigments that provide color to the eyes, and is composed of connective tissue cells with muscle fibers that control the size of the pupil [5]. In addition, the iris develops from the formation of the fetus, and has the peculiarity of shrinking the pupil if there is too much light and expanding it if there is not enough, a function controlled by the brain [6]. These considerations are fundamental for the development of the present study.

ISO / IEC 29794-6 establishes the methods used to quantify the quality of iris images, as well as software and hardware regulatory requirements. This helps regulate the samples during the analysis; However, medical conditions that could affect the study of the iris, such as: excessive dilation or constriction of the pupil, congenital diseases, surgical procedures and many other pathologies must be considered. All this affecting directly or indirectly the results obtained by the iris recognition systems [7].

Of the several studies reviewed for the realization of this research, many highlight the importance of the eye to assess medical disorders associated with problems such as: alcoholism, smoking, hypertension, diabetes and obesity [8]; for this case, the relationship between the study of the iris and neurological disorders such as Alzheimer's will also be determined.

The advantage of the iris is that due to its anatomical structure it varies between each individual, making it difficult to falsify it. That is the importance as a biometric authentication method, using pattern recognition based on images of the iris of the eye [5]. In addition, with the support of sophisticated camera systems and computer-aided image processing, they capture high-quality samples that are converted and stored as digital templates. With its objective characteristics that have greater precision, reliability and high efficiency in the results [5].

On the other hand, iridology has been used for years in alternative medicine as a diagnostic method for diseases, which is based on establishing related patterns between the organs of the body and the iris of the eye; therefore, it allows a mathematical representation of the iris allowing a positive and unequivocal identification of the individuals analyzed [6]. Therefore, applying a neural network will help validate the alternative method proposed and the results of the diagnosis, gathering the largest amount of information stored in the iris of the eyes (Fig. 1).

The remainder of the article has been organized as follows:

Section I addresses a systematic review of literature and related work with respect to the central theme and other works that contribute to the understanding of the field of study. Section II deals with the Methodology used to obtain the characteristics, as well as analyzing the learning algorithms and classifiers that adapt better. Section III presents the results obtained, as a result of qualitative exploratory analysis, in addition to the use of mathematical models that validate the proposed model and software. Finally, Section IV presents studies related to this research, as well as the incorporation of some ideas and future lines of research.

1.1. Literary Review and Functionality

The reviewed works focus much of their attention on the enormous utility of digital image processing, since they allow to determine certain health disorders related to vision. For this, digital templates play a major role in determining the characteristics of location and removal of the eyes, which are necessary to determine diagnoses and evaluate them [9]. These studies also reveal that during the analysis stage for the recognition of the iris, aspects such as surgeries, cataracts, among others must be considered. All this, compared with eyes without any problem or pre-existing surgery, as it influences the performance of recognition, with a non-coincidence rate equal to 11% percent [10].

Table 1: Comparative table of work related to the proposal made

Related Work	Approach	Results
New methods of verification and identification using iris patterns	Proposal of a new method for the verification and recognition of identity with the help of the modalities of the iris of the eyes	The use of light in process and performance techniques can recognize the inner limit of the iris.
Various iris recognition algorithms for biometric identification: a review	A vision of the different methods of iris recognition.	Daugman algorithm offers maximum accuracy and minimum error rates among the four algorithms
Assessment of iris recognition reliability for eyes affected by Ocular pathologies	More exhaustive analysis of the performance of iris recognition in the presence of various eye diseases	Some different diseases do not depend on the taxonomy of the disease, but on the type of damage inflicted on the eye
Review of Image Processing and Machine Learning Techniques for Eye Disease Detection and Classification	Review of the processing of medical images and automatic learning techniques to detect and classify images of eye diseases for the recognition of diseases	The detection and recognition of eye diseases can be achieved through the proposed system with the use of image processing techniques and data mining
The application of retinal fundus camera imaging in dementia: A systematic review	Investigate cerebrovascular disease and the contribution of microvascular disease to dementia with camera-background images.	Ten studies evaluating the properties of the retina in dementia were included. Retinopathy was more prevalent in dementia.
Early Detection of Alzheimer's using digital image processing through Iridology, an Alternative method	It aims to show that there are alternative methods to detect certain neuronal disorders through the digital processing of the images in such a way that criteria related to the alternative diagnosis on said pathology can be issued.	The potential of the image analysis and the quality of the data depend on the adjustment of the criteria and the decision making when issuing the diagnosis.

Other works present different mechanisms of analysis to locate and segment the ocular sector, but in the case of the present study it is intended to contribute with the analysis of the different characteristics extracted from the human eye; and how they influence the determination of certain pathologies such as Alzheimer's [9] [11]. Alzheimer's disease is a progressive disorder that causes the brain cells to degenerate and die, this being the most common form of senile dementia accompanied by a continuous decrease in the abilities of: thinking, behavior and social influencing in a negative way in the life of the individual. Alzheimer's is composed of two systemic cerebrovascular factors, such as: diabetes and hypertension [2]. In addition, recent studies reveal that part of the population suffer from some type of eye disease such as: Macular Degeneration, Cataracts, Retinitis, Diabetic Macular Edema, Ocular Hypertension, Retinal Detachment, Glaucoma and Neurodegenerative Alteration, among others [12].

From the related works, the contribution for the proposed research can be evidenced, either by its approach and techniques which would allow to improve the results and the use of associated tools. In addition to considering the criteria that were taken into account to contrast in some way the results obtained.

2. Materials and methods

The data was analyzed and published with the help of a Foundation in the city of Quito, which collaborated with the staff for the collection of the samples. In this way, a cross-validation was carried out with the information obtained, to perform the processing of the images and subsequently obtain the characteristics of the iris of the people.

From the literature reviewed, conclusive points can be established regarding the presence of a state of alteration in any organ of the body, seeing this reflected in the iris of the human eye; and this condition is contrasted with templates of iridology to verify if it suffers from an alteration corresponding to the area where said alteration is located. Figure 1 illustrates a map or template of iridology, which allows to address the various diseases related to the organs of the body, according to alternative medicine. For the present study, the region where the alteration and anomalies that cause Alzheimer's occur is found in the brain area.

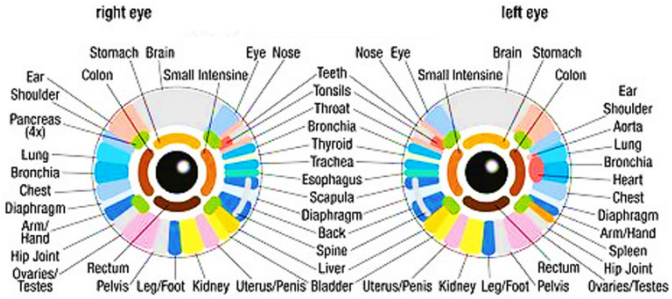


Figure 1. Iridology of the eyes map (source: Shutterstock)

The first thing that is done is to detect the color changes in the area of the brain (Fig. 1), then the relevant parameters for the analysis and subsequent evaluation are extracted. For this case, six parameters are extracted (Table 4), which are relevant for the recognition of the pathology. Then it is necessary that the system learns to establish based on criteria a possible diagnosis automatically, but through a learning, whether supervised or unsupervised applying the theory of Artificial Intelligence and other aspects. Figure 2 shows the general process by which can be described as a process diagnosis is described.

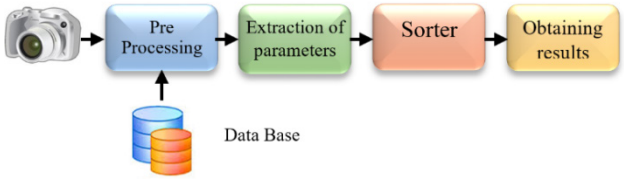


Figure 2. Proposed process for the detection and diagnosis of Alzheimer's

2.1. Process Description

The process begins when the system acquires an image of the iris, which can be obtained from an existing camera or database (for example, Iriso). For the present work we use the images that were obtained by the Foundation, it is worth mentioning that these are in JPG format due to pixelated themes when processing the image (Fig. 3). During the pre-processing of the image, some steps are carried out to determine the characteristics that will be used later.

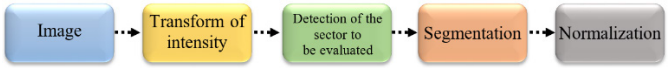


Figure 3. Algorithm of image pre-processing

In the next step, the detection of the inner (iris) and external (pupil) edges is performed, using the Hough Transform, which basically detects the circular shape of the sectors of the eye (Fig. 4). Then, the image is segmented applying discontinuity to divide the image where sudden changes of gray appear (Fig. 5), to establish the area of the brain that will be analyzed. An additional process that is included is the Logarithmic Intensity Transformation applied to the image, allowing to clarify the affected area.

Once the region is established, applying the Hough Transform, the LaGrange equations are applied to determine which would be the best curve that approximates the required characteristics. Then, it can be concluded that the approximations for the area of the pupil will be verified from the circumferences of 20, 30 and 40 pixels around it; and in the same way a second circumference for the interior area to the iris, which will be verified from the circumferences of 90, 100 and 110 pixels around it. The entire process will lead to obtain the corresponding region of the brain area and, this whole process aims to obtain the best sensitive area of analysis and the results are biased as little as possible.

Figure 4. Hough transform for detection of pupil and iris

Tables 2 and 3 show the approximation processes proposed for obtaining the most suitable area of the brain, which will later be used during the testing stage and the results obtained. It must be borne in mind that the ranges identified have the purpose of covering the first sector corresponding to the brain as shown in Figure 6.

Figure 5. Process of segmentation of the image

Figure 6. Sector of analysis corresponding to the area of the brain

Table 2. Process of obtaining the brain area (second ring)

Approach to the brain area (IRIS)	
Pupil Range	Average pupil pixels
[30–70]	40
[40–70]	30
[50–70]	20

Table 3. Process of obtaining the brain area (third ring)

Approach to the brain area (IRIS)	
Pupil Range	Average pupil pixels
[90–200]	110
[100–200]	100
[110-200]	90

The process of segmentation of the image is evaluated, verifying if it was fulfilled without problems and that oblige to adjust the process. Otherwise, it proceeds to transform to fixed dimensions that allow comparing and extracting the standardized characteristics, which in this case are white lines identified in the image that is in gray scale (Figure 7).

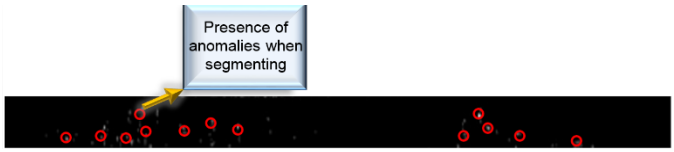


Figure 7. Normalized image for feature extraction

At this point it is intended that the system is able to feed their learning, taking as input the parameters or anomalies within the area of the brain. During the process it was possible to identify and extract 6 parameters (Table 4) as a basis for their evaluation, of which the functionality is detailed below [13]: (1) Average intensity, is the depth of the gray tone determined by the number of bits used to define each pixel, that is, the deeper it proportionally will fulfill the quantity of tones that can be represented in an image. (2) The average contrast of the standard deviation is a range of brightness or average variant in the bitmap, which is evidenced when there is a propagation of the gray levels and the texture does not have a direct proportionality with the softness. (3) The softness, is responsible for minimizing variations or intensity differentials between the closest pixels, eliminating noise or details, which includes techniques that provide a smoother appearance of the image. (4) Third moment, measures the bias or asymmetry of the central periods normalized in a histogram, from which they are derived to a set of moments: invariants, translations, rotations and changes of scale. (5) Uniformity of the histogram, verify if there is correspondence between the number of pixels and the gray level of a monochromatic image, where the elongation or shrinkage of the curvature is analyzed. (6) Entropy, it measures the randomness of the pixels in the image, and corresponds to the average of information it has after applying the segmentation process, to identify information relevant to its analysis.

Table 4 describes the parameters of the image associated with the characteristics that were extracted in the previous stage to be processed.

Table 4. Name of the parameter and number of features extracted

Name of the evaluated parameter	Number of extracted characteristics
Average intensity	1
Average contrast of standard deviation	1
Mildness	1
Third moment	1
Uniformity of the histogram	1
Entropy	1
TOTAL	6

2.2. Sorter

To evaluate the 6 parameters or characteristics extracted, supervised learning techniques will be used, however, it is verified that it is simple and agile during the processing. For which purpose it is tried to review some proposals found in the literature, in such a way that when evaluating and comparing its performance, it allows to realize a correct discrimination and detection of thresholds, thus we have the most suitable techniques: ZeroR, Multilayer Perceptron and Naive Bayes. (1) ZeroR: is a classifier that predicts the majority (nominal) or average (numerical) value class according to the purpose. In addition, it is easy to implement and serves as a basis for measuring the performance of other classifiers [8]. (2) Multilayer perceptron: it is a logistic regression classifier; with a neural network of backward propagation formed a single hidden layer, which is used to solve problems that cannot be separated linearly, reducing as much as possible the threshold in the values of bias [8]. (3) Naïve Bayes: it is an automatic probabilistic learning model that is used for the task of simple classification and widely used for its sophisticated way of working, it is useful for quite large data [8].

3. Evaluation and validation results

At this point, feedback is given on the process that has been carried out to carry out the validation tests with the diagnosis obtained: in the first instance, the images are acquired in a

Foundation of the city of Quito; these images being those corresponding to elderly people who have a clinical Alzheimer's disease, with which a Database is structured.

As a second point, the images are taken to apply digital processing, thus obtaining the parameters or characteristics necessary for the analysis of said criteria. The derivative of this process resulted in obtaining 3 characteristics that do not influence the results (Table 5), on the other hand, the remaining 3 characteristics were considered of high relevance at the time of presenting the diagnosis (Table 6). Analyzing the results scenario, we choose to exclude the less relevant characteristics to avoid the over-training of the classifier.

Then, in the analysis stage of the possible pathology, the proposed classification methods are used, to evaluate the efficiency of each one and to determine which is ideally suited to the particularities of the prototype. Obtaining as a result a system that is capable of performing recognition through automatic learning, during training.

Table 5. Unnecessary characteristics in the diagnosis of Alzheimer's

Name of the evaluated parameter	Number of extracted characteristics
Mildness	1
Third moment	1
Average intensity	1
TOTAL	3

Table 6. Characteristics needed in the diagnosis of Alzheimer's

Name of the evaluated parameter	Number of extracted characteristics
Average contrast of standard deviation	1
Entropy	1
Uniformity of the histogram	1
TOTAL	3

Next, the parameters within the classifiers are evaluated and thus determine if they adapt to the proposed diagnostic model. In the case of ZeroR, the characteristics of medium intensity, smoothness and standard deviation in RGB are analyzed. On the other hand, in the case of the multilayer perceptron, the characteristics of the standard deviation of RGB and the third moment are analyzed. Finally, similar tests were carried out in Naïve Bayes, unlike the other classifiers, in the latter it was possible to establish decision rules as shown in Figure 8.

At this point, it is concluded that the efficiency of the ZeroR and Multilayer Perceptron classifiers are not suitable for the process of identifying the characteristics, since their predictive capacity is approximately 0.45 in relation to the Naïve Bayes classifier with a predictive capacity of 0.63, this being much higher and reliable in terms of precision, sensitivity and specificity, which will allow to establish an appropriate diagnosis. analyzed and presented in terms of accuracy, sensitivity, specificity and predictive capacity; all this analysis is achieved by applying a tool known as "ROC Curve" [13], used in clinical/medical research, which allows to measure the overall performance of a test and compare it with the other tests; for this, the area under the curve called "receiver performance characteristic" to verify the performance of the system (see equations 1-3).

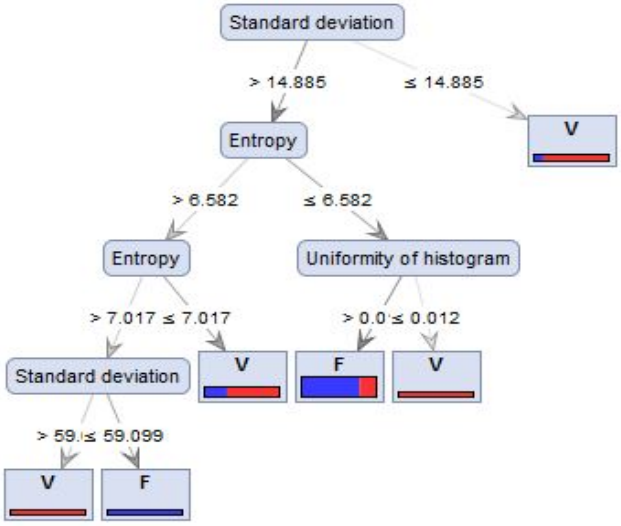


Figure 8. Decision Tree

Accuracy is the ratio between the addition of true positives and true negatives, out of a total number of cases inspected:

$$Accuracy = \frac{(VP + VN)}{VP + VN + FP + FN}\ 100\% \quad (1)$$

Sensitivity measures the model's ability to discover true positives:

$$\text{Sensitivity} = \frac{\text{VP}}{\text{VP} + \text{FN}}\ 100\% \quad (2)$$

The specificity is the percentage value that represents the group of healthy people or who do not have the correctly identified conditions:

$$\text{Specificity} = \frac{\text{VN}}{\text{FP} + \text{VN}}\ 100\% \quad (3)$$

Given the following: i) VP – true positive; ii) VN – true negative; iii) FP – false positive; and, iv) FN – false positive.

The result of the ROC curve shows the sensitivity to the specificity that each classifier has for all threshold values, in which they can take the different values to yield a possible decision based on characteristics. During the analysis of the area under the curve, we have the following observations: (1) the area as close as possible or equal to 1, indicates that the predictor has ideal prediction characteristics, which does not occur with an area near or equal to 0.5. Where the predictor has more random characteristics of prediction, what in a clinical system is relevant in the diagnosis. Then, the data is taken to elaborate the confusion matrix of the classifiers and evaluate them in the ROC curve.

Table 7 shows the average of the precision, sensitivity, specificity and predictive capacity obtained by each classifier, during the analysis that the ROC curve showed for each case. Thus it can be identified that the Naïve Bayes classifier has a better performance compared to the other two classifiers. Therefore,

Table 8 presents the confusion matrix of this classifier, with which the values of precision, sensitivity and specificity of the system were obtained.

Table 7. Efficiency of classifiers

Sorter	Accuracy %	Sensitivity %	Specificity %	Predictive ability
ZeroR	61.9565	100	0	0.454
Multilayer perceptron	61.9565	92.7272	16.2162	0.449
Naïve Bayes	61.9565	74.0000	47.6191	0.632

The results shown in Table 7 show the efficiency level of the classifiers, during the analysis the ROC curve was shown for each case. Then, it can be concluded that the Naïve Bayes classifier is more efficient in relation to the other two classifiers. Therefore, the confusion matrix of the Naïve Bayes classifier is elaborated, as shown in Table 8, from which the precision, sensitivity and specificity values of the proposed system were obtained.

Table 8. Matrix of confusion of Naïve Bayes

		Alzheimer's Prediction	
		Yes	No
Alzheimer	Yes	37	20
	No	22	13

During the analysis of the results shown in the ROC curve (Figure 9), all possible positive values (true positives and false positives) were identified to generate the diagnosis. On the other hand, the area value of the ROC curve oscillates between 0.632, which corresponds to 63.2% of the result of the diagnosis that was made to the patient with the pathology is more precise than that of healthy patient who is chosen at random, all this taking into account the quality of the diagnostic test.

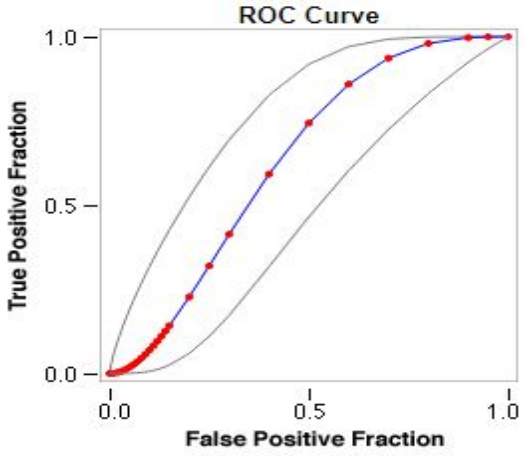


Figure 9. ROC curve for affirmative Alzheimer's values (Naïve Bayes classifier)

We must mention some considerations to establish the reasons why the system cannot get an optimal diagnosis: (1) The first reason involves the anatomy of the human eye, where the alterations or deformations in the pupil are due to genetic processes or degenerative age, since some images obtained and sent to the segmentation process do not allow to identify the circumference; (2) the second reason involves the specialized team with which the images of the eye are captured, in which the main influence is the reflection of the light in the pupil of the eye, since during the detection of the associated characteristics in said area can be extracted partially, or none that contribute with the diagnosis.

The results of Naïve Bayes obtained in the process will be used as input data to determine the best approximation to the area of the brain. For this, Newton's interpolation is used, which allows to establish and model curves for each proposed scenario in which abnormalities occur in the area of the brain and determine the diagnostic.

During the interpolation of Newton, mathematical functions that represent the diagnostic will be obtained, then functions are compared with each other, to determine the model that has the least error in approaching the area of the brain. The purpose is to improve in some way the level of precision in the preventive diagnosis.

3.1. Stage 1: Analyzing the data in image 1 of the database

Result: "No abnormalities have been found in the brain area".

Table 9. Naïve Bayes range 40-pixel pupil and 110-pixel brain.

X	R	Y
1	2.4650	99.8260
2	2.5970	23.5004
3	2.5070	0.0084
4	0.2550	-0.0009
5	0.2550	0.0135
6	0.2550	6.5143

Figure 10 shows the result of the interpolation of the points obtained in the image processing using Naïve Bayes shown in Table 9, in such a way that the following function is interpreted:

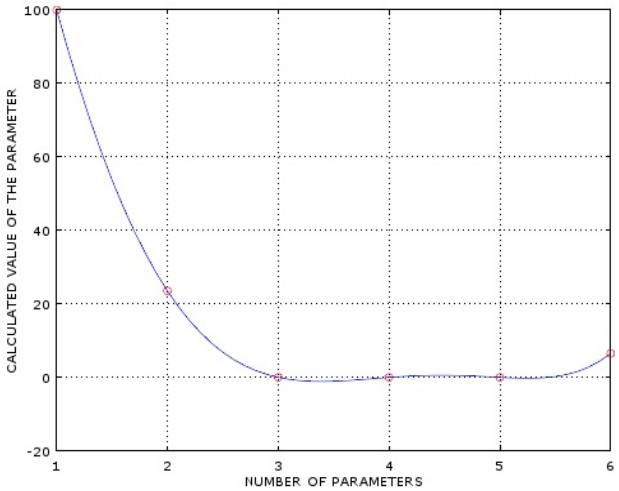


Figure 10. Modeling of the function for Naïve Bayes for the range of 40-pixel pupil and 110-pixel brain

$$0.20024x^5 - 2.7581x^4 + 9.6737x^3 + 19.306x^2 - 166.79x + 240.2 \quad (4)$$

Table 10. Naïve Bayes range 30-pixel pupil and 100-pixel brain.

X	R	Y
1	2.4650	131.251429
2	2.5970	25.564477
3	2.5070	0.009951
4	0.2550	-0.227350
5	0.2550	0.013389
6	0.2550	6.569150

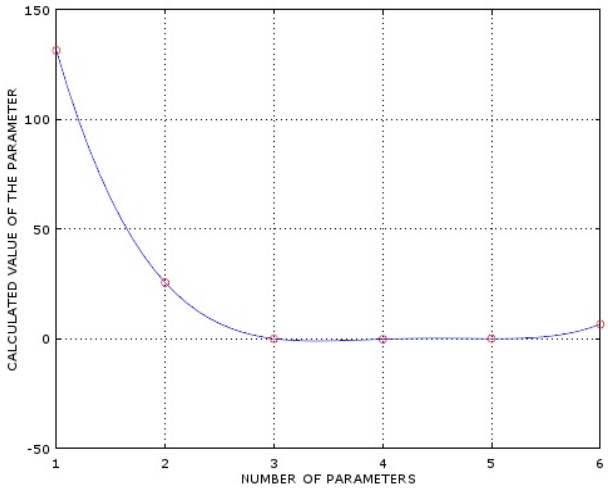


Figure 11. Modeling of the function for Naïve Bayes for the range of 30-pixel pupil and 100-pixel brain

Figure 11 shows the result of the interpolation of the points obtained in the image processing using Naïve Bayes shown in Table 10, in such a way that the following function is interpreted:

$$0.0058346x^5 + 1.1615x^4 - 21.13x^3 + 137.28x^2 - 387.23x + 401.16 \quad (5)$$

Table 11. Naïve Bayes range 20-pixel pupil and 90-pixel brain.

X	R	Y
1	2.4650	74.029405
2	2.5970	17.280810
3	2.5070	0.004571
4	0.2550	0.053247
5	0.2550	0.019178
6	0.2550	6.040649

Figure 12 shows the result of the interpolation of the points obtained in the processing of the image using Naïve Bayes (See Table 11), in such a way that the following function is obtained:

$$0.15672x^5 - 2.1533x^4 + 7.6549x^3 + 13.534x^2 - 123.49x + 178.33 \quad (6)$$

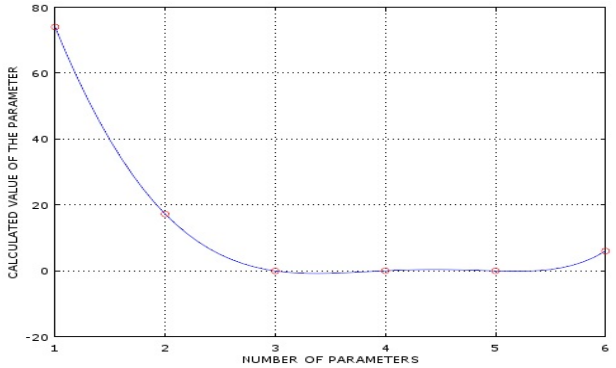


Figure 12. Modeling of the function for Naïve Bayes for the range of 20-pixel pupil and 90-pixel brain

3.2. Result of the comparison of functions Stage 1

For the comparison of the results of image 1 of the database with criteria: "No abnormalities have been found in the area of the brain"; For this reason, it is proposed to calculate the levels of error in the approximation of functions, using a code of own authorship that will basically be evaluated in the same intervals of functions.

In order to appreciate the graph and analytically the error that would later be compared with a table that explains this process.

3.3. Results of the analysis of image 1 of the database

Result: “No abnormalities have been found in the brain area”.

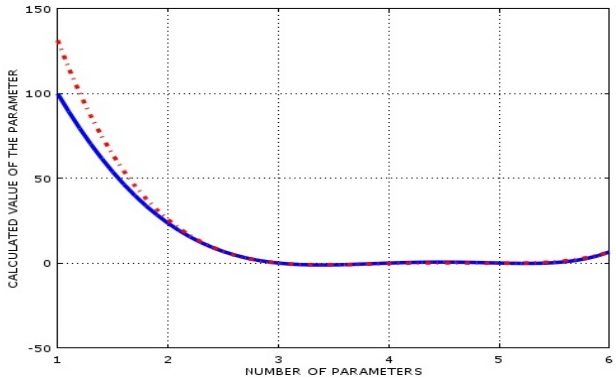


Figure 13. Comparison between the functions: Range of 40 pixels (iris) and 110 pixels (brain) in red; Range of 30 pixels (iris) and 100 pixels (brain) in blue of the pupil

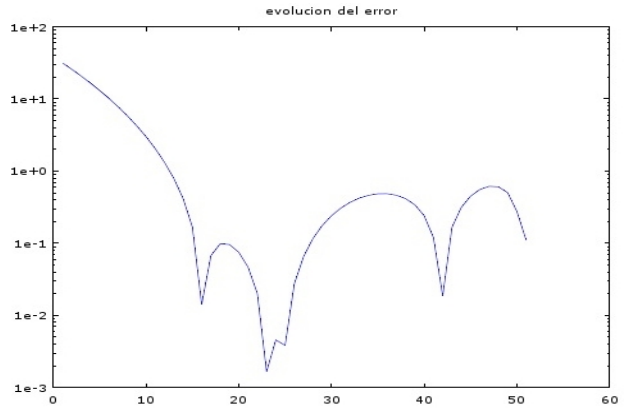


Figure 14. Approximation of the error between the functions: Range of 40 pixels (iris) and 110 pixels (brain); Range of 30 pixels (iris) and 100 pixels (brain) of the pupil

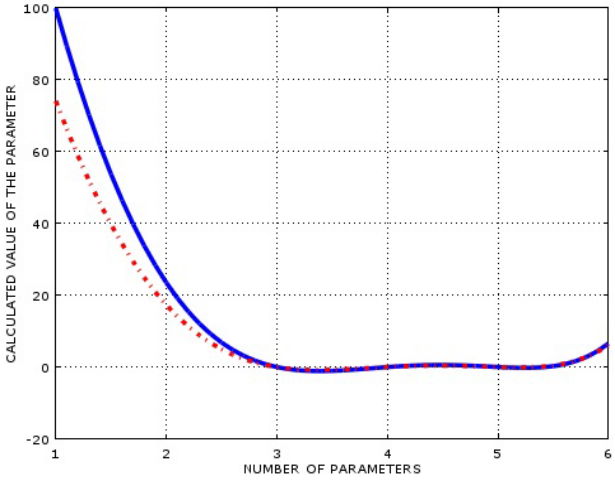


Figure 15. Comparison between the functions: Range of 40 pixels (iris) and 110 pixels (brain) in red; Range of 20 pixels (iris) and 90 pixels (brain) in blue of the pupil

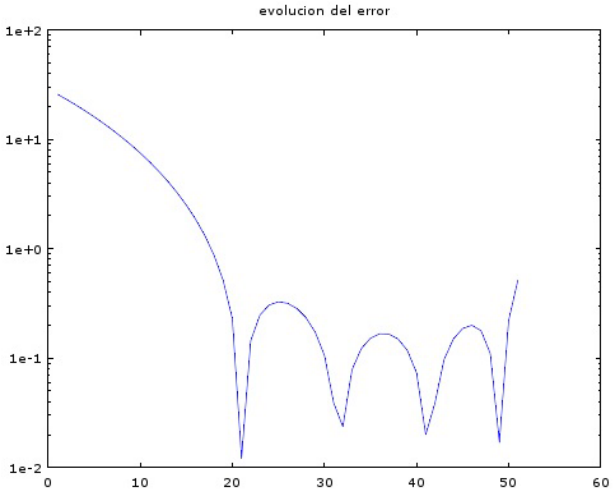


Figure 16. Approximation of the error between the functions: Range of 40 pixels (iris) and 110 pixels (brain); Range of 20 pixels (iris) and 90 pixels (brain) of the pupil

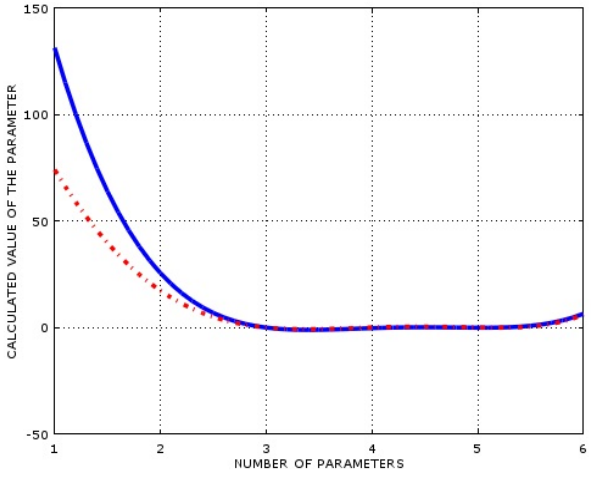


Figure 17. Comparison between the functions: Range of 30 pixels (iris) and 100 pixels (brain) in red; Range of 20 pixels (iris) and 90 pixels (brain) in blue of the pupil

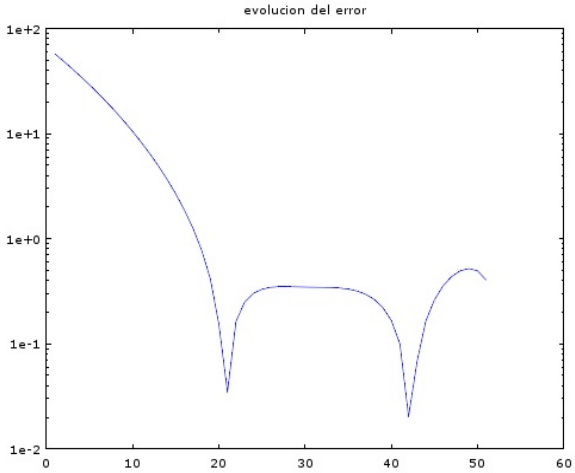


Figure 18. Approximation of the error between the functions: Range of 30 pixels (iris) and 100 pixels (brain); Range of 20 pixels (iris) and 90 pixels (brain) of the pupil

After appreciating the results of the same program where stage 1 was evaluated error reference values as shown below it was obtained Table 12:

Table 12. Referential error values analyzed with respect to image 1 of the database

Error reference values Result: "No abnormalities have been found in the brain area"	
Function	Error
(1) range of 40 pixels (iris) - 110 pixels (brain) with 30 pixels (iris) - 100 pixels (brain)	0.48533
(2) range of 40 pixels (iris) - 110 pixels (brain) with that of 20 pixels (iris) - 90 pixels (brain)	0.52800
(3) range of 30 pixels (iris) - 100 pixels (brain) with that of 20 pixels (iris) - 90 pixels (brain)	0.57215

During the evaluation of the pairs of functions (x, y) for the ranges that delimit the study sector of the brain, the error values that are generated in an image without anomalies detected were estimated and in this case which is adapted as a contribution of improvement to the proposed diagnostic model (Table 12).

3.4. Stage 2: Analyzing the data in image 2 of the database

Result: "Abnormalities have been found in the brain area".

Table 13. Naïve Bayes range 40-pixel pupil and 110-pixel brain.

X	R	Y
1	2.4650	108.3348
2	2.5970	29.2406
3	2.5070	0.0130
4	0.2550	-0.6294
5	0.2550	0.0176
6	0.2550	6.2318

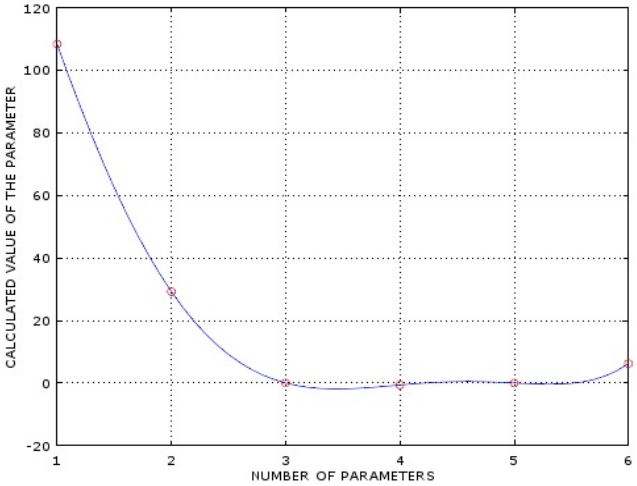


Figure 19. Modeling of the function for Naïve Bayes for the range of 40-pixel pupil and 110-pixel brain

Figure 19 shows the result of the interpolation of the points obtained in the image processing using Naïve Bayes shown in Table 13, in such a way that the following function is obtained:

$$0.31323x^5 - 4.9491x^4 + 25.584x^3 - 33.034x^2 - 94.553x + 214.97 \quad (7)$$

Table 14. Naïve Bayes range 30-pixel pupil and 100-pixel brain.

X	R	Y
1	2.4650	99.571548
2	2.5970	29.976068
3	2.5070	0.013630
4	0.2550	0.350826
5	0.2550	0.012525
6	0.2550	6.722807

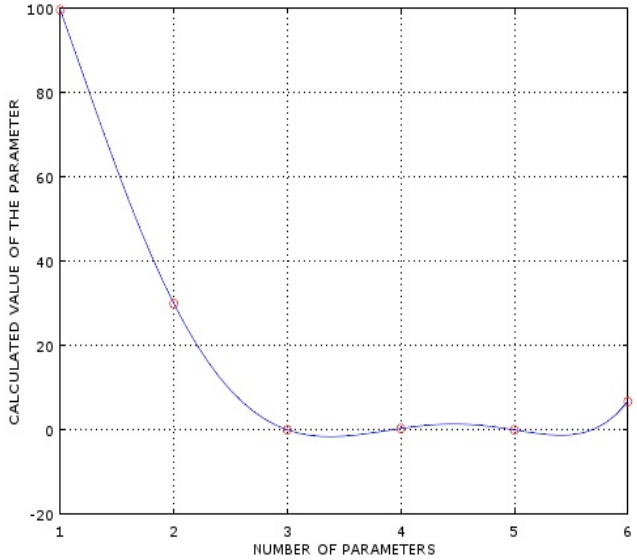


Figure 20. Modeling of the function for Naïve Bayes for the range of 30-pixel pupil and 100-pixel brain

Figure 20 shows the result of the interpolation of the points obtained in the image processing using Naïve Bayes shown in Table 14, in such a way that the following function is obtained:

$$0.50284x^5 - 8.4444x^4 + 50.203x^3 - 115.55x^2 + 36.709x + 136.15 \quad (8)$$

Table 15. Naïve Bayes range 20-pixel pupil and 90-pixel brain.

X	R	Y
1	2.4650	84.176905
2	2.5970	23.862226
3	2.5070	0.008681
4	0.2550	0.266040
5	0.2550	0.016836
6	0.2550	6.352425

Figure 21 shows the result of the interpolation of the points obtained in the image processing using Naïve Bayes shown in Table 15, in such a way that the following function is obtained:

$$0.36647x^5 - 6.0081x^4 + 34.203x^3 - 69.764x^2 - 11.68x + 137.06 \quad (9)$$

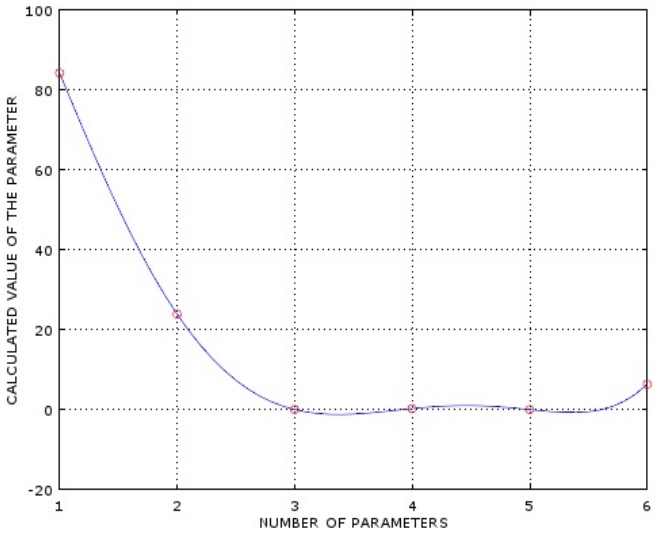


Figure 21. Modeling of the function for Naïve Bayes for the range of 20-pixel pupil and 90-pixel brain

3.5. Result of the comparison of functions Stage 2

In this case for the comparison of the results of image 2 of the database with criteria: "anomalies have been found in the area of the brain"; For this reason it is proposed to calculate the error levels in the approximation of the functions using the same own authorship code applied in scenario 1, in order to evaluate in the same intervals pairs of functions. To appreciate graphically and analytically the error that was later compared in a table that explains this process.

3.6. Results of the analysis of image 2 of the database

Result: “Abnormalities have been found in the brain area”.

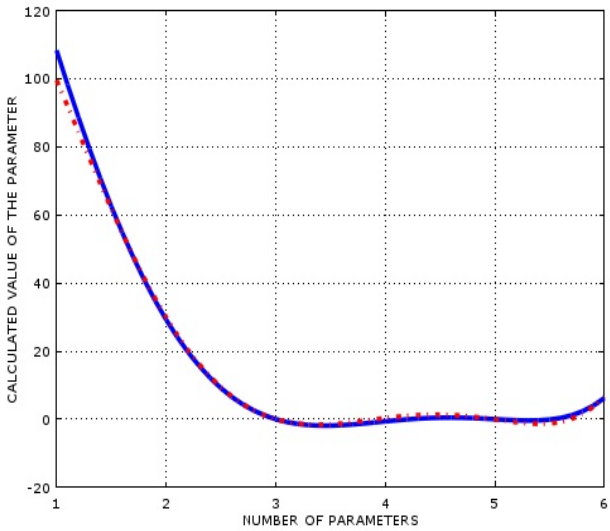


Figure 22. Comparison between the functions: Range of 40 pixels (iris) and 110 pixels (brain) in red; Range of 30 pixels (iris) and 100 pixels (brain) in blue of the pupil

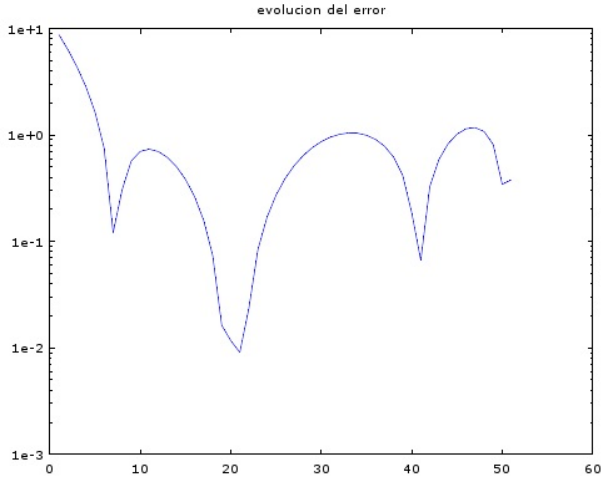


Figure 23. Approximation of the error between the functions: Range of 40 pixels (iris) and 110 pixels (brain); Range of 30 pixels (iris) and 100 pixels (brain) of the pupil

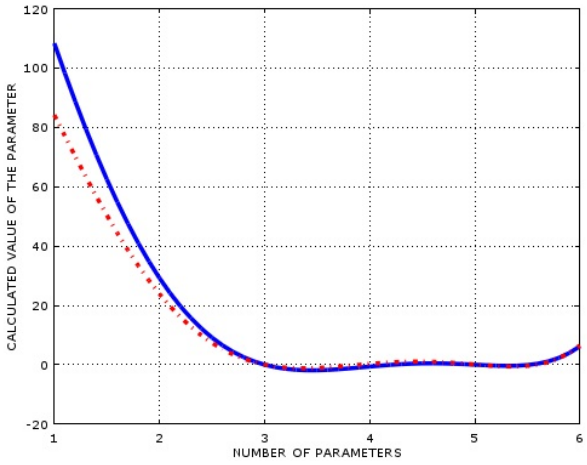


Figure 24. Comparison between the functions: Range of 40 pixels (iris) and 110 pixels (brain) in red; Range of 20 pixels (iris) and 90 pixels (brain) in blue of the pupil

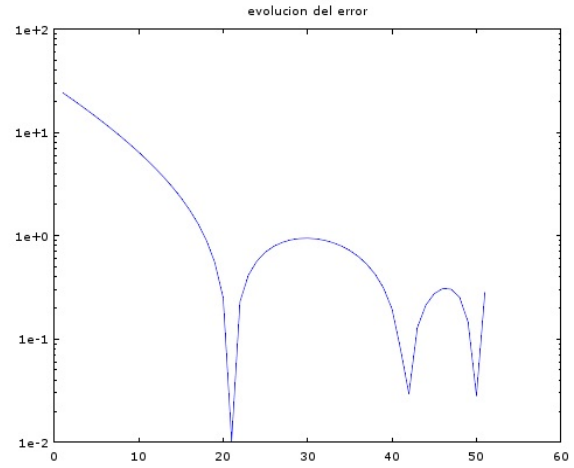


Figure 25. Approximation of the error between the functions: Range of 40 pixels (iris) and 110 pixels (brain); Range of 20 pixels (iris) and 90 pixels (brain) of the pupil

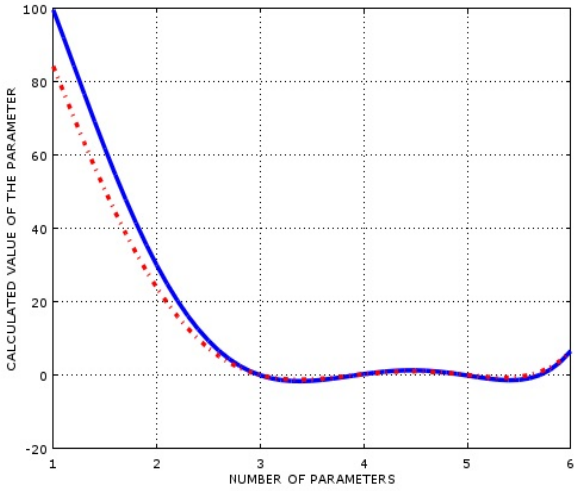


Figure 26. Comparison between the functions: Range of 30 pixels (iris) and 100 pixels (brain) in red; Range of 20 pixels (iris) and 90 pixels (brain) in blue of the pupil

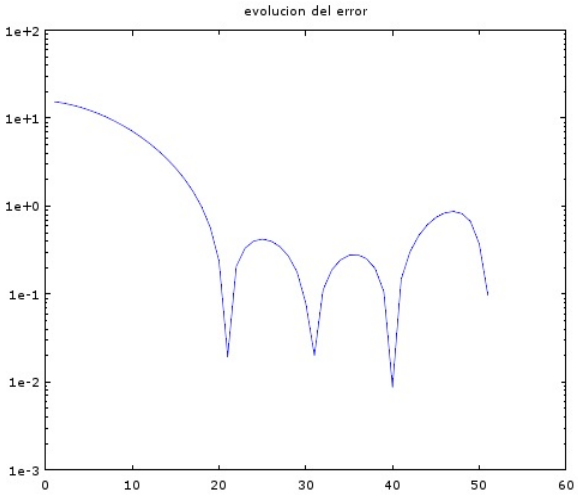


Figure 27. Approximation of the error between the functions: Range of 30 pixels (iris) and 100 pixels (brain); Range of 20 pixels (iris) and 90 pixels (brain) of the pupil

After appreciating the results obtained from the same program where scenario 2 was evaluated, reference error values were obtained as shown below (Table 16):

Table 16. Referential error values analyzed with respect to image 2 of the database

Error reference values Result: "Found abnormalities in the brain area"	
Function	Error
(1) range of 40 pixels (iris) - 110 pixels (brain) with 30 pixels (iris) - 100 pixels (brain)	0.87607
(2) range of 40 pixels (iris) - 110 pixels (brain) with that of 20 pixels (iris) - 90 pixels (brain)	0.24154
(3) range of 30 pixels (iris) - 100 pixels (brain) with that of 20 pixels (iris) - 90 pixels (brain)	0.15393

During the evaluation of the pairs of functions (x, y) for the ranges that delimit the study sector of the brain, the error values that are generated in an image with anomalies detected were estimated and in this case which is adapted as an improvement to the model of proposed diagnosis (Table 16).

3.7. General analysis of the results obtained

The results of the pairs of functions (x, y) reviewed and evaluated, were analyzed to conclude that those with a lower level of error will be taken in the study, in order to allow a tolerable approximation to the solution being investigated. For this case, the functions are taken for the range (1) (Table 12) and the functions for the range (3) (Table 16) with error levels of 0.48533 and 0.15393 respectively.

During the validation of the model, the permissible error level with which the system works was estimated (Table 17), analyzing the pairs of functions (x, y) the range of 30 pixels, shown in Fig. 28.

Table 17. Referential value of error

Referential error value permissible for the model	
Function	Error Value
(3) 30 pixeles (iris) – 100 pixeles (cerebro)	0.075019

The characteristics or parameters that were previously extracted during the processing stage are evaluated with the best approximation model to adjust the diagnostic results. In Fig. 28 can be analyzed aspects such as: possible pathology or some anomaly located in the iris of the eye that corresponds to the brain area, and achieve compare the results thrown by the system.

Table 17 indicates the level of error, which is within a margin of 1%, being permissible at the time of presenting correct results, including understanding that the functions are adjusted to the requirements of the system. In any case, the error can be controlled and improved by using processing data in other types of models that can adapt to the proposed predictive system.

These results are favorable to the problem of the lack of early detection in patient has Alzheimer's disease; the diagnosis will be verified when attending a medical appointment and receiving the appropriate treatment avoiding the complication of the future pathology. The ROC curve indicates a good level of prediction, however, it must be optimized in order to implement a fully automatic system, and that is, a high value in the ROC curve is required. Based on this the present work determined that the predominant value is the sensitivity to specificity, since during the process it was analyzed and determined that a pathology of easy analysis and interpretation is better in healthy patients. With the implementation of patterns in corrective models on the approach function to the brain area, the level of approach proposed in this work can be improved. The numerical method of Newton's interpolation is already one of the modeling processes of a function that allows coupling certain approximation criteria for detection of the brain area.

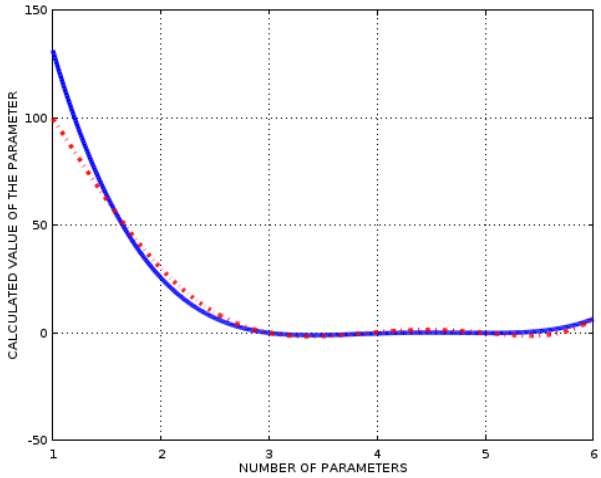


Figure 28. Comparison between functions with a range (3). Has Alzheimer's disease (blue), does not have Alzheimer's disease (red)

The results presented in the study are a contribution to the analysis and medical support, as a response to the lack of alternative methods of early detection of Alzheimer's disease. During the diagnosis, the doctor will use the results to hire and help them, as well as when the patient attends a medical appointment, will receive the indications and treatment to prevent future complications and a possible condition of this pathology.

On the other hand, we have the results of the ROC curve as satisfactory in terms of its predictive capacity, however, it is necessary to optimize it by improving the growth of the curve to obtain an area closer to 1, turning it into a fully automatic system. Taking these results, it can be determined that there is a higher level of preponderance in terms of sensitivity and specificity, since during the study it can be verified that the pathology can be analyzed and contributes with the doctors in a better way. Another observation regarding the model of approach to the area of the brain we can analyze that the method of interpolation of Newton allows coupling based on the criteria of approach to detect the area that corresponds to the brain, however, it can be optimized with the use of patterns based on the models generated by the functions.

4. Conclusions and future work

One of the most significant differences between each classification algorithm is the percentage of accuracy, which may vary depending on the characteristics that are implemented in each of them.

In order to reach an adequate characterization of subtypes of this pathology, more significant and more relevant characteristics must be used, that is, a better specific treatment based on patterns that can present each of these characteristics provided by the processing of the iridology template.

The adequate extraction of characteristics is proportional to an adequate recommendation of the pathology in question, it can be achieved with an exhaustive work of spatial image processing.

In the work in question, the most relevant classifier is Naïve Bayes with a 61.96% of accurate diagnoses, a 74.00% probability of success that a patient with this pathology obtain a diagnosis with an assertive result and a 47.62% of a patient in optimal conditions, this pathology does not.

The values of this study allow to have a better panorama of the situation that is being handled and of the variants that a future could intervene for the improvement of the proposed algorithm.

For associated works that are going to be carried out, more sophisticated tools such as convolutional neural networks should be used, specifically if you have adequate data for the work in question.

The use of artificial vision tools is limited for the present work since it is not strictly counted with a sufficiently robust data for the use of it.

It is recommended the use of google APIs like Tensor Flow, as an artificial vision tool to compare the results.

Supervised and unsupervised learning can significantly help the crystallization of a project, the parameters to be used should always be defined by a multidisciplinary group where specialists and technicians of the area are always involved so that the emission of results always contains a significant value.

References

[1] Hernández, F., Vega, R., Tapia, F., Morocho, D., & Fuertes, W. (2018). Early detection of Alzheimer's using digital image processing through iridology, an alternative method. 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), 1-7.

[2] McGrory, S, Cameron, J, Pellegrini, E, Warren, C, Doubal, F, Deary, I, Dhillon, B, Wardlaw, J, Trucco, E & MacGillivray, T 2017, "The application of retinal fundus camera imaging in dementia: a systematic review" Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, vol 6, pp. 91-107. DOI: 10.1016/j.dadm.2016.11.001.

[3] N.D.Pergad , Miss. S. B. More,* E&TC, Shri Tulja Bhavani College of Engineering, Tuljapur, India, Pergad*, 4(7): July, 2015, "Detection of diabetic presence from iris by using support vector machine", ISSN: 2277-9655.

[4] Casals, J. (2013-2018). Datelobueno.com: La Web de Jordi Casals. New York, EU.: Curso de Iridologia. Recuperado de http://datelobueno.com/.

[5] Gao Xiaoxing1, Feng Sumin2, Cui Han, Shijiazhuang University of Economics, Shijiazhuang 050031, China, Shijiazhuang Engineering Vocational College ,Shijiazhuang 050031, China, June 2015, 'Enhanced iris recognition based on image match and hamming distance', VOL. 8.

[6] Najmeh dashti nejad, June 2015, 'Diagnosis of heart disease and hyperacidity of stomach through iridology based on the neural network', Vol. 2, No. 6, 2015, pp. 17-25, ISSN 2454-3896.

[7] Mateusz Trokielewicz, Adam Czajka, Piotr Maciejewicz, Poland-2014, "Cataract influence on iris recognition performance", Proc. of SPIE Vol. 9290 929020-1, doi: 10.1117/12.2076040.

[8] Sarika G. Songire, Madhuri S. Joshi, PhD, India-431001, January 2016, "Automated Detection of Cholesterol Presence using Iris Recognition Algorithm", Volume 133 – No.6.

[9] Ali Azimi Kashani*, Alimohamad Monjezi Nori, Iman Mosavian, 2015, "New methods of verification and identification using iris patterns", pp. 118-122, ISSN 1115-7569.

[10] Mateusz Trokielewicz, Adam Czajka and Piotr Maciejewicz, "Assessment of iris recognition reliability for eyes affected by ocular pathologies", The IEEE Seventh International Conference on Biometrics: Theory, Applications and Systems (BTAS 2015), September 8 – 11, 2015, Arlington, USA.

[11] Chandranayaka, I.R, Jul-Aug, 2016, "Various iris recognition algorithms for biometric identification: a review", Vol. 03, Issue 04, Pg. 286 – 290, ISSN – 2394-0573.

[12] Mr. Langade Umesh, Ms. Malkar Mrunalini, Dr. Swati Shinde, March 2016, "Review of Image Processing and Machine Learning Techniques for Eye Disease Detection and Classification", Volume: 03 Issue: 03, e-ISSN: 2395 - 0056, p-ISSN: 2395-0072.

[13] Gonzalez, Rafael C.; Woods, Richard E. "Digital Image Processing", 2007, Addison-Wesley. Publishing Company, Inc.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 138-144 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

A Comprehensive Review on the Feasibility and Challenges of Millimeter Wave in Emerging 5G Mobile Communication

Cosmas Kemdirim Agubor[*,1], Isdore Akwukwuegbu[1], Mathew Olubiwe[1], Chikezie Onyebuchi Nosiri[1], Atimati Ehinomen[1], Akande Akinyinka Olukunle[2], Samuel Okechukwu Okozi[2], Longinus Ezema[2], Benjamin Chukwujekwu Okeke [2]

[1]*Department of Electrical and Electronic Engineering, Federal University of Technology, Owerri, Nigeria*

[2] *Department of Information Technology, Federal University of Technology, Owerri, Nigeria*

ARTICLE INFO

Article history:
Received: 15 February, 2019
Accepted: 08 May, 2019
Online: 27 May, 2019

Keywords:
Beamforming
Massive MIMO
Millimeter wave
Microwave
Pathloss
Signal attenuation

ABSTRACT

This article presents a comprehensive review on the feasibility and challenges of millimeter wave in emerging fifth generation (5G) mobile communication. 5G, a multi-gigabit wireless network is the next generation wireless communication network. The mmWave cellular system which operates in the 30-300 GHz band has been proposed for use as the propagation channel. Its large bandwidth potential makes it a candidate for the next-generation wireless communication system which is believed to support data rates of multiple Gb/s. High frequency bands such as mmWave have channel impairments. These impairments are challenges that are necessary to be properly understood. Employing mmWave as a propagation channel requires dealing with these challenges which this paper is aimed at reviewing. One aim of the work is to discuss these challenges in a more elaborate manner using simple mathematical equations and graphics to ensure clarity. To achieve this, current related works were studied. Challenges and solutions are identified and discussed. Suggested research directions for future work are also presented. One is developing suitable electronic such as fast analog-to-digital (ADC) and digital-to-analog (DAC) systems necessary for the transmitter/receiver (TX/RX) system.

1. Introduction

The demand for high-speed reliable communications has always been on the increase. This demand has been a challenge to existing third generation (3G) wireless network as well as the fourth generation long term evolution – advance (4G LTE-A) which is the most current network. These ever increasing traffic demand, combined with significantly improved user experience have resulted to the drive towards the next generation 5G mobile communication networks due to its large capacity. It has been widely accepted that the capacity of the 5G wireless communication system will be able to handle 1000 times the capacity of the 4G (LTE-A) wireless communication [1]. The 5G network will therefore serve as a key enabler in meeting the continuously and ever increasing demands for future wireless applications.

It has also been the consensus that future 5G network should realize the goals of thousand-fold system capacity, hundredfold energy efficiency, ultra-high data rate, ultra wide radio coverage and an ultra-low latency [2], [3]. It is commonly assumed today that around the year 2020, a new 5G mobile network will be deployed [4]. The ability to have massive number of devices processed will be compulsory as there will be billions of connected devices in the 5G wireless communication network by 2020 [5]. This is because there will be an increase in the popularity of various intelligent or smart devices resulting to huge traffic demand.

The implementation of 5G has come with some concerns which have generated some interests. The provision of secure network infrastructure is one of such many areas of interest. In [6] methods of providing substantial security requirements were investigated. Spectral efficiency and energy efficiency requirements were investigated in [1]. Elsewhere in [7], the adoption of software defined network (SDN) in 5G as a platform to achieve efficient end-to-end (E2E) latency, authentication, hand over and privacy protection was studied. Interference challenges and mitigation techniques were extensively discussed in [8], [9].

System architecture is also an area of great interest. This is because a robust system architecture to enable Gbps user experience, seamless coverage, and green communications is a must for an aggressive 5G version [10]. This will be possible based on advanced technologies which are necessary for the above requirements to be practicable. These technologies will form the key elements of 5G wireless systems.

Some of these technologies have also gathered great interest and are seen as promising candidates for 5G wireless communication systems [11], [12]. One of such is the HetNet (heterogeneous network) technology. Described in [3], the HetNet creates a multi-tier topology where multiple nodes are deployed

[*]Corresponding Author: Cosmas Agubor, Email: aguborcosy@yahoo.com

with dissimilar characteristics such as transmit power, coverage areas, and radio access technologies.

Other technologies discussed in several literature include millimeter wave (mmWave) techniques, denser small cells (DSC), software defined air interface (SDAI), and high-efficiency multiple antenna techniques known as massive multiple input multiple output (mMIMO) [13]-[15].The area that have appealed to many researchers is the mmWave and its application in the next generation mobile network.

Available sources in public domain have records of the application of mmWave technique in some specialized areas. Its use in Radio-over Fiber (ROF) technology was discussed in [16]. The discussion centered on the main mmWave signal generation technique for ROF technology. In [17], mmWave was presented as a technology that has to be supported by signal processing in mmWave wireless systems and some challenges that may be faced in doing so. The paper laid emphasis on using MIMO at higher carrier frequencies.

Another area of interest is its application in 5G mobile communication. The common view here is its ability to support larger bandwidth compared to microwave frequencies. This attractive feature of mmWave was discussed in [18] as well as the advantages and disadvantages of its application in 5G networks. In another study, extensive propagation measurements were carried out at 28GHz and 38 GHz to determine the path loss, delay spread and penetration characteristics [19]. The measurements made were to obtain results that could be useful in the design of future 5G mmWave communication systems. Here, like in other related work, mmWave applications were documented but the likely challenges especially its propagation characteristics were not well discussed.

Since mmWave has been proposed to drive the 5G mobile network, it is necessary to have sufficient knowledge of these challenges that must be addressed. This is because understanding the radio channel is a fundamental requirement to developing future mmWave mobile communication systems [19]. This paper is out to address the above by focusing on some of the several major challenges and possible solutions of mmWave as a proposed propagation medium for future 5G networks.

The novelty of this work is that, challenges of mmWave as a transmission medium were identified from several sources and discussed in a more simplified and elaborate manner. Mathematical expressions obtained from these sources were used to carry out computations which were not so in the original literature. The purpose for the computations is to help in proper understanding of the subject. To carry out this work, current related scholarly works were consulted.

The contributions of this paper are listed as follows:

- We identified and discussed in a simplified and elaborate manner the challenges and possible solutions for the implementation of mmWave for 5G mobile communication.
- Unlike other literature where similar issues were either mentioned or listed, this work presents itself as a one source where much information on the subject matter could be obtained.
- We suggested possible directions for future work based on the reviewed articles.

The rest of the paper is organized as follows: section II covers a discussion on mmWave, a brief comparison with microwave frequency and propagation problems. In section III, the solutions and suggestions for further investigations are presented. Section IV is the conclusion.

2. Millimeter Wave

Generally the radio spectrum for the millimeter wave (mmWave) is between 30 GHz to 300 GHz. This band of frequencies utilizes wavelengths between 1 and 10mm. In practice however the frequencies suitable for wireless communication are between 71-76 GHz and 81-86 GHz bands which are referred to as E-band or the 70GHz and 80GHz bands [20]. The 5 GHz spectrum available in each of these bands makes mmWave a propagation medium with ultimate bandwidth that can be compared only to fiber optic (FO). Aggressive deployment of FO by operators may be restricted due to geographical constraints and economic reasons. Millimeter wave technology presents itself as the next attractive alternative capable to overcome such constraints.

Its suitability for wireless backhaul, immunity to interference, high capacity and inexpensive nature are discussed in [2]. Millimeter wave frequencies present signals with small wavelengths. This characteristic makes it potentially suitable for the deployment of large number of antennas for signal directivity and link reliability improvement by compensating severe path loss to achieve larger coverage [21]. It has a potential Gigahertz transmission bandwidth incomparable to other microwave band used in conventional cellular networks [22].

Table 1: Comparison of millimeter wave and microwave frequencies

Parameter	Millimeter wave Frequency	Microwave Frequency
Frequency band	30GHz-300GHz	300MHz-30GHzWave
Wavelength	10mm-1mm	1m-0.01m
Bandwidth	Ultrahigh	high
Antenna size/weight	Smaller due to very short wavelength	Large especially at the lower part of the band
Coverage	Suitable for short distance	Long distance application especially at 4-13 GHz band (Long haul)
Frequency reuse option	Suitable for Frequency reuse	Frequency reuse will likely cause interference
System Gain	Very high gain (Gain is proportional to frequency)	High gain
Attenuation	High during heavy rainfall	Good resistance to rain at lower frequencies
Peak rate	10-100 Gbps	1-5 Gbps
Application	Radar, mmWave imaging, medicine, mmWave scanner	Radio and television broadcasting, cellular telephony, satellite and terrestrial communication, radar, navigation

Although microwave which covers the band of frequency from 300MHz-300GHz has been widely used in wireless communication, the mmWave due to its higher frequency range has a greater prospect in terms of capacity delivery. Table 1 shows a brief comparison between mmWave and microwave frequencies. It is seen that based on their respective frequencies, mmWave has an advantage over microwave in bandwidth and antenna size. This two features have been accepted as very useful in the realization of 5G networks [13], [18].

2.1. Weather and Environmental Effects

(I) Propagation Losses

Other wireless technologies use microwave frequencies which have lower carrier frequencies compared to mmWave communication (Table 1). It is known that the higher the signal frequency the more likely it becomes susceptible to adverse atmospheric conditions. Therefore, in the GHz band of frequencies the atmosphere is seen as a propagation medium characterized with the presence of atmospheric constituents such as molecules, water vapor and suspended water droplets. Millimeter wave signals are absorbed by these atmospheric constituents. Also signals at GHz band suffer from rain attenuation. This is illustrated in Figures 1 and 2.

Figure 1 [23], shows the degree of attenuation suffered by frequencies between 1 to 1000 GHz band. At 75GHz in the E-band, it is observed that a mere rain drizzle results to 0.4dB loss and increases to about 30dB for a typical tropical rainfall. It is also shown that signal attenuation increases with rain intensity as depicted in Figure 2 [24]. At 200 mm/h of rain, 3GHz suffered 0.1dB loss as against 92dB loss for 30GHz. Details of mmWave signal attenuation are documented in [25] where experimental data obtained from both the rain intensity and rain attenuation measurements were statistically processed. Apart from rain attenuation, atmospheric absorption is also a major impediment to mmWave communications [26].

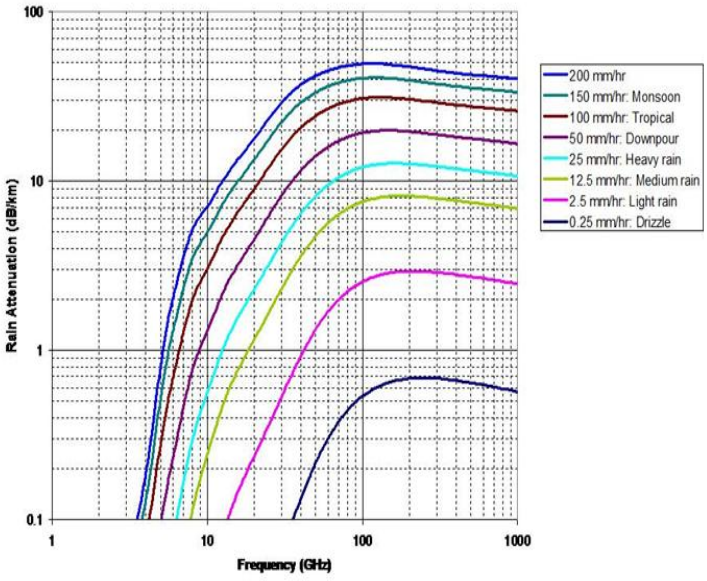


Figure 1. Effect of rain attenuation on mmWave frequencies [23]

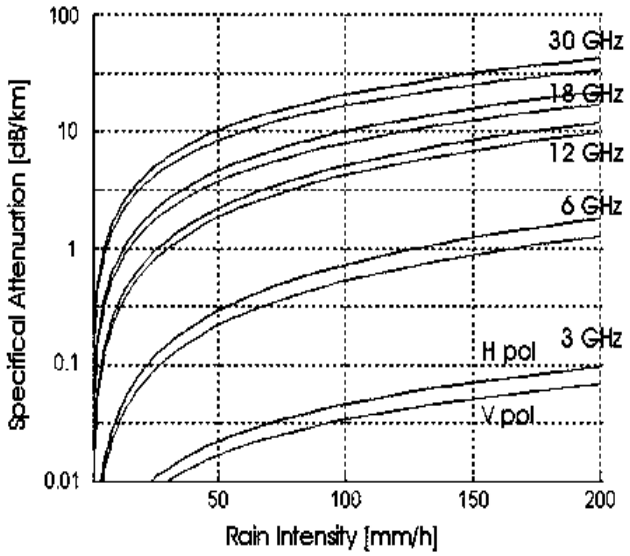


Figure 2. Rain attenuation vs frequency [24]

The rain attenuation and atmospheric absorption characteristics of mmWave propagation limit the range of mmWave communications [23], [27].

The propagation characteristics of mmWave communications in different bands are summarized in [23]. They showed the level of loss due to both rain attenuation and oxygen absorption under line-of-sight (LOS) and non-line-of-sight (NLOS) channels. Table 2 presents a summary of signal loss due to oxygen absorption at 200m range. It is observed that the propagation losses at 28GHz and 38GHz are not as significant as those of 60GHz and 73GHz.

Table 2: Absorption loss in mmWave frequencies [23]

Frequency Band (GHz)	Range (m)	Oxygen Absorption (dB)
28 GHz	200	0.04 dB
38 GHz	200	0.03 dB
60 GHz	200	3.2 dB
73 GHz	200	0.09 Db

(II) Free Space Loss

Due to its nature, mmWave frequencies experience greater free space loss than lower frequencies. In [28], the Free Space Loss *(FSL)* is shown to be inversely proportional to the square of the operating wavelength, ie

$$FSL = \frac{4\pi R}{\lambda^2} \quad (1)$$

Here, R in Km, is the link distance between transmit and receive antennas and λ the wavelength of the operating frequency. In decibel form and after converting to units of frequency, the equation can be expressed as [28]

$$FSL\ (dB) = 92.4 + 20\ log\ f + 20\ log\ R \quad (2)$$

Where f is the frequency in GHz.

Computed FSL in the E-band using (2) is tabulated in Table 3. The Table indicates a proportional increase in FSL with both frequency and distance. In particular, it indicates that for a given frequency the FSL increases with distance. Thus, the more the distance or range the higher the signal attenuation.

Table 3. Free space loss at mmWave frequencies

E-Band (GHz)	FSL (dB)				
	R=1m	R=2Km	R=3Km	R=4Km	R=5Km
71	129.4	135.4	139	141.5	143.4
72	129.5	135.6	139.1	141.6	143.5
73	129.7	135.7	139.2	141.7	143.6
74	129.8	135.8	139.3	141.8	143.8
75	129.9	135.9	139.4	141.9	143.9
76	130.0	136.0	139.6	142.0	144.0

(III) Foliage Loss

Foliage loss takes into account the effect of vegetation within the propagation environment such as tree size and nature or roughness of plant leaves. As the leaves become comparable in size relative to the wavelength, there is decrease in signal penetration through the leaves, while scattering off the leaves increases. [29].

Work in [28], showed an empirical relationship that can be used to predict or determine foliage losses. This was developed by CCIR Rpt 236-2 which reported that for a depth of less than 400m, the loss is given by

$$L = 0.2f^{0.3}R^{0.6}\ dB \quad (3)$$

Here, f is the frequency in MHz and covers the range 200-95,000MHz (0.2-95GHz). R is the foliage depth in meters (R< 400 m).

Using (3), the foliage loss for the E-band was computed and tabulated in Table 4. For a range of 300m the loss at 70GHz is 174dB which increased to 177dB at 75GHz. Similar computation in [28] indicated that at 40 GHz, a penetration of 10m taken to be equivalent of a large tree or two in tandem, the foliage loss is about 19 dB. These values show that foliage loss in mmWave is significant enough and like other forms of losses should not be neglected in overall network design.

Table 4: Foliage loss at mmWave frequencies

E-Band (GHz)	F (dB)		
	R=100m	R=200m	R=300m
70	90.06	136.51	174.11
71	90.45	137.10	174.86
72	90.83	137.67	175.59
73	91.20	138.24	176.32
74	91.58	138.81	177.04
75	91.95	139.37	177.78

(IV) Blockage Loss

Blockage in communication is caused by obstructions in the path of propagation which are man-made or natural physical structures. Such obstructions introduce losses which cannot be neglected. With a small wavelength, links in the 60 GHz band are sensitive to blockage by obstacles (e.g., humans and furniture) [23]. For example, blockage by a human penalizes the link budget by 20-30 dB [23][30]. This was also confirmed in [31] where it was shown that mmWave systems suffer from significant loss in performance due to blockages caused by humans or other obstacles along its propagation path.

(V) Penetration Loss

This loss arises as transmitted frequencies attempt to propagate through objects along its path. At mmWave frequencies the losses are more significant compared to UHF/microwave bands. Current works contained in most literature on this subject covers outdoor-to-outdoor and outdoor-to-indoor penetration loss measurements. More importantly is the outdoor-to-indoor measurements. A 28GHz outdoor-to-indoor measurements made and recorded in [32][33] were done using a rotating horn antenna channel sounder which can provide an accurate absolute delay information. The result indicated clusters, larger excess delays and larger angular spreads indoor.

Another work on 28GHz is described in [34] while [35] investigated the penetration loss for the band of 0.8GHz to 28GHz. Measurements described in [34][36][37] and[38] for both outdoor and indoor environments are summarized in Table 5.

Table 5: Summary of mmWave penetration losses

Ref	Frequency (GHz)	Material Type/Environment	Loss (dB)
[29]	28	Outdoor-indoor	3-60
[31]	38	Tinted glass Glass door	25.0 37.0
[32] [33]	28	Tinted glass(Outdoor) Brick pillar(Outdoor) Clear glass (Indoor) Dry wall (Indoor)	40.1 28.3 3.6 6.8

The Table shows penetration losses for common materials found in buildings. By observation, outdoor materials recorded higher penetration losses than indoor materials.

2.2. Hardware Implementation Challenges

Apart from atmospheric losses there are other challenges resulting from the type of hardware that can suitably and efficiently function at the mmWave frequencies. With high carrier frequency and wide bandwidth, there are several technical challenges in the design of circuit components and antennas for mmWave communications [23], [39]. Highlighted in [2], are the cost of electronic components and the complexity of transceiver including high-speed analog-digital converters (ADCs), digital analog converters (DACs), synthesizers, mixers, etc., which are much larger than that in conventional microwave communications.

In [37], the nonlinear distortion of power amplifiers (PA) especially at 60GHz was discussed as an impediment to the implementation of mmWave. Also in [40], [41], phase noise and IQ imbalance are seen too as challenging problems faced by radio frequency integrated circuits (RFIC).

3. Solutions and Suggested Areas for Further Investigations.

3.1. Solutions

The shortcomings of millimeter wave in terms of propagation and system hardware implementation have been discussed. Despite these challenges, its large bandwidth makes the usage of mmWave communications in the 5G cellular access still attractive

[42]. Some remedial steps are necessary for the practical realization of mmWave systems.

(I) Short Distance Communication (SDC)

In mmWave, path loss increases with distance as earlier shown. Its limited range makes it suitable for short distance applications. This gives room for the shrinking of cells which is an effective way to increase area spectral efficiency [43], [44]. Small cells are base stations (BS) that cover a small geographical area and are meant for low power, short range wireless communication. Cell size shrinking reduces the number of users per cell, thus more spectrum is made available to each user. The small cell sizes and short distances will improve frequency reuse [2]. Frequency reuse is a cellular concept which allows the use of same radio frequencies on BS within a geographical area. These BS are separated by sufficient distances to avoid or minimize signal interference with each cell.

(II) Reduced Inter cell Interference

Attenuation due to rain, foliage and atmospheric absorption has been discussed as major impediments or challenges to mmWave application. Deployment of small cells known as ultra dense cell (UDC) can help overcome such problems. This is because atmospheric absorption will efficiently increase the isolation of each cell by further attenuating or reducing the background interference from other distant base stations [23], [26], [45]. Densification of small cells has been proposed in [46], [47] as a technique to achieve increase in network capacity in the future and in [23], as the promising solution for the capacity enhancement in the 5G cellular networks.

(III) Hardware solution

Work in overcoming the technical challenges associated with electronic components of mmWave is ongoing and have been discussed in several research reports. Progress work on radiofrequency (RF) power amplifiers (PAs), low-noise amplifiers (LNAs), voltage-controlled oscillators (VCOs), etc., are documented in [39]. Radio frequency integrated circuits and other low-cost electronic components are believed to bring about the evolution of massively broadband 5G millimeter wave communications [39], [44]. Recorded in [48] are efforts made in providing a practical phased array antenna solution.

3.2. Suggested Areas for Further Investigation

Challenges of mmWave have been highlighted. Propagation loss is a major impediment in its practical implementation. A solution to this is to improve in antenna technology. This will result in the use of greater gain antennas. Beam steerable antenna technologies have the capability for such gains. The workings of such antennas are fully described in [49], [50]. These are antennas that have the ability to compensate the path loss caused by blockage from dynamic obstacles.

Since mmWave is associated with small wavelength and the size of an antenna greatly relates to the operating wavelength, large number of antennas or MIMO technology can also be made use of in mmWave communications. An example is the Massive MIMO (mMIMO) antennas widely discussed in [51]-[53]. This is also known as large-scale antenna systems (LSAS) [51], hyper-MIMO (HMIMO), or full-dimension MIMO systems [54].

This technology which allows the BS to be equipped with up to a hundred or more antennas is meant to leverage on the benefits of the traditional MIMO antenna system known for its ability to significantly improve the capacity and reliability of wireless systems [51] [55]. The challenges defined for mmWave can be tackled with the implementation of massive MIMO which theoretically promises increase in spectral and energy efficiencies [56]. It is believed that the antenna gain will be high enough to overcome high propagation loss in mmWave bands,

However, further development is required to make this concept practically functional. This is because a large antenna structure such as massive MIMO will have individual array elements each with active transceiver modules.

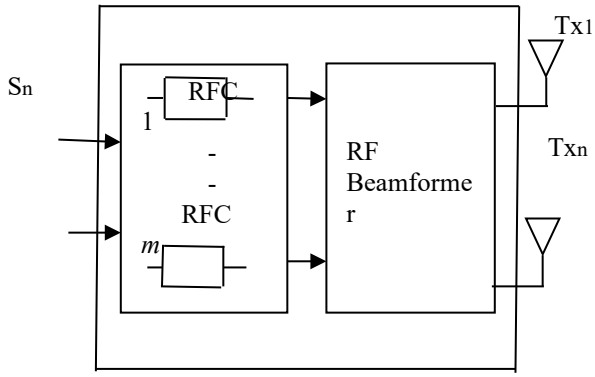


Figure 3.Transmit antennas with dedicated RF chains

That is each transmitter (TX) and receiver (RX) will have a dedicated RF chain (RFC) as shown in Figure 3. For massive MIMO with a hundred or more antennas it will result to high system cost as well as increase in system complexity.

Therefore, further work in developing low cost, low profile and appropriate electronic components is necessary. Examples of such electronic components are high-speed analog-to-digital and digital-to-analog converters (ADC/DAC), RF amplifiers, TX/RX switches, filters, etc. for each individual antenna element [57]. The direction here should be in hardware unit design that will reduce overall system complexity and cost.

Methods of solving large outdoor-to-indoor penetration loss caused by certain building materials (Table 5) is also another direction for further work.

4. Conclusion

Millimeter wave band of frequencies with ultimate bandwidth has been discussed as a potential candidate for 5G mobile communication. Key challenges in propagating at this frequency band as well as some hardware issues have also been discussed. The implementation of mmWave in the emerging 5G wireless communication will have to address these shortcomings. To this end some solutions have been suggested. Based on the study and since the mmWave 5G wireless communication is an emerging network still at its early stage of implementation, some selected areas for further work have also been suggested.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] C. X. Wang, S. Wu, L. Bai, X. You, J. Wang and I. Chih-Lin, "Recent advances and future challenges for massive MIMO channel measurements and models," Special Focus on 5G Wireless Communication Networks, 59, 1-16, 2016, Doi: 10.1007/s11432-015-5517-1).

[2] Z. Gao, L. Dai, D. Mi, Z. Wang, M. Ali Imran, and M. Z. Shakir, "MmWave massive MIMO based wireless backhaul for 5G ultra-dense network," arXiv:1508.03940v3 [cs.IT], 1-7, 2015

[3] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J.Yuan and M. Di Renzo, "Safeguarding 5G wireless communication networks using physical layer security," IEEE Communications Magazine, 20-27, 2015, DOI:10.1109/MCOM.2015.7081071.

[4] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd and T. Svensson, "The Role of Small Cells, Coordinated Multipoint and Massive MIMO in 5G," IEEE Communications Magazine, 2014

[5] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J Sachs and Y. Selen, "5G radio access," Ericsson Review, 6(1), 2014

[6] P. Schneider, "5G Security research at Nokia Bell Labs," Nokia Solutions and Networks, 2016.

[7] B. Uma and S. Sumathi, "High throughput, privacy and security for handover in 5G networks using software-defined networking," International Journal of Innovative Research in Science, Engineering and Technology, 5(2), 2016

[8] A.K. Mishra and S. Gaur (2016), "Review of the pilot contamination problem for massive MIMO and possible solution," International Journal Of Engineering Sciences & Research Technology, 5(7), 2016

[9] N. Mehrotra and A. K. Chaubey, "Pilot contamination effect in massive MIMO and analysis of mitigation techniques," International Journal of New Technology and Research, 3(2), 19-23, 2017.

[10] C.-L. I, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," IEEE Commun. Mag., 52(2), 66- 73, 2014.

[11] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," IEEE Commun. Mag., 52(2), 186-195, 2014.

[12] H. Quoc, N. Linköping, "Massive MIMO: Fundamentals and system designs," Linköping Studies in Science and Technology Dissertations, (1642), 2015

[13] R. Taori and A. Sridharan, "Point-to-multipoint in-band mmWave backhaul for 5G networks," IEEE Wireless Commun.53 (1), 195-201, 2015

[14] Q. Sun, C. Lin, S. Han, Z. Xu and Z. Pan, "Software defined air interface: a framework of 5G air interface," IEEE Wireless Comm. 2015. DOI: 10.1109/WCNCW.2015.7122520

[15] E. G. Larsson, O. Edfors, F. Tufvesson and T..L.Marzetta, "Massive MIMO for next generation wireless system," IEEE Comm magazine 52(2), 186-195, 2014. DOI: 10.1109/MCOM.2014.6736761

[16] D. Singh and P. Singh, "Techniques of millimeter-wave signal generation in ROF system: A Review," International Journal of Computer Applications and Information Technology,1(2), 45-49, 2012

[17] R .W. Health Jr, N.Gonzalez-Prelcic, S.Rangan, W.Roh and A.Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," Computer Science Information Theory, 2015

[18] D.D.Pai, "Survey on millimeter wave mobile communication for 5G cellular networks," International Journal of Innovative Research in Electrical, Electronic, Instrumentation and Control Engineering, 5(6), 278-284, 2017

[19] T.S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G . N. Wong, J.K. Schulz, M. Samimi and F. Gutierrez, "Millimeter wave mobile communication for 5G cellular: It will work," IEEE access, 1, 335-349, 2013

[20] P. Adhikari, "Understanding millimeter wave wireless communication," Loea Corporation, 1-6, 2008,

[21] S. Han, C.-L. I, Z. Xu and C. Rowell, "Large-scale antenna systems with hybrid precoding analog and digital beamforming for millimeter wave 5G," IEEE Commun. Mag., 53(1), 186-194, 2015.

[22] L. Wei, R. Q. Hu, Y. Qian and G. Wu, "Key elements to enable millimeter wave communications for 5G wireless systems," IEEE Wireless Commun., 21(6), 136-143, 2014

[23] Y. Niu, Y. Li, D. Jin, L. Su and A. V. Vasilakos,"A Survey of millimeter wave (mmWave) communications for 5G: opportunities and challenges," GrXiv: 502.07228v1 [cs.NI], 1-17, 2015

[24] "E-band technology" 2012. [Online]. Available www.e-band.com

[25] G. Timms, V. Kvičera and M.Grábner, "60 GHz band propagation Experiments on terrestrial paths in Sydney and Praha,"Radio Engineering, 4(4), 2005.

[26] F. Al-Ogaili and R.M. Shubair, "Millimeter wave mobile communications for 5G: challenges and opportunities,"2016. DOI:10.1109/APS.2016.7696210.

[27] Q. Zhao and J. Li, "Rain attenuation in millimeter wave ranges," in IEEE Int. Symp. Antennas, Propag. EM Theory,1-4, 2006

[28] "Millimeter wave propagation: Spectrum Management Implications," Federal Communications Commission, 1997

[29] M. Shafi, J. Zhang, H. Tataria, A. F. Molisch, S.Sun, T. S. Rappaport, F. Tufvesson, S.Wu, and K. Kitao, "Microwave vs. millimeter-wave propagation channels: key differences and impact on 5G cellular systems," IEEE Communications Magazine, 2018. DOI:10.1109/MCOM.2018.1800255

[30] S. Singh, F. Ziliotto, U. Madhow, E. M. Belding and M. Rodwell, "Blockage and directivity in 60 GHz wireless personal area networks: from cross-layer model to multihop MAC design," IEEE J. Sel. Areas Commun., 27(8),1400–1413,2009.

[31] A. Samuylov, M. Gapeyenko, D. Moltchanov, M. Gerasimenko, S.Singh, N. Himayat, S.Andereev, Y. Koucheryavy, "Characterizing spatial correlation of blockage statistics in urban mmWave systems," Globecom Workshop (GC Wkshps), IEEE, 1-7, 2016. DOI: 10.1109/GLOCOMW.2016.7848859

[32] C. U. Bas, R. Wang1, T. Choi, S. Hur, K. Whang, J. Park, J. Zhang and A. F. Molisch, "Outdoor to indoor penetration loss at 28 GHz for fixed wireless access," arXiv:1711.0168v1 [cs.IT], 2017.

[33] J. Ko, K. Lee, Y. J. Cho, S. Oh, S. Hur, N. G. Kang, J. Park, D. J. Park, and D. H. Cho, "Feasibility study and spatial-temporal characteristics analysis for 28 GHz outdoor wireless channel modelling, IET Communications, 10(17), 2352–2362, 2016.

[34] C. Larsson, F. Harrysson, B. E. Olsson, and J. E. Berg, "An outdoor-to-Indoor propagation scenario at 28 GHz," in 8th European Conference on Antennas and Propagation (EuCAP), 3301–3304, 2014.

[35] I. Rodriguez, H. C. Nguyen, I. Z. Kovcs, T. B._ Srensen, and P. Mogensen, "An empirical outdoor-to-indoor path loss model from below 6 GHz to cmwave frequency bands," IEEE Antennas and Wireless Propagation Letters, 16, 1329–1332, 2017

[36] I.R. Larrad, H.C. Nguyen, T.B. Sorensen, J.A. Holn, J.Elling, P. Mogensen and B.Vejlgaard, "Analysis of 38 GHz mmwave propagation characteristics of urban scenarios," in 21th European Wireless Conference; 1-8, 2015.

[37] T. S. Rappaport, S.Sun, R. Mayzus, H. Zhao, Y,Azar,K. Wang,G.N. Wong, J.K. Schulz, M. Samimi and F.Gutierrez "Millimeter wave mobile communications for 5G Cellular: It Will Work! "IEEE Access, 1, 335–349, 2013. DOI: 10.1109/ACCESS.2013.2260813

[38] H. Zhao, R. Mayzus, S. Sun, M. Samimi, J.K. Schulz, Y.Azar, K.Wang, G.N Wong, F. Gutierrez and T.S. Rappaprt , "," in IEEE ICC, 5163–5167, 2013. DOI: 10.1109/ICC.2013.6655403

[39] T. S. Rappaport, J. N. Murdock, and F. Gutierrez, "State of the art in 60-GHz integrated circuits and systems for wireless communications," in IEEE, 99(8), 2011, 1390– 1436, 2011

[40] S. K. Yong, P. Xia, and A. Valdes-Garcia, 60GHz Technology for Gbps WLAN and WPAN. Chichester, United Kingdom: John Wiley & Sons Ltd., 2011.

[41] B. Razavi, "Design considerations for direct-conversion receivers," IEEE Trans. on Circuits and Systems II, 44,(6), 428–435, 1997.

[42] S. Rangan, T.S. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: potentials and challenges," in IEEE, 102(3), 2014, 366–385, 2014

[43] J. G. Andrews, S. Buzzi, W. Choi, S.V. Hanly,A.Lozano, A.C.K. Song and J.C. Zhang, "What will 5G be?" IEEE Journal on Selected Areas in Communications, 32(6), 1065–1082, 2014. DOI: 10.1109/JSAC.2014.2328098

[44] T. S. Rappaport, Y. Xing, G. R. MacCartney, Jr., A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," IEEE Transactions on Antennas and Propagation, Special Issue on 5G, 2017.

[45] A. Gupta and R. K. Jha, "A survey of 5G network: architecture and emerging technologies," Access, IEEE, 3, 1206–1232, 2015.

[46] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter wave enhanced local area systems: A high-data-rate approach for future wireless networks," IEEE Journal on selected areas in communications, 32(6), 1152–1163, 2014.

[47] R. Baldemair, T. Irnich, K. Balachandran, E. Dahlman, G. Mildh, Y. Seln, S. Parkvall, M. Meyer, and A. Osseiran, "Ultra-dense networks in millimeter-wave frequencies," IEEE Communications Magazine, 53(1), 202– 208, 2015.

[48] W. Hong, K. -H. Baek, Y. Lee, Y. Kim, and S. T. Ko, "Study and prototyping of practically large-scale mmWave antenna systems for 5G Cellular Devices," IEEE Communications Magazine, 52(9),63–69, 2014.

[49] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," IEEE Transactions on Microwave Theory and Techniques, 64(7), 2207–2225, 2016

[50] S. Sun, G. R. MacCartney, Jr., and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5G wireless communications," in IEEE International Conference on Communication (ICC),1-7, 2017.

[51] L. Lu, G. Ye Li, A. Lee Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," IEEE Journal of Selected Topics in Signal Processing, 8(5), 2014.

[52] S. Dierks_, W. Zirwas, M. J¨ager, B. Panzner, and G. Kramer, "MIMO and massive MIMO – analysis for a local area scenario," in 23rd European Signal Processing (EUSIPCO)

[53] J. Hoydis, S. T. Brink and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" IEEE Journal on Selected Areas in Communications, Institute of Electrical and Electronics Engineers, 31(2), 160 – 171, 2013

[54] C. K Agubor, M. C. Ndinechi, G. A. Chukwudebe and F. K. Opara, "Performance evaluation of multiple antenna systems with diversity techniques using BER analysis," Academic Research International, 5(5) 2014

[55] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. Wong, F. Tufvesson, V.O. Wall and O. Edfors, "The world's first real-time testbed for massive MIMO: Design, implementation, and validation," arXiv:1701.01161v2 [cs.IT], 2017.

[56] Panzner, W. Zirwas, S. Dierks, M. Lauridsen, P. Mogensen, K. Pajukoski and D. Miao, "Deployment and implementation strategies for massive MIMO in 5G. in Proc. Globecom Workshops (GC Wkshps), 346-351, 2015, DOI:10.1109/GLOCOMW.2014.7063455

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 145-149 (2019)

www.astesj.com

Special Issue on Recent Advances in Engineering Systems

ASTESJ

ISSN: 2415-6698

Transfer Function Analysis of Fractional-Order Three-Dimensional Electrically Coupled Cell Network

Mahmut ÜN*

Department of Electrical-Electronics Engineering, Faculty of Engineering and Architecture, İstanbul Yeni Yüzyıl University, Istanbul 34010, Turkey, mahmut.un@yeniyuzyil.edu.tr

A R T I C L E I N F O

Article history:
Received: 13 December, 2018
Accepted: 19 May, 2019
Online: 10 June, 2019

Keywords:
Fractional-order
Transfer function analysis
Impedance characteristics
Cell network
Comparative analysis

A B S T R A C T

In this paper, a novel method is proposed for the dynamic analysis of fractional- order three-dimensional electrically coupled cell network. In general, three- dimensional cell network is constructed by combining three one-dimensional circuit networks. Analysis method is based on the principles of the dynamic analysis with transfer function approximation. Although fractional-order three-dimensional circuit network model contains nonlinear fractal elements such as fractional-order capacitors and inductors, transfer function approximation is employed in dynamic analysis of the network by using the Laplace transform.

First by using nodal analysis method, general expression in matrix forms for the transfer function and typical equivalent impedance of the fractional-order three-dimensional circuit network are derived in fractional domain. Transfer function and Equivalent network impedance of the cell network model are obtained in the form of matrix equation with the implicit analytical expression. Secondly the effects of five network parameters such as inductance L, capacitance C, number of cell unit n and fractional- orders (α, β) on the impedance and electrical network characteristics such as transfer function and output responses are investigated by means of MATLAB Simulation programs. Finally, the validity of the proposed method is done by using PSPICE simulations which show the experimental performance and PSPICE simulation results is presented.

1. Introduction

Due to the use of electrical and non-electrical systems for modeling, in recent years circuit networks have become more attractive [1,2]. A study on the resistance network of the grapyhene showing the presence of a planar circuit network in nature received the Nobel Prize in Physics in 2010 [3-5]. In the last few decades, researchers have published publications on fixed-coefficient circuit networks. These publications focus on the analysis of circuit networks consisting of single-member resistors and capacities with constant coefficients [6,7].

The fractional order mathematical models developed for inductances and capacities have more accurate electrical characteristics for the better realization and best fit. Especially actual inductors and capacitors are in fractional order elements in nature [8]. In recent years, some researchers have been working on the realization and design of fractional elements [9-11]. In addition, in recent years, researchers have focused on fractional order theory [12 -15]. However, few researchers have dealt with the issue of electrical characteristics of multidimensional circuit networks. Therefore, we focus on the equivalent network impedance and transfer function of the three-dimensional cell networks in the fractional domain.

Recently, some researches concentrate on the fractional-order three-dimensional circuit network [16]. Moreover, there has been little research on the transfer function and input equivalent impedance characteristics of the three-dimensional circuit network in the fractional order sense with different cases. That is why, we concentrate on the electrical characteristics of the fractional-order three- dimension network, which may lay the bases for both the electric circuits and dielectrics communities [17].

*Mahmut ÜN, Electrical-Electronic Engineering Department, İstanbul Yeni Yüzyıl University, 34010 İstanbul, Turkey Email: mahmut.un@yeniyuzyil.edu.tr

This study is organized as follows: Basic definitions of the fractional order capacitors and inductors are introduced in Section 2. Also the general expressions of the transfer function and equivalent input impedances of the circuit-network are derived by using the nodal analysis method. MATLAB and PSPICE simulations of the fractional order three- dimensional network are done and simulation results is given in Section 3. Finally the paper is concluded in the last section.

2. Fractional-Order Three Dimensional Circuit Network Model

Electrical networks consisting of one-dimensional and two-dimensional RC elements mentioned in the literature have been proposed and used. Since there are no fractional elements in these electrical networks, the models have been insufficient to precisely characterize the system. RLC elements are used in electrical equivalent circuits in one-dimensional and two-dimensional networks used in the dynamic analysis of electrical interactive cells [18]. Three dimensional random RC networks have been proposed for modeling electrical circuits and dielectric material communities [19]. Nodal-matrix analysis was used to model the three-dimensional network of dispersed capacitance and resistance elements for the multiple bound layers [20]. This article shows that large three-dimensional RC networks can be modeled in terms of performance using fractional order models and integrals. In the fractional domain, the fractional order 2 × n RLC circuit network model has been proposed to examine the impedance characteristics of the circuit network [21]. Electrical RC equivalent circuits are proposed to examine the behavior of the human body impedance against contact currents in a wide frequency range. With the proposed fractional-order three-dimensional RLC networks, expressions for calculating the typical equivalent impedances of the network using the differential equation model and the matrix transform method for different cases in the fractional domain are derived [22].

The circuit diagram of fractional-order three-dimensional electric circuit network used in this study is given in Figure 1.

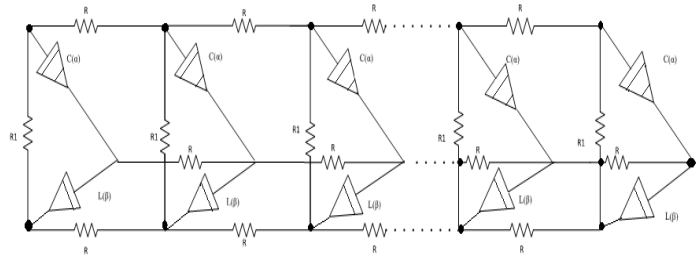


Figure 1: RLC circuit network diagram

In general, this circuit network is obtained by combining three one-dimensional circuit networks. There are n cells in the network. The electrical coupling between the cells is indicated by the resistances R. The cells were modeled with resistance (R_1), fractional- order inductance (L_β) and capacities ($C\alpha$). Fractional impedance values of capacitance and inductance are given below respectively.

$$Z(C_\alpha) = 1 / (C\, s^\alpha) \quad (1)$$

$$Z(L_\beta) = L\, s^\beta \quad (2)$$

Fractional-order three-dimensional cell network is modelled with lumped parameters passive elements such as resistance, fractional-order inductance and capacitance. Thus, this network is composed of passive elements and is a passive network.

In calculating the characteristic values of the circuit network, such as transfer function and the equivalent impedance, the well-known node voltages analysis in circuit theory is used in the fractional domain. For the circuit network depicted in Fig. 1, the equation of the node voltages in matrix form in the fractional domain is written as

$$\begin{bmatrix} Y & -IG & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ -IG & Y_1 & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} \\ & & \vdots & & & & \\ & & \vdots & & & & \\ & & \vdots & & & & \\ \underline{0} & \underline{0} & \underline{0} & \cdots & -IG & Y_1 & -IG \\ \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} & -IG & Y \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ \vdots \\ \vdots \\ V_{n-1} \\ V_n \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ \vdots \\ \vdots \\ I_{n-1} \\ I_n \end{bmatrix} \quad (3)$$

Here, the I_1 (3 × 1) indicates the source currents connected to the nodes and the Vi (3 × 1) indicates the node voltages. Vi = $[V_a, V_b, V_c]^t$ shows the node voltages for each cell, where a,b,c nodes are terminals of abc triangle, I (3 × 3) unit matrix, G = 1 / R is resistance conductivity. Elements of the node admittance matrix are;

$$Y = \begin{bmatrix} s^\alpha C + G + G_1 & -G_1 & -s^\alpha C \\ -G_1 & \frac{1}{s^\beta L} + 2G + G_1 & -\frac{1}{s^\beta L} \\ -s^\alpha C & -\frac{1}{s^\beta L} & s^\alpha C + \frac{1}{s^\beta L} + G \end{bmatrix} \quad (4)$$

$$Y_1 = \begin{bmatrix} \frac{1}{s^\alpha C} + G + G_1 & -G_1 & -\frac{1}{s^\alpha C} \\ -G_1 & s^\beta L + 2G + G_1 & -s^\beta L \\ -\frac{1}{s^\alpha C} & -s^\beta L & s^\beta L + \frac{1}{s^\alpha C} + 2G \end{bmatrix} \quad (5)$$

From (3) and (4) and (5) ones can obtain (6) in the matrix form.

$$AV_d = I \quad (6)$$

The node A (3n × 3n) in size is the admittance matrix and is in the form of a symmetric band matrix. It illustrates the vector of the V_d (3n × 1) node voltages vector and I (3n × 1) dimension of the node voltages vector and I (3n × 1) the current vector. If the vector of the node voltages is resolved from the equation (6),

$$V_d = A^{-1} I \quad (7)$$

is obtained. When a single source is applied to the network, the equivalent driving point impedances and transfer functions for various situations can be found from equation (7) by using the basic definition.

3. Computer Simulations

The element and parameter values selected for the circuit network shown in Figure 1; are given as follows. The cell number is taken as n=5 in all computer simulations. The element values are chosen as R_1= 1Ω, R= 0.1 Ω, L= 1 H and C = 1 F.

3.1. MATLAB Simulations

A constant current source I_1 is applied to the Va node and is taken as Va = V_1. If V_1 input voltage is written in terms of node voltages, V_1= [100 000 000 000 000].V_d . The equivalent impedance is calculated as Z = V_1 The equivalent impedance is calculated as Z = V_1. For fractional order terms, the second order approach was obtained using the value of $s^{0.5s} = (5s^2 + 10s + 1) / (s^2 + 10s + 5)$, Z equivalent input impedance of the network was obtained as a rational function in s [23]. The frequency response, step response and impulse response are obtained from Z network impedance function obtained as a rational function. Then MATLAB simulation results for these rational functions are shown in Figure 2, Figure 3, Figure 4, respectively.

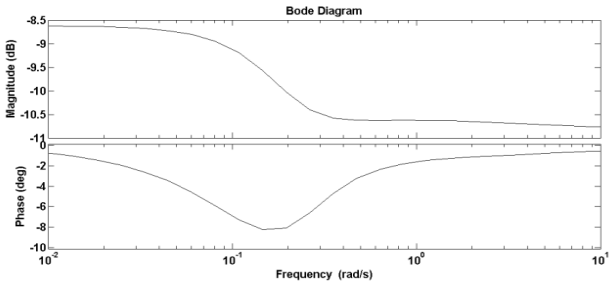


Figure 2: Frequency responses for network input impedance

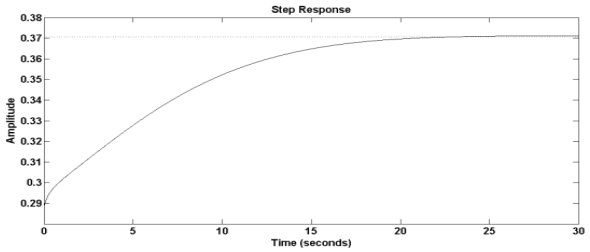


Figure 3*:* Step response of the input impedance

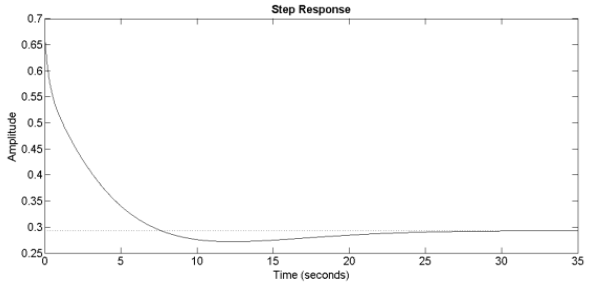


Figure 4*:* Impulse response of the network impedance

Let the output voltage V_o get the input voltage V_1 when the circuit network is excited by a source. The input voltage is V_1= [100 000 000 000 000] .Vd and the output voltage is V_o= [000 000 000 000 001]. Vd. The transfer function (as voltage gain) is defined as TF = V_o / V_1. For fractional order terms, TF is also obtained as a rational function of the TF using the quadratic fractional approach in [24]. The frequency response, step response and impulse response obtained by MATLAB simulations of TF are depicted in Figure 5, Figure 6, Figure 7, respectively.

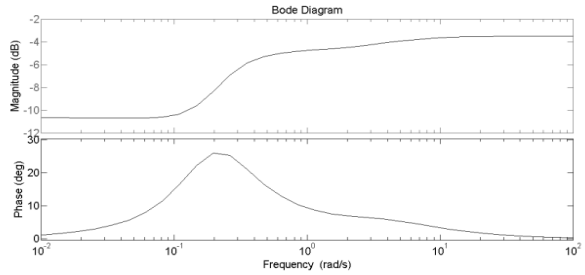


Figure 5*:* The transfer function frequency responses

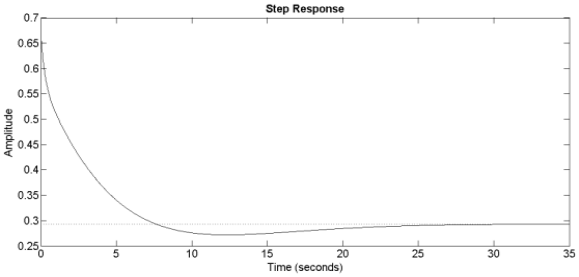


Figure 6*:* The transfer function step response

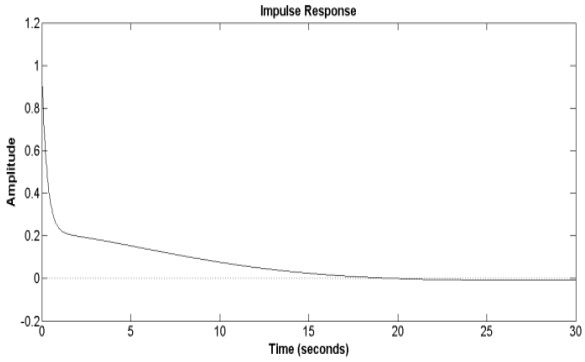


Figure 7*:* The transfer function impulse response

3.2. PSPICE Simulations

The experimental realization of the study was done by PSPICE program using element values taken before in the MATLAB simulations. Developed SPICE model of the network circuit for n=5 cell is presented in Figure 8.

In PSPICE simulations, the quadratic approximation equivalent circuits given in [24] are used instead of fractional capacitor and inductance elements, and the simulation results are depicted in Figure 9, Figure 10, Figure 11, Figure 12, respectively.

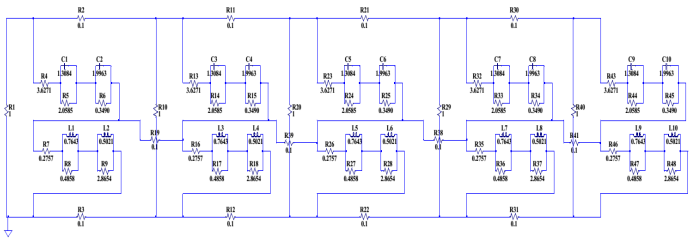


Figure 8: PSPICE model of the cellular network for n=5 cell

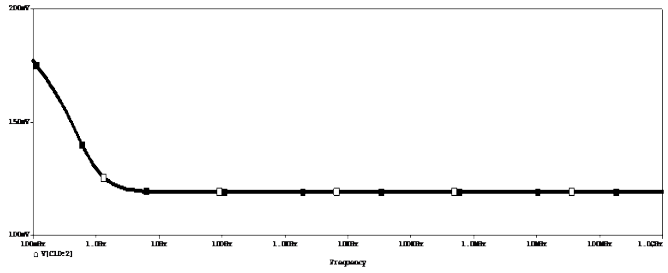


Figure 9: Frequency response of network transfer function with PSPICE

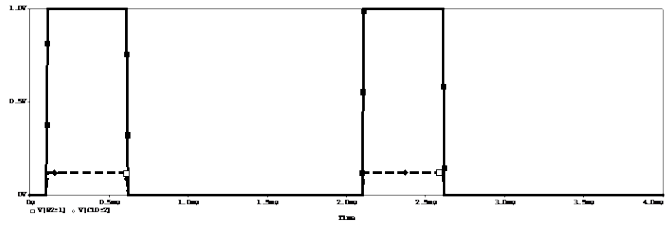


Figure 10: Step response of network transfer function with PSPICE

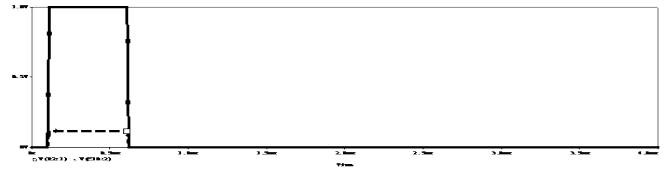


Figure 11: Impulse response for the network transfer function with PSPICE

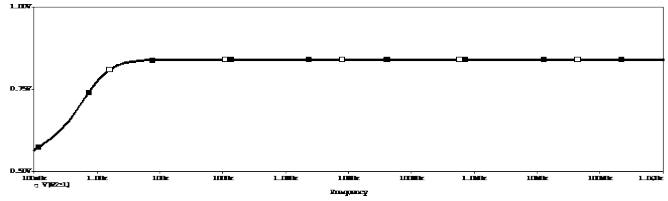


Figure 12: Frequency response of equivalent network impedance with PSPICE

4. Conclusions

In this paper, mathematical models were introduced for the analysis of the fractional-order three-dimensional network in fractional domain. The transfer function and the equivalent input impedance of the fractional-order three-dimensional network were obtained in form of a rational transfer function using rational function approximation for each fractional term. Although the circuit network includes nonlinear fractal elements such as fractional-order capacitors and inductors transfer function approximation method was offered by means of some advanced properties of Laplace transform. A new mathematical method was developed to model and analyze fractional- order three-dimensional circuit network in the fractional domain.

In general, the electrical characteristics of the three-dimensional cellular circuit network were investigated in the fractional domain. Using the nodal analysis method which is well known in circuit theory first transfer function and equivalent input impedance of the fractional order three dimensional circuit network were derived in fractional domain. After that employing second order approach for each fractional order term, equivalent input impedance and transfer function were found as a rational function in s. As a consequence of this process we could be able to calculate transfer function and equivalent input impedance of the network as a rational transfer function.

The electrical system characteristics of the transfer function and the driving point impedance of the fractional- order three–dimensional circuit network were investigated by using MATLAB simulation programs in detail. Moreover network system characteristics such as impulse, frequency, and step responses were graphically obtained for both transfer function and equivalent input impedance. Dynamic characteristics and stability of fractional –order three –dimensional circuit network were also examined with respect to model parameters such as inductance, capacitance, resistances, fractional order values and cell numbers. In order to compare MATLAB and PSPICE simulations the same network parameter values were employed in all simulations. Finally, when PSPICE and MATLAB simulation results are compared graphically, basic similarities between PSPICE and MATLAB simulation results were observed.

References

[1] Von Busse V., Swartz S.M., and C.C, ''Flight metobolism in relation to speed in Chiroptera: Testing the U-shape paradigm in short-tailed fruit bat Carollia perspicillata,'' *J.Exp.Bio.*, 216(11), s. 2073-2080, 2013.

[2] Simpson R., Jaques A., Nunez, C.H. Ramirez, and Almonacid A., ''Fractional calculus as a mathematical tool to improve the modelling of mass transfer phenomena in food,'' *Food Eng.Rev.*, 5(1) s. 45-55, 2013.

[3] Bolotin K.I., Ghahari F., Shulman M.D., Stormer H.L., and Kim p., ''Observation of fractional quantum Hall effect in graphene ,'' *Nature*, 475(7), s. 7354-7358, 2011.

[4] Feldman B.E., Krauss B., Smet J.H., and Yacoby A., "Unconventional sequence of fractıonal quantum hall states in suspended graphene,'' *Science*, 337 (6099), s. 1196-1199, 2012.

[5] Owaidat M.Q., Hijjawi R.S., and Khalifeh J.M., "Network with two extra interstitial resistor," *Int.J.Theor. Phys.*, vol. 10, s. 3152-3159, 2012.

[6] Yao N.Y., Gorshkov A.V., Laumann C.R., Lauchli , J. Ye, and Lukin M.D., "Realizing fractıonal Chern insulators in dipolar spin systems,'' *Phys .Rev.Lett.*, 110(18), 2013.

[7] Hijjawi R.S., Assad J.H., and Sakaji, H.J., "Perturbation of an infinite network of identical capacitors,'' *Int.J. Mod. Phys.*, vol. B2, s. 199-209, 2007.

[8] Carlson G., and Halijaj C., "Approximation of fractional capacitors$(1/s)^{1/n}$ by regular Newton process,'' *IEEE T. Circuit System*, s. 210-213, 1964.

[9] Tenreiro M.J.A., and Galhano A.M.S.F., "Fractional order inductive phenomena based on the skin effect,'' *Nonlinear Dyn.*, vol. 1-2, s. 107115, 2012.

[10] Radwan A.G., and Salama K.N., "Passive and active elements using fractional Lβ,Cα,'' *IEEE T. Circuits-1*, vol. 10, s. 2388-2397, 2011.

[11] Das S., Sivaramakrishna M., Biswas K., and Goswami B., "Performance study of a constant phase angle based impedance sensor to detect milk adultcration. Sensors and actuators,'' *A-Physical*, vol. 2, s. 73-278, 2011.
[12] Krishna M.S., Biswas K., and Goswami B., "Fabrication of a fractional order capacitor with desired specifications; A study on process identification and characterization,'' *IEEE Trans.Electron Dev.*, vol. 11, s. 4067-4073, 2011.
[13] Haba T., Ablart G., Camps T., Olivie F., "Influence of the electrical parameters on the input impedance of a fractal structure reaslised on silicon,'' *Chaos Solitons Fract.*, vol. 2, s. 479, 2005.
[14] Elshurafa A.M., Almadhoun M.N., Salama K.N., and Alshareet H.N., "Microscale electrostatic fractional capacitors using reduced graphene oxide percolated polymer compasites,'' *Appl.Phys.Lett.*, vol. 23, s. 1234-1237, 2013.
[15] Radwan A.G., and Salama K.N., "Fractional–order RC and RL circuit,'' *Circuits System Signal Process*, vol. 6, s. 1901-19015, 2012.
[16] A. Doghri, T.Djerafi, A. Ghotto, and K.Wu, "Substrate integrated wave guide directional couplers for compact three-dimentional integrated circuitsş" IEEE Trans. Microw. Theory Tech., Vol.63, No.1, pp. 209-219, Jan. 2015
[17] R, K, H. Galvao, S. Hadjloucas, K.H. Kienitz, H. M. Paiva, and R.J. M: Afonso,"Fractional order modeling of large three-dimentional RC networks," IEEE Trans. Circuits Syst. I, Reg. Papers, Vol.60, no.3, pp.624-637 , 2013
[18] Chen P., and He S.B., "Analysis of the fractional order paralle tank circuit,'' *J.Circuit System. Comput.*, vo. 6, s. 452-458, 2013.
[19] Merichel M.B., Eyman E.D., and Katar S.B., "Analysis of a Network of Electrically coupled neorans producing Rhytmic Activity in the snail helisoma trivolvis," *IEEE Trans on Biomedical Engineering,* vol. BMM-24, no. 3, s. 277-287, May.1977.
[20] V.Torre, W.G. Owen, G. Sandini, "The Dynamics of electrically interacting cell," *IEEE Trans on Systems, Man and Cybernetics*, vol. 13, no. 5, s. 757-765, September/October 1983.
[21] R. Zhou, D. Chen , Herbert H., C. Lu, "Fractional order 2×n RLC Circuit Network," *Journal.f Circuit Systems and Computers*, vol. 24, no. 9, DOI.10.1142/5021812661550142X(25 pages), 2015.
[22] T.J. Freeborn, A.S. Elwakil, B. Macendy, "Compact wide frequency range fractional order models of human body impedance against contact current," *Hindawi Publishing Corporation, Mathcmatical Problems in Engineering*, vol. 2016, Article ID. 4567937.
[23] R. Tanudar, S. Kumar, "Analysis &Design of Fractance based fractional order filter," Int Journal of Innovative Research in Electrical," *Electronics, Instrumentation and Control Engineering*, vol. 1, Issue. 3, June 2013.
[24] Radwan A. G. And Salama K. N.,"Fractional-order RC and RL circuit", Circuits Sys. Signal Process, Vol:6, 1901-1915, 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 150-155 (2019)

www.astesj.com

Special Issue on Advanced Electrical and Communication Technologies

ASTESJ

ISSN: 2415-6698

Modeling of Grid-Connected Photovoltaic System Installation in Moroccan Ibn Tofail University

Maroua Bouksaim*,[1], Yassin Acci[2], Mohamed Nabil Srifi[1]

[1] *The Electronics and telecommunications Research Group ENSA, Ibn Tofail University of Kenitra, 14000, Morocco*

[2] *Laboratory of Electrical Engineering and Energy Faculty of Science, Ibn Tofail University Kenitra, 14000, Morocco*

ARTICLE INFO

Article history:
Received: 28 March, 2019
Accepted: 18 May, 2019
Online: 10 June, 2019

Keywords:
Photovoltaic system
MPPT controller
Boost converter
Inverter
Three phase grid

ABSTRACT

The main importance of solar photovoltaic energies research is to meet the many environmental demands of the energy challenge, to maximize power and to reduce the costs of photovoltaic (PV) systems to reply the energy needs of population. In the present research work, the first objective is to study the performance and the output energy can be produced by a photovoltaic panel installed in the parking of Ibn Tofail University at Kenitra - Morocco. These provide the basis for developing a simple and efficient model for the PV panel electrical behavior. As the output powers of photovoltaic system are influenced of the solar irradiances (G) and cell temperature (T); the effects of varying the two variable factors, series and shunt resistances, and partial shading on the output of the PV system are presented. Then, the PV system is connected to three phases grid using Boost converter controlled by the perturb and observe technique of Maximum Power Point Tracking and using the DC-AC inverter. The Matlab/Simulink software is used to model the system and to show the simulations result.

1. Introduction

Up to date it is known that, the renewable energies must be developed, so that it can meet the energy needs which are in ascending order continuously [1]. The use of these green energies sources plays a very important role in keeping the planet clean as possible of pollution [1]. Specially, the application of solar PV energy has increasing in different fields [2]. Morocco intends to exploit this clean and inexhaustible energy on a massive scale over the next ten years (Raphaëlle Grouix-Monvoisin, CIFE, 2015). In this work, a PV system installed in the parking of the Ibn Tofail University in Morocco was chosen, to study its performance and to make the system as competitive as possible. So, the objective of this work is to present a design of a boost converter controlled by perturb and observe method of MPPT to extract the maximum power whatever the variations of sunshine. To do this, the work is divided in four parts. The first part describes the modeling of the photovoltaic system consists of a photovoltaic generator, boost converter and MPPT controller. The second one describes the modeling of the PV grid-connected system; this section is divided into two steps; the first is a boost converter controlled by MPPT which allows the PV system to operate in its best condition. The second step is a DC-AC converter that allows a connection to the grid [3]. The output current has to be sinusoidal and in phase with the grid voltage. The boost converter and the inverter operate independently to facilitate the system control [4]. The third part of this work gives the results of the simulations. Last section presents and discusses the results.

*Corresponding Author: Maroua BOUKSAIM, B.P 242 Kenitra, 14000, Morocco. Contact No: 00212662180047. Email: maroua.bouksaim@uit.ac.ma

2. Methodologies

In this section of this search work, the different components of the system studied are presented and modeled. The figure bellow shows the block diagram of the proposed system

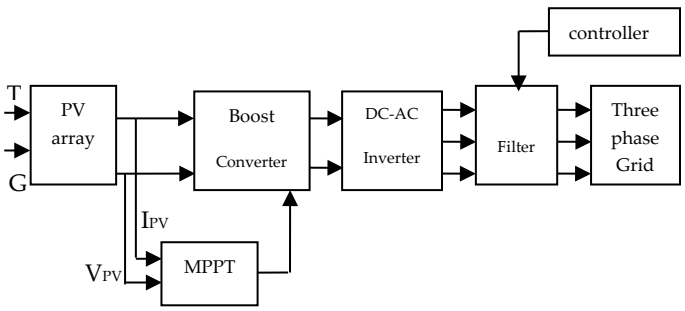


Figure 1: Block diagram of the proposed system

2.1. Photovoltaic panel model

A PV module is PV cells connected in parallel, which are defined by a p-n junction in a film semiconductor that produce electricity from solar radiation [1].

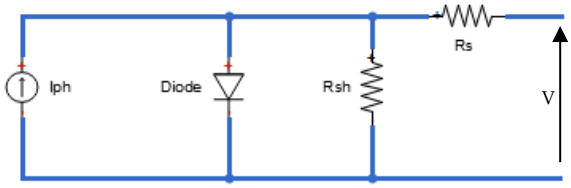


Figure 2: Equivalent circuit of PV panel with single diode

The Kirchoff's law is applied to get the equation below:

$$I = I_{ph} - I_D - I_{RS} \quad (1)$$

Knowledge that:

I_{ph} : The photocurrent, influenced by illumination, its expression is following :

$$I_{Ph} = [I_{SC} + K_I \times (T - T_r)] \times \frac{G}{T} \quad (2)$$

I_D : The current through the diode, its equation below :

$$I_D = I_0 e^{\frac{q(V+Rs\times I)}{nKT}} - 1 \quad (3)$$

I_{RS} : The current flowing in the resistor R_{Sh} ,it's given by :

$$I_{RS} = \frac{V + I \times R_s}{R_{Sh}} \quad (4)$$

The voltage-current characteristic of the solar cell is given below:

$$I = I_{ph} - I_0 e^{\frac{q(V+Rs\times I)}{nKT}} - 1 - \frac{V+I\times R_S}{R_{Sh}} \quad (5)$$

And knowledge that:

I_0 : The dark saturation current, influenced by temperature.

I_{SC}: The current short circuit

q : The charge of an electron is equal to 1.6×10^{-19} C.

K : The constant of Boltzmann is equal to $1.38.\ 10^{-23}$J/K.

T : The ambient temperature, in Kelvin.

R_S : Cell series resistance

R_{Sh} : The cell (shunt) resistance

G : The illumination in W/m^2.

V_{CO}: Open circuit voltage (I=0).

In this paper, PV panels installed in the parking of Ibn Tofail University in Morocco were studied based on the electrical characteristics values shown in the table 1; their reference is Jinko JKM280M-60-DV.

Table 1: Values of electrical characteristics of panel used

Electrical characteristics	Values
STC(Standard Test Conditions) Power Rating	280 W
PTC Power Rating	302.21 W^{-1}
STC Power	15.8 W/ft^2(170.2W/m^2)
Peak Efficiency	17.02%
Power Tolerance	0% /+3%
Number of cells	60
Nominal Voltage	Not applicable
I_{mp}	8.81A
V_{mp}	31.8V
I_{sc}	9.49A
V_{OC}	38.6V
NOCT	45°C

The PV panel was modeled in Matlab/Simulink software using the values of the electrical characteristics presented in the table 1, with the previous equations at Standard Test Conditions (STC), when illumination is equal to G= 1000 W/m^2 at temperature T=25°C.

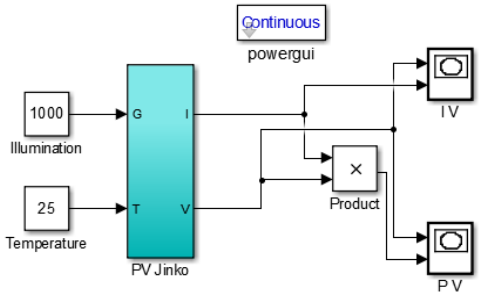


Figure 3: The model of PV panel in Matlab/Simulink

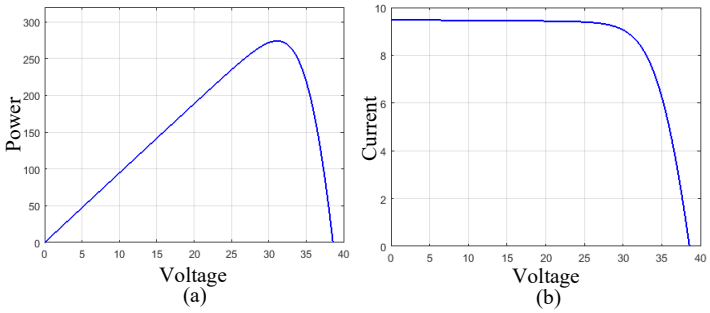


Figure 4: Power-voltage and power-current waveforms of photovoltaic generator; (a) power-voltage waveform of photovoltaic module at STC conditions; (b) current-voltage waveform at STC conditions

In this part of work, the current produced by the cell is almost proportional to the illumination G. But, the voltage V across the junction varies because it is a function of the potential difference at the NP junction of the material itself [5]. However, the open circuit voltage decreases only with irradiation. As a result, the optimum power of the cell is almost proportional to the illumination and its point is at approximately the same voltage of different values of irradiation [6]. The output powers of PV system are influenced of the solar irradiances and cell temperature [7]; and, the effects of varying the two variable factors are presented in the figure 4.

2.2. Influence of illumination and temperature

To treat these important parameters, the simulations on the figure 5 are done to obtain the current-voltage and power voltage curves of the PV field for different values of illuminations and temperatures.

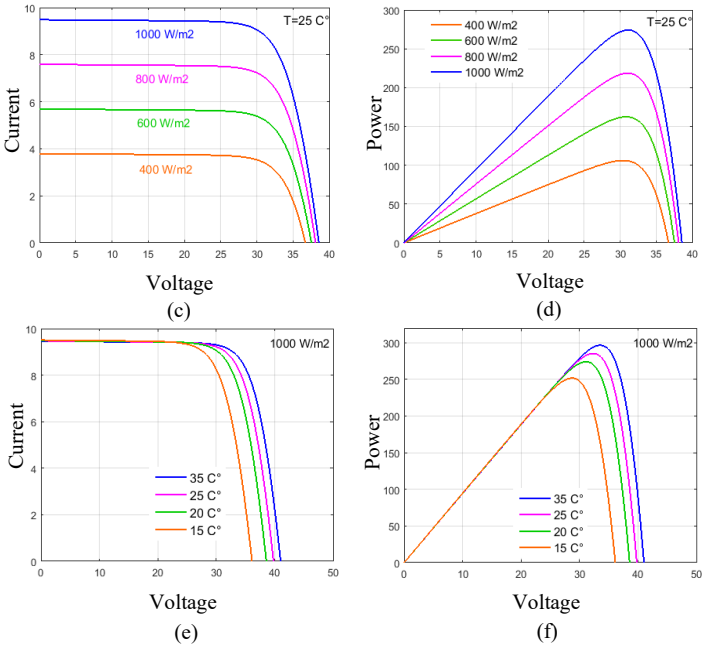


Figure 5: (c) Current-Voltage waveform of a Photovoltaic panel by changing illumination (d) Power-Voltage waveform of a Photovoltaic panel by changing Illumination; (e) Current-Voltage waveform of a Photovoltaic panel at different values of temperature; (f) Power-Voltage waveform of a Photovoltaic panel at different values of temperature

It is noticed that, by changing the illumination, it is observed that when G = 400 W/m2, 600 W/m2, 800 W/m2 and 1000 W/m^2, the respective maximum powers of the PV field are 105 W, 165 W, 220 W and 280 W. The maximum power of the PV field increases with sunshine respectively. It is even for the current that believes with illumination.

With changing the temperature, and G=1000 W/m^2 ; it is observed that when T = 15 ° C, 20 ° C, 25° C and 35 ° C the maximum voltages of the PV field are respectively 36 V, 38 V, 39 V and 41 V, and the respective maximum powers are 250 W, 270 W, 280 W and 300 W. Then, the tension decreases as the temperature increases. It's the same for power maximum that decreases with temperature. We notice that the maximum power of the PV panel at STC conditions is equal 280 W which is compatible with the maximum power P_{MPP} [8].

2.3. Boost converter model

A BOOST converter is a DC to DC voltage converter; that converts a DC voltage into another DC voltage from lower to higher value [1-9]. The figure 6 shows the model and components of the converter used.

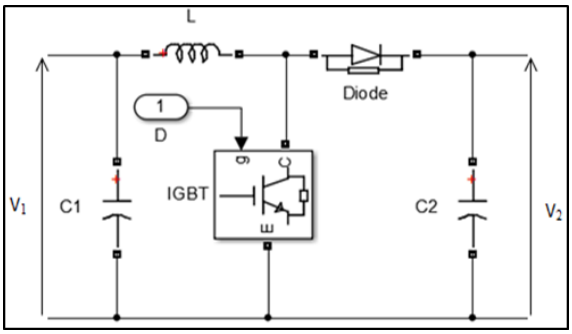


Figure 6: The circuit of DC-DC converter modeled in Matlab/Simulink

This converter operates according to the following equations. The mathematical expression between V_1 and V_2 is following [1]:

$$V_2 = \frac{V_1}{1-D} \tag{6}$$

The expressions of inductor and capacitors are following:

$$C = \frac{D \times V_2}{F \times R \times \Delta V} \tag{7}$$

$$L = \frac{D \times V_1}{F \times \Delta I} \tag{8}$$

Knowledge that:

F: Frequency

D: Duty cycle

R: Load resistance

2.4. Maximum Power Point Tracking

It is known that, the Maximum Power Point Tracking (MPPT) is a technique to search and find the optimum maximum power that can deliver an electrical generator; its principle is based on the automatic variation of duty cycle continuously to maximize the power. There are different methods of MPPT, the most known in the literature are: perturb & observe incremental conductance, fractional open circuit voltage, fractional short circuit current, fuzzy logic and artificial neural network. In this paper, the perturb and observe technique of MPPT was applied. The principle of this method is to perturb by decreasing or increasing the cyclic ratio α

and to observe the effect on the power delivered by the photovoltaic generator [10]. Its algorithm is shown in the figure 7.

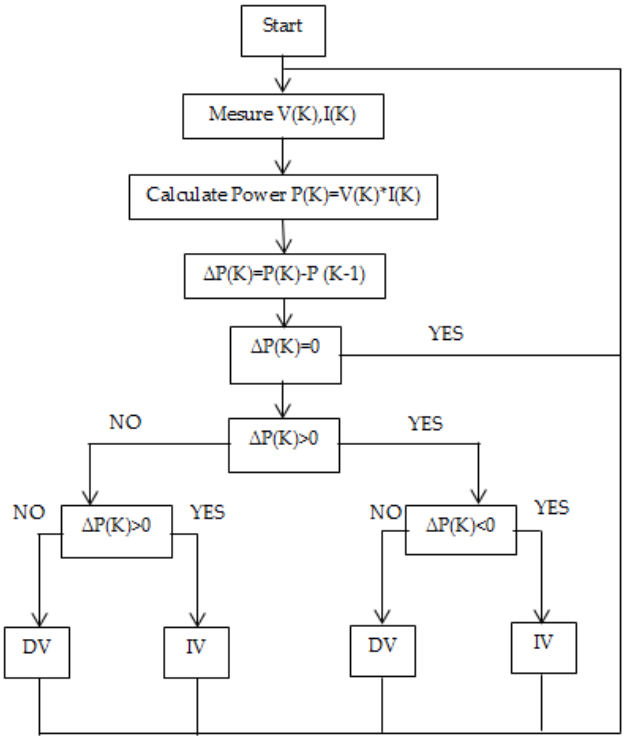


Figure 7: The perturb & observe algorithm

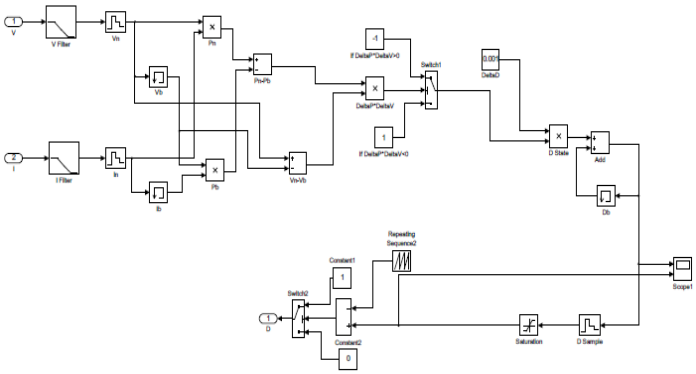


Figure 8 : The model of P&O MPPT in Matlab/Simulink

2.5. Inverter

The principal challenge of PV system connected to grid is to have compatibility between PV arrays and the electricity grid [11]. The inverter converts the direct current of the output of PV array into a synchronizes sinusoidal waveform. The creation of a sinusoid from a DC voltage is obtained through voltage pulses of specific width. This technology uses the Pulse Width Modulation or Pulse width Modulation [12]. The objective of the inverter is converting a DC voltage similar to that of a battery into a single-phase or three-phase alternating voltage similar to that of the electricity grid [13]. It consists of a set of active components (electronic switches) and passive components (transformer) [9]. In this paper three-phase and three level voltage source converters has been used to convert DC power into AC. Three-phase converter symbol is following:

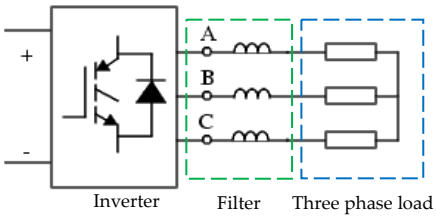


Figure 9: The three-phase inverter proposed

To control the Voltage inverter and reduce harmonics, several switching techniques are used; Pulse width modulation (PWM) technique is the best of these techniques, to produce a sinusoidal variable voltage to the load [14]. It is controlled in closed loop with a PI regulator [15-17].

3. Results

After modeling of the different blocks of the grid-connected photovoltaic system, the figure 10 presents the block diagram of the global model in matlab/Simulink.

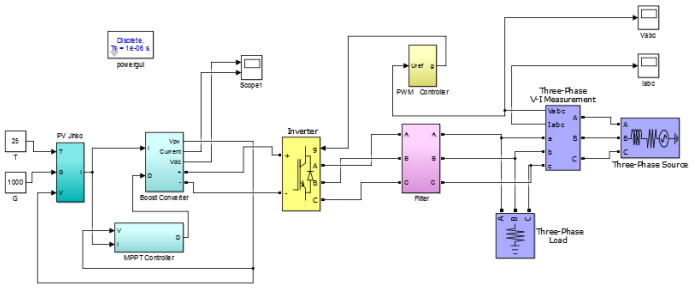


Figure 10: The global model of the system in Matlab/Simulink

This section presents the results of simulations using the STC conditions; the temperature is set at 15°C and illumination is set at 1000 W/m^2.

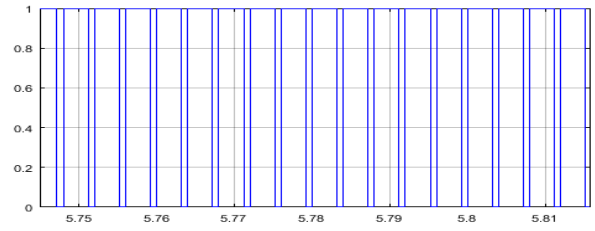
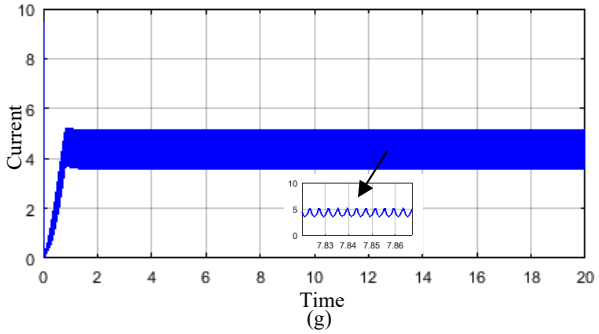


Figure 11: The curve of duty cycle



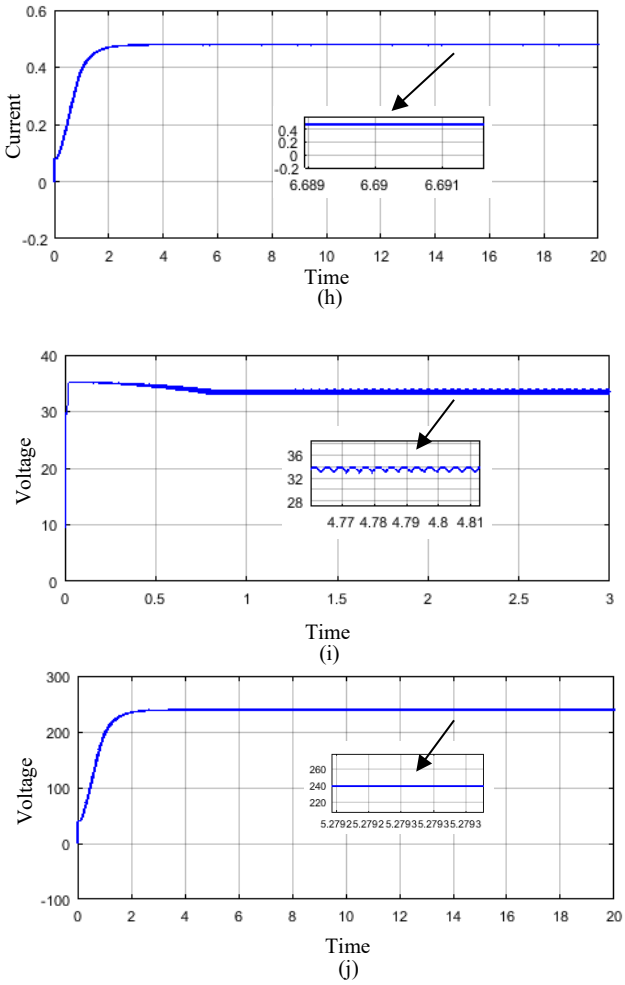


Figure 12: The curves of current and voltage; (g) the current of photovoltaic system before MPPT; (h) the current after MPPT; (i) Voltage before application of Boost converter; (j) Voltage after boosting with the converter

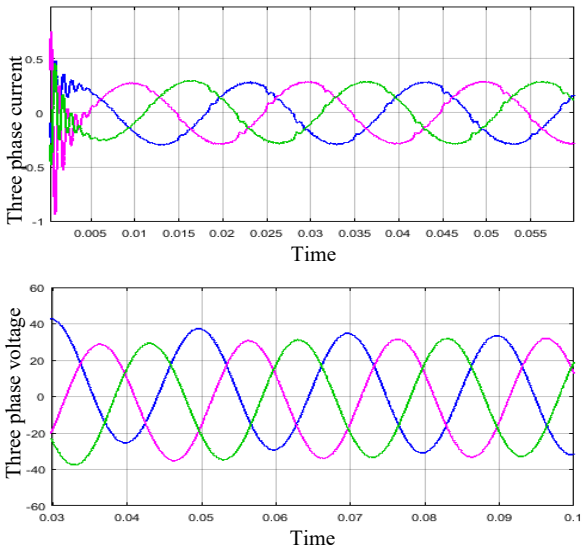


Figure 13: The output of voltage and current with filter

4. Discussions

In this section of the present research work and as shown by the figure 11; the duty cycle is varied using perturb and observe technique of MPPT to generate the best value of voltage for attaining the maximum power.

Then, the figures 12 shows the outputs of voltage and current before application of BOOST converter controlled by MPPT controller; that the voltage was equal to 33.48 V DC; and after boosting, its value was increased from that value 33.48 V DC to 239.7 V DC. That's the objective of applying the DC-DC converter and MPPT controller. Also, it's clearly to see the amelioration of curves without steady-state oscillations.

For the three-phase current and voltage, it is noticed that the figure 13 presents the output of DC current inverted into a synchronized sinusoidal waveform. That's the objective of applying the DC-AC converter (Inverter).

In the present search work, it is observed that, the simulation results of the proposed system are encouraging and as expected, they meet the requirements of the studied system.

In the future, according to these results, the plan is to improve the work by the use of other controllers based MPPT or the combination of two controllers for better performance.

5. Conclusion

This paper presents a study and modeling of photovoltaic system connected to grid by the boost converter (DC-DC converter) controlled by the MPPT controller, to increase, raise the voltage and extract the maximum power point in different variations of irradiation and temperature; and a DC-AC inverter to inverts the continue output current generated by the PV arrays into a synchronizes sinusoidal waveform. The inverter used in this research study is PWM. To search for the maximum power point; Perturb and Observe based MPPT is the technique used in the present paper. The simulation results of the system studied in this paper are encouraging compared to the results obtained in literature. Therefore, in that case our system may be really recommended for practical application in larger electricity generation projects.

References

[1] Maroua Bouksaim, Nissrine Krami, Yassin Acci, Mohamed Nabil Srifi, Abdelkader Hadjouja."Modeling of photovoltaic module using maximum power point tracking controller", in 2018 IEEE International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 2018.

[2] Zainal, Nurul Afiqah, Ajisman, and Ahmad Razlan Yusoff. "Modeling of photovoltaic module using matlab simulink", IOP Conference Series Materials Science and Engineering, 2016.

[3] Fadil,H.E.,F.Giri, and J.M. Guerrero."Grid-connected of photovoltaic module using nonlinear control", 2012 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), 2012.

[4] H. El Fadil, F.Giri. "MPPT and utility power factor achievement in grid-connected PV System Using Nonlinear control", IFAC Proceedings Volumes, 2012.

[5] P.Moraitis, "Review of the Operational Performance of Grid Connected PV Systems", Master's Thesis, Utrecht University, Utrecht, the Netherlands, 2014.

[6] Mau, S.; Jahn, U. "Performance analysis of grid-connected PV systems". In Proceedings of the European Photovoltaic Solar Energy Conference

and Exhibition (EU PVSEC), Dresden, Germany, 4–9 September 2006; pp. 2676–2680.

[7] Mahammad, Abd Kadir, Sharifah Saon, and Wong Swee Chee. "Development of optimum controller based on MPPT for photovoltaic system during shading condition", Procedia Engineering, 2013.

[8] "Proceeding of the 1st International Conference on Electronic Engineering and Renewable Energy", Springer Nature America,Inc,2019

[9] Zainuri, M.A.A.M.; Radzi, M.A.M.; Soh, A.C.; Rahim, N.A. "Development of adaptive perturb and observe-fuzzy control maximum power point tracking for photovoltaic boost DC–DC converter". IET Renew. Power Gener. 2013, 8, 183–194.

[10] Saad Motahhir; Abdelaziz El Ghzizal; Aziz Derouich. "Modélisation et commande d'un panneau photovoltaïque dans l'environnement PSIM" ,2ème Edition du congrès international de génie industriel et management de système, CIGIMS'15,2015.

[11] Antonio Carlos Zambroni de Souza, Miguel Castilla Editors, "Microgrids design and implementation", Cham, Switzerland : Springer, 2019.

[12] Lecture Notes in Electrical Engineering; 2015.

[13] Ghazanfar Shahgholian, Jawad Faiz, Pegah Shafaghi. "Nonlinear control techniques in uninterruptible power supply inverter: A Review", 2009 Second International Conference on Computer and Electrical Engineering, 2009

[14] Nazmul Islam Raju, Md. Shahinur Islam, Ahmed Ahsan Uddin. "Sinusoidal PWM signal generation technique for three phase voltage source inverter with analog circuit & simulation of PWM inverter for standalone load & micro-grid system", INTERNATIONAL JOURNAL of RENEWABLE ENERGY RESEARCH, 2013.

[15] Ramos-Paja, C.; Gonzalez, D.; Carrejo, C. "Predictive control of a photovoltaic dc/dc converter". In Proceedings of the 6th IET International Conference on Power Electronics, Machines and Drives (PEMD2012), Bristol, UK, 2012; pp. 1–6.

[16] Titri, S.; Larbes, C.; Toumi, K.Y.; Benatchba, K. "A new MPPT controller based on the Ant colony optimization algorithm for photovoltaic systems under partial shading conditions". Appl. Soft Comput. 2017, 58, 465–479.

[17] S. Z. Hassan, H. Li, T. Kamal, J. Ahmad, M. H. Riaz, M. A.Khan, "Performance of Different MPPT Control Techniques for Photovoltaic Systems" in 5th International Conference on Electrical Engineering, Pakistan 2018; pp 1-6.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 156-163 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Linearity Improvement of VCSELs based Radio over Fiber Systems utilizing Digital Predistortion

Muhammad Usman Hadi[*,1], Jacopo Nanni[1], Olivier Venard[2], Geneviève Baudoin[2], Jean-Luc Polleux[2], Pier Andrea Traverso[1], Giovanni Tartarini[1]

[1]*Department of Electronic and Information Engineering, University of Bologna, Bologna, 40136, Italy*

[2]*University of Paris-Est, ESYCOM, ESIEE Paris, France, 2 boulevard Blaise Pascal, Noisy-le-Grand, France*

ARTICLE

Article history:
Received:30 April, 2019
Accepted:18 May, 2019
Online: 10 June, 2019

Keywords:
Digital Predistortion
Generalized Memory Polynomial
VCSELs
Radio-over-Fiber

ABSTRACT

The article proposes a Digital Predistortion (DPD) methodology that substantially meliorates the linearity of limited range Mobile Front Haul links for the extant Long-Term Evolution (LTE) and future (5G) networks. Specifically, the DPD is employed to Radio over Fiber links that contrive of Vertical Cavity Surface Emitting Lasers (VCSELs) working at 850 nm. Both, Memory and Generalized Memory Polynomial models are implied to Single Mode (SM) and Multi-Mode (MM) VCSELs respectively. The effectiveness of the proposed DPD methodology is analyzed in terms of Normalized Mean Square Error, Normalized Magnitude, Normalized phase and Adjacent Channel Power Ratio. The demonstration has been carried out with a complete (Long Term Evolution) LTE frame of 10 ms having 5 MHz bandwidth with 64-QAM modulation configuration. Additionally, the effectuality of the proposed DPD technique is evaluated for varying levels of input power and link lengths. The experimental outcomes signify the novel capability of the implied DPD methodology for different link lengths to achieve higher system linearization.

1. Introduction

This paper is extension of the work originally presented in 2018 IEEE International Topical Meeting on Microwave Photonics (MWP) titled "Experimental evaluation of digital predistortion for VCSEL-SSMF-based Radio-over-Fiber link" [1]. The fifth generation (5G) technology is envisaged to provide faster internet access with low latency, cost effectiveness and pervasive mobile coverage [2-3]. Radio over Fiber (RoF) is a pertinent technology to deal with exorbitant requirement of bandwidth multivariate wireless services both for outdoor and indoor scenarios [4-6], and is regarded as a significant technology for the next generation networks [7].

In particular, RoF technology can provide an essential platform for building centralized/cloud radio access network (C-RAN) which should be able to control the centralized base band units (BBU) coming from different base stations and remote radio heads (RRHs) [8].

The interconnectivity of these BBUs with RRHs is economically viable with the distribution network known as the 'fronthaul' [9-10]. The RoF technology is then a suitable candidate for the fronthauling due to its inherent capillary properties. Figure 1 shows the C-RAN utilizing optical fronthauls (OFHs).

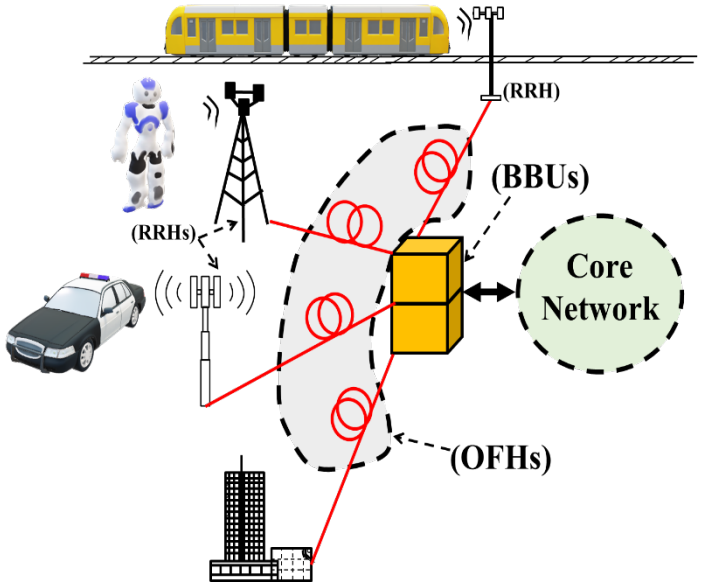


Figure 1: Basic C-RAN architecture showing optical fronthauls (OFHs)

Within these short-medium reach networks, OFHs based on the RoF technology can be realized utilizing Standard Single mode

[*]Muhammad Usman Hadi, Department of Electronic and Information Engineering, University of Bologna, Bologna, Italy, Tel: +390512093951
Email: muhammadusman.hadi2@unibo.it; usmanhadi@ieee.org

Fiber (SSMF) or also Multimode Fiber (MMF) [11]. In both cases, a possible solution which keeps low levels cost and power consumption can be obtained utilizing Vertical Cavity Surface Emitting Lasers also known as VCSELs which emit in the first optical window (wavelength λ=850 nm) as the optical source.

While the use of VCSELs operating at short wavelengths is a relatively consolidated choice finding application e.g. in combinations with Plastic Optical Fibers within in-building networks [12-13] or in combination with MMFs within data centers [14], care must be taken in front of the possible impairments due to the multimodal behavior of SSMFs that operate at the 850 nm [15]. The presence of multimodal behavior would indeed be absent if expensive VCSELs operating at 1.3 μm and 1.55 μm were utilized , in which the major SSMF-related impairments could just be ascribed to optical nonlinear effects and/or to chromatic dispersion [16-17]. However, appropriate countermeasures have been proposed, which can prevent this last impairment cause to be critical [18-19].

Besides the cited OFH, VCSEL-based RoF systems can be found in multivariate scenarios, being utilized for radio astronomic signal transmission [20], multi service indoor wireless signal distribution [21], or machine learning detection [22].

In all these cases, a cardinal issue is represented by the impairments of the opto-electronic devices and particularly the whole VCSEL based Radio over Fiber system. Such nonlinearities, that arise due to relatively stable causes like nonlinear attributes coming from laser diode and perhaps from photodiode [23]. These nonlinearities can also arise due to little dynamics of VCSELs [24], e.g., when high Peak-to-Average Power Ratio (PAPR) signals are transmitted and can have an important role especially for multi-channel transmission. They indeed can cause high in and out of band distortion, which leads to higher interference among near channels.

In order to minimize the nonlinearities in RoF transmission, different methodologies have been proposed that comprises of digital and analog electrical techniques [25-27]. Among them, Digital Predistortion (DPD) has been regarded to be effectual method that leads to reduction in nonlinearities of RoF systems. The primary intension behind this method exploits a digital predistorter (PD) which has inverse and nonlinear profile as that of the Radio over Fiber system. Consequently, when cascaded with the RoF system, it will lead to linearization of the overall cascaded system. DPD linearization technique using memory polynomials based on Volterra series has been applied to Non linear RoF link [26-28]. Similarly, the Digital Predistortion technique based on canonical piecewise-linear (CPWL) function was proposed for intensity modulated/direct detection RoF system [29]. All the aforementioned techniques apply the distortion compensation considering the RoF systems as "black box".

In our previous work reported in [1,30], we demonstrated the novel implementation of DPD implied to considered radio over fiber link which was based on single mode (SM) based VCSEL for LTE applications utilizing Memory Polynomial (MP) architecture. The highlighted results were obtained in back-to-back (B2B) configuration, i.e. utilizing a SSMF fiber patch-cord of a few meters length.

In this paper, the novelty of the current work is summarized as follows: (1) for the first time, the evaluation that was performed in [1] is extended by considering the same SM VCSEL-based RoF system, not in B2B configuration, but equipped with SSMF spans of different lengths; (2) Moreover, also Multimode (MM) VCSEL-based RoF systems are analyzed, utilizing different lengths of MMFs, (3) in addition, to the DPD technique based on Memory Polynomials (MP) utilized in [1], another architecture named as Generalized Memory Polynomial (GMP), is utilized as well, which in general shows an improved effectiveness with respect to MP; (4) the DPD is trained and tested near to threshold to check the efficacy of the proposed technique.

As a performance index evaluation, the link linearity improvements are observed by calculating the Adjacent Channel Power Ratio (ACPR), Normalized Mean Square Error (NMSE), spectral regrowth, Normalized Magnitude (AM/AM) and Normalized phase (AM/PM). The paper is arranged as follows. In Section 2, modeling methodology is highlighted while Section 3 outlines the predistortion learning architecture. Section 4 discusses the estimation algorithm while Section 5 explains briefly the characterization and experimental setup. Section 6 presents the results in detail. Finally, conclusions are drawn in Section 7.

2. Modeling Methodology

The architecture of the predistortion technique is shown in Figure 2 which utilizes Indirect Learning Architecture for the estimation of PD training coefficients [31-34]. Since the statistic of the waveform will remain the same over time, therefore, it can be assumed that system nonlinearity is not time varying and the training of the predistortion can be applied in an off-line practice. This will not only reduce the expenditures and over heads of the predistorter but also avoids the need of an identification algorithm. The identification of DPD is performed in one step. Hence, a linear estimation of PD coefficients result in straightforward identification.

At first, the training phase calculates the predistorter coefficients. At this point, the RoF system output $y(n)$, becomes input $z(n)$ to the Pre-Distorter Training block which is defined as $z(n) = \frac{y(n)}{G}$ where G denotes the gain of the system. The coefficients appraisal can be made using least-squares-based algorithm and is described in Sec. 3. The convergence point of the training is achieved by minimizing the error function. Eventually, the computed training coefficients are employed to the predistorter referred as Digital Pre-Distorter in Figure 2. The predistorter models utilized are derived from the classical Volterra series.

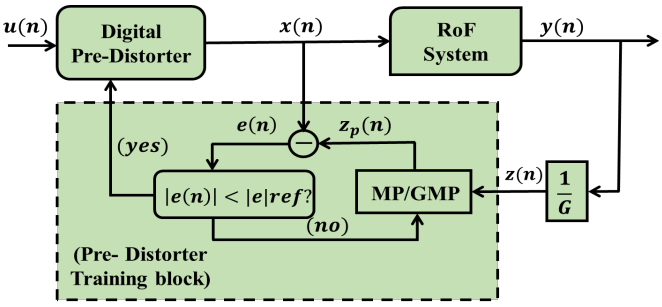


Figure 2. DPD identification for RoF utilizing ILA

3. Digital Predistortion Model

A. Memory Polynomial (MP) Model

Memory Polynomial (MP) model is referred as an inverse non-linear model that has been exploited previously as a powerful model for both, inverse and direct modeling of power amplifier

(PA) nonlinearities. Applying this model for RoF has an additional advantage since memory less model might have problems to characterize the electro-optical (EO) conversion phenomena accurately [34]. The MP model is generally referred as a compromise between memoryless nonlinearity and full Volterra series due to presence of diagonal memory. The output in this case is referred as:

$$z_p(n) = \sum_{k=0}^{K-1} \sum_{q=0}^{Q-1} c_{kq} z(n-q) |z(n-q)|^k$$

Here K represents order of non-linearity, Q is referred as the memory depth, z represents the predistorter input sequence and c_{kq} denotes the model coefficients.

B. Generalized Memory Polynomial (GMP) Model

The use of Generalized Memory Polynomial (GMP) model has been widely utilized for the linearization of PAs [35]. However, GMP has not been yet evaluated for RoF with varying lengths. The GMP model basis functions possess memory for both, for the diagonal terms as well as for the crossing terms i.e., $x(n-q)|x(n-r)|^{k-1}$, where $q \neq r$. The output of the post inverse block $z_p(n)$ modeled with GMP can be expressed as:

$$\begin{aligned} z_p(n) = & \sum_{k=0}^{K_a-1} \sum_{q=0}^{Q_a-1} c_{kq} z(n-q) |z(n-q)|^k \\ & + \sum_{k=1}^{K_b} \sum_{q=0}^{Q_b-1} \sum_{r=1}^{R_b} d_{kqr} z(n-q) |z(n-q-r)|^k \\ & + \sum_{k=1}^{K_c} \sum_{q=0}^{Q_c-1} \sum_{r=1}^{R_c} e_{kqr} z(n-q) |z(n-q+r)|^k \end{aligned} \quad (2)$$

where $z_p(n)$ and $z(n)$ represents the DPD output and input respectively. Similarly, $c_{kq}; d_{kqr}$ and e_{kqr} denotes the complex coefficients for the signal and the envelope; signal and lagging envelope and signal and leading envelope respectively. k represents the index of nonlinearity, and q, r represents the indices of the memory. While K_a, K_b, K_c are the maximum orders of nonlinearity, Q_a, Q_b, Q_c are the memory depths, while R_b and R_c exhibits the lagging and leading delay tap lengths, respectively. GMP has been applied choosing K_a=K_b=K_c= K, Q_a=Q_b=Q_c= Q and $R_b = R_c = R$.

The main objective of the predistorter is to minimize the error ($e(n) = z_p(n) - x(n)$) between the $z_p(n)$ and the input x of the RoF system.

4. Estimation Algorithm

There are number of least squares (LS) algorithm for estimation of model coefficients that take the linear weighting of nonlinear signals [36]. The formulation of the estimation initiates with collecting the coefficients e.g., c_{kq}, d_{kqr} and e_{kqr} in to a $R \times 1$ vector $\boldsymbol{v}$. R represents the total number of coefficients. $\boldsymbol{v}$ means a signal whose time will sample over the same period. Considering (2), coefficients c_{21} denotes the signal $x(n-1)|x(n-1)|^2$. $\boldsymbol{Z}$ characterizes the collection of all such vectors into a Nx R matrix. Once the convergence condition is obtained, the output of the predistorter training block becomes: $z_p(n) = x(n)$ and hence $z(n) = u(n)$. For total samples N, the output can be written in the following way:

$$\boldsymbol{z_p} = \boldsymbol{Zv}$$

Where $\boldsymbol{z} = [z(1),..,z(N)]^T$, $\boldsymbol{z_p} = \left[z_p(1),..,z_p(N)\right]^T$, while $\boldsymbol{v}$ displays a $R \times 1$ vector that contains the coefficients c_{kq}, $d_{kqr}$and e_{kqr}.The LS solution will then become a solution for the equation expressed as:

$$[\boldsymbol{Z}^H \boldsymbol{Z}]\hat{\boldsymbol{v}} = \boldsymbol{Z}^H \boldsymbol{x} \quad (4)$$

The LS solution in (4) should minimize the cost function

$$C = \sum_{n=1}^{N} \left| z_p(n) - x(n) \right|^2 \quad (5)$$

The advanced variations to Volterra series in form of MP, GMP and others increases the effectiveness of predistortion. However, this advanced variation can come in to existence by increasing the memory depth and nonlinearity order. However, this will extortionate the computational complexity and this has to be weighed against other simpler expediencies. Moreover, the performance would generally be more efficient if the predistortion model has higher number of coefficients. This means that while selecting the model and its complexity, a smart tradeoff between complexity and performance can be made accordingly. The effect of increase in complexity will be shown in Sec. 7.

5. Experimental Setup

The experimental test bed utilized for demonstrating and validating the proposed DPD technique is shown in Figure 3. The baseband LTE signal of 5-MHz with 64 QAM modulation format is emulated through a domestic software on MATLAB compliant with 3GPP release TS 36.104 V15.2.0 [37]. The signal is oversampled at 38.4 MSa/s. After this, the sampled sequence of signals passes through the DPD block, which is then RF-transformed at 800 MHz by a Vector Signal Generator (VSG) (see Figure 3) and is then sent to the optical link.

There are two different optical links that are considered for testing DPD sequentially. In the first, a Single Mode VCSEL (Optowell), that operate at 850-nm, which is followed by a fiber span (SSMF) and a PIN photodiode, having 2.5 GHz bandwidth and responsivity factor of 0.6 A/W. The SSMF for varying lengths has been utilized so that the effect of length on the DPD technique can be observed. In the second a 850-nm MM VCSELs in combination with MMFs was tested, maintaining the same photodetector of the other link.

The reason performing the test with reference to the two RoF configurations is related to the fact that each one of them presents aspects that, according to the particular applicative context, may make it preferable with respect to the other. The SM-VCSEL-SSMF based link can indeed feature convenience in terms of cost and flexibility, because of the lower cost per meter of SSMF with respect to MMF, and because of the huge transmission bandwidth of the SSMF. The MM-VCSEL-MMF based link can take advantage of a typically higher dynamic range and

emitted/coupled power of the MM-VCSEL over MMF with respect to the SM-VCSEL over SSMF case.

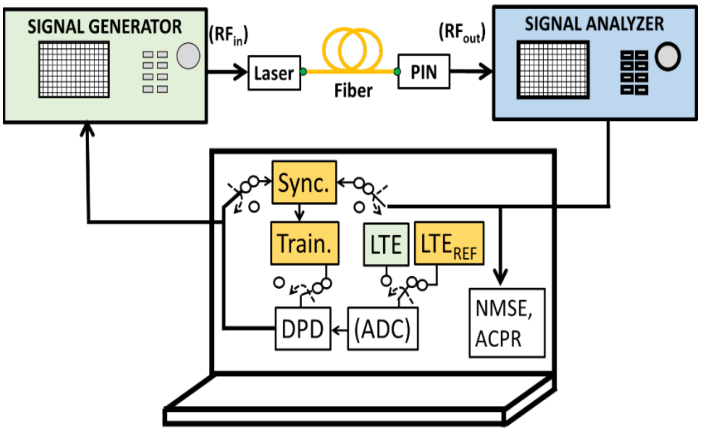


Figure 3. Experimental evaluation comprising of DPD training and application phase.

The reason performing the test with reference to the two RoF configurations is related to the fact that each one of them presents aspects that, according to the particular applicative context, may make it preferable with respect to the other. The SM-VCSEL-SSMF based link can indeed feature convenience in terms of cost and flexibility, because of the lower cost per meter of SSMF with respect to MMF, and because of the huge transmission bandwidth of the SSMF. The MM-VCSEL-MMF based link can take advantage of a typically higher dynamic range and emitted/coupled power of the MM-VCSEL over MMF with respect to the SM-VCSEL over SSMF case.

The process of predistortion is carried out in two steps. The first step is referred as a training phase. In training phase, the parameter identification for predistorter is carried out as referred in Figure 2. During the training phase, reference LTE frames are utilized (see LTE$_{REF}$ block in Figure 3). Primarily, input and output signals are synchronized i.e. time aligned (see Sync. in Figure 3). This block finds the time delay estimation by utilizing the Synchronization Signals (Primary and Secondary) accessible in the LTE framework. Then, the predistorter coefficients are procured through the PD models utilized (see Train. in Figure 3).

In the second step i.e. the testing phase, the training coefficients obtained in previous step are utilized to apply DPD in MATLAB. This means that all switches are turned to opposite position in Figure 3 and the testing is done for LTE frames followed by sampling, pre-distortion, and then uploaded to the Vector Signal Generator (VSG). Finally, they are channelized through the optical link and performance metrics are compared with the one without predistortion. It must be noted that the DPD testing and validation is not only evaluated for the reference LTE signals that were used not only for training, but also for generalized LTE frames.

As mentioned, the quantities utilized to evaluate the effects of the DPD introduced are ACPR and the NMSE. At first, the efficiency of the proposed methods are presented by the ACPR that computes the distortion components outside the useful bandwidth. It is expressed as [1]: (6)

$$ACPR_{dBc} = 10\log_{10}\left[\frac{\int_{ab\_l}^{ab\_u} S(f)\,df}{\int_{ub\_l}^{ub\_u} S(f)\,df}\right]$$

where ab_l and ab_u are the lower and upper frequency limits of the adjacent channel; ub_l and ub_u are the frequency bounds of useful bands. $S(f)$ denotes Power Spectral Density (PSD) of the output signal $y(n)$.

NMSE is defined as follows [27]:

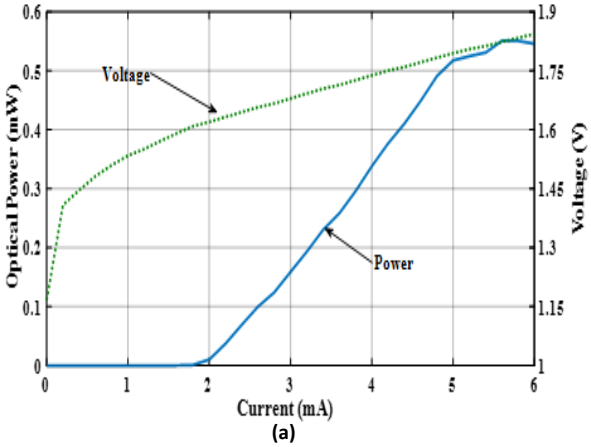
$$NMSE_{dB} = 10log_{10}\left[\frac{\sum_{n=1}^{N}|x(n)-z_p(n)|^2}{\sum_{n=1}^{N}|x(n)|^2}\right] \quad (7)$$

The NMSE value is estimated between $z_p(n)$ and $x(n)$ of the RoF where N embodies the total signal length.

Exploiting these definitions, ACPR and NMSE are calculated for the predistorted and without predistortion output signals. The comparison of the evaluated quantities is presented in the next section. The description of the specifications utilized are given in Table 1.

6. Results and Discussion

The efficacy of the linearization method is appraised and brought into discussion in this section. A primary attention should be given while selecting the parameters of the predistorter such as



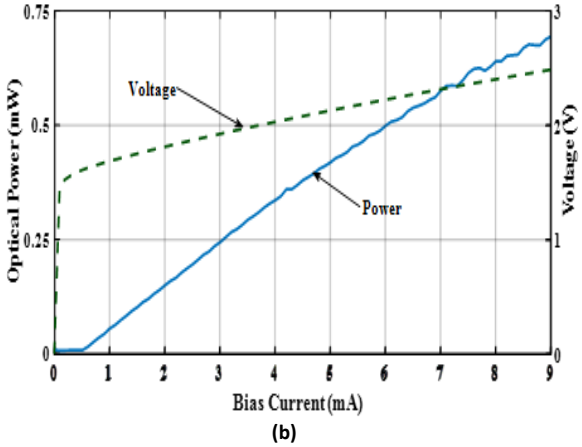


Figure 4. L-I-V characterization utilized in the evaluation of DPD experiments for (a) SM-VCSEL and (b) MM-VCSEL.

memory depth (Q) and nonlinearity order (K) (see Sec. 5 for the details).

Table 1: System Parameters

Optical Link Component	Parameters Values
Laser	**SM-VCSEL** Wavelength= 850 nm $I_{Bias} = 4\ mA$ $I_{Threshold} = 2\ mA$ RIN = -130 dB/Hz **MM-VCSEL** Wavelength= 850 nm $I_{Bias} = 4\ mA$ $I_{Threshold} = 0.8\ mA$ RIN = -125 dB/Hz
Fiber	**SSMF** Maximum Length = 1.5 Km Attenuation = 2.5 dB/km **MMF** Maximum Length = 1 Km Attenuation = 2.5 dB/km
Photo-detector	Responsivity = 0.6 A/W Bandwidth = 2.5 GHz

The L-I-V characteristic curve of SM-VCSEL and MM-VCSEL utilized are illustrated in Figure 4 (a) and (b) respectively. In SM-VCSEL, the maximum saturation current (I_s) is 5 mA while threshold current (I_{th}) is 2 mA. The bias point (I_{bias}) is chosen at 4 mA.

Similarly, In case of MM-VCSEL, the (I_s) is 8 mA and threshold current (I_{th}) is 0.8 mA. The bias point (I_{bias}) is chosen to be 4 mA. The signals having higher PAPR will give rise to higher distortions owing to small dynamics. The PAPR of signal utilized is 9.1 dB.

As explained in Sec. 5, the predistorter complexity is dependent upon the model and order of Q and K chosen. In order to show the trend of this fact, Figure 5 elaborates the experimental NMSE results for different choices of Q and K using both models.

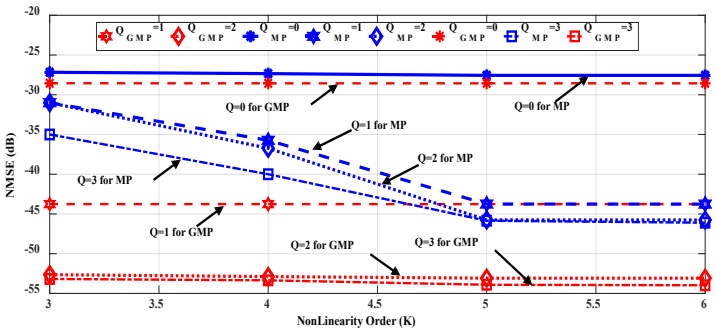


Figure 5. Normalized Mean Square Error results with varying values of memory depth Q and Nonlinearity Order K.

The evaluation in Figure 5 has been done for varying orders of Q (0, 1, 2 & 3) and non-linearity K (3, 4, 5 & 6) by applying a higher RF input power (0 dBm). It can be perceived that GMP results in higher reduction of NMSE in comparison to MP at lower orders of K. From this initial finding, the optimal values $Q = 1,2$ and $K = 3$ are chosen, resulting from a tradeoff between performance achieved and complexity required to the system.

In addition to NMSE, to show the effect of increasing complexity of the proposed distorter, keeping the value of $R = 1$ fixed, Figure 6 represents the ACPR for MP ($Q = 1,2$ and $K = 3$) and GMP ($Q = 1,2$ and $K = 3$) for 1 Km of SSMF. It is perceived that ACPR for MP model doesn't satisfies 3GPP ACPR requirement for MP (both $Q = 1,2$ and $K = 3$) after -5 dBm of input power. While, GMP with $Q = 1$and $K = 3$ satisfies the requirement, however, it is exactly on the borderline at 0 dBm of RF input power.

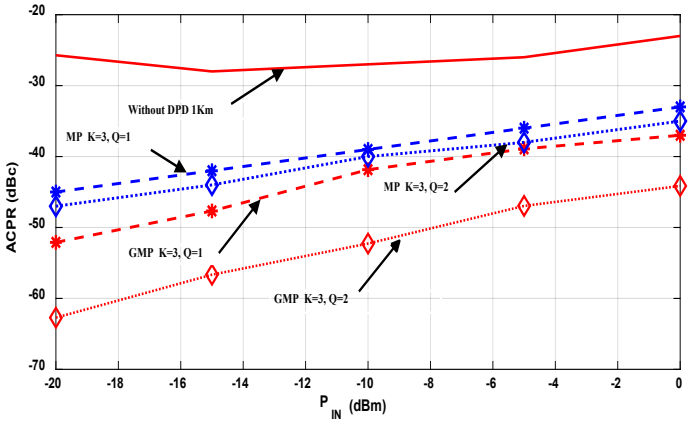


Figure 6. ACPR outcomes vs. input RF power. (Q=1,2, R=1 and K=3 for MP & GMP).

Considering GMP with $Q = 2$ and $K = 3$, the ACPR is well below the requirement set by 3GPP, hence the higher complexity enhances the performance. This demonstration shows that the performance can be enhanced at the cost of higher complexity of the PD model.

Similarly, the Power Spectral Density (PSD) of the received sequence is analyzed when both models are implemented. The length is fixed to 1 Km and RF input power is 0 dBm. The results have been arranged for comparison of MP and GMP with same order of K and Q.

Consider Figure 7(a), the DPD for received output signal is demonstrated for both models by fixing K=3 and Q=1. It is noticeable that GMP has higher reduction in spectral regrowth reduction as compared to MP. Then, in order to highlight the effect of increasing the complexity order, the PSD of GMP and MP are compared for K=3 and Q=2 in Figure 7 (b). This shows that DPD for K=3 and Q=2 is stronger as compared to K=3 and Q=1 due to which spectral regrowth reduction in latter case is higher. Similarly, GMP for K=3 and Q=2 results in ACPR of -44 dBc as compared to -37 dBc when GMP is utilized for (K=3 and Q=1). This substantiates that GMP has better ACPR reduction than MP. This conclusion seems justified due to the fact that GMP carries memory not only in the diagonal but also in the crossing terms.

In light of this discussion and results in the Figure 5, 6 and 7, GMP with optimal values $Q = 2$ and $K = 3$ keeping $R = 1$ fixed have been selected from a tradeoff between performance achieved and complexity required to the system.

In Figure 8 (a) and (b), the ACPR experimental outcomes for several RF input powers (P_{in}) are reported for SM-VCSEL with SSMF and MM-VCSEL with MMF respectively. First consider Figure 8 (a), there are three different cases discussed: Patch cord (Back to back), 1 Km fiber length and 1.5 Km fiber length case. By increasing the length of the fiber, the leakages in adjacent channels rise and therefore the worsening occurs with higher lengths. By employing linearization, it can be seen that fiber length with different lengths have different linearization profiles. For instance, let's consider 1.5 Km length case, the ACPR without linearization at 0 dBm is -18 dBc, and with linearization employed, it is reduced by 22 dBc to -40 dBc.

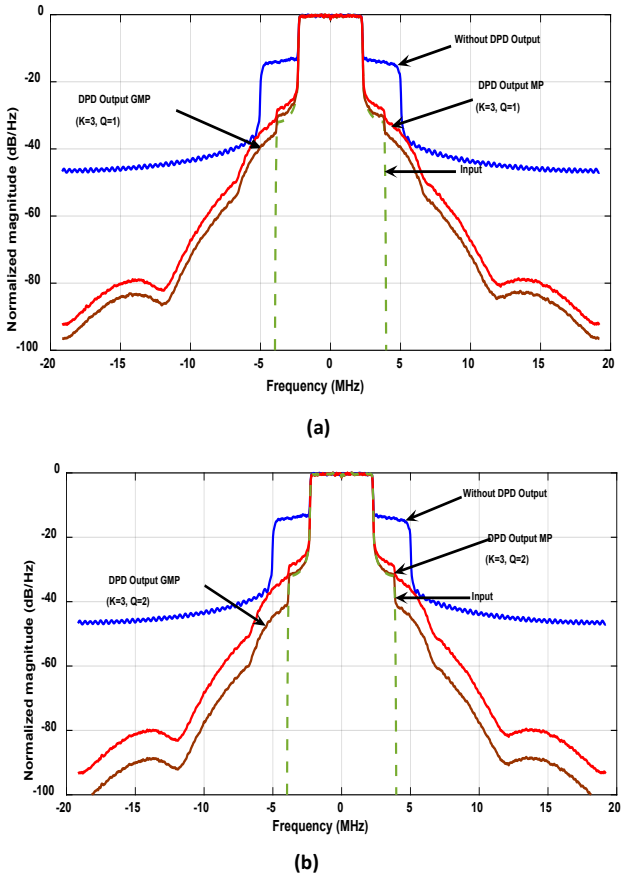


Figure 7. PSD evaluation of output signal without and with DPD (MP/GMP) 0 dBm for 1 Km SSMF with SM-VCSEL having: (a) Q=1, R=1, K=3 and (b) Q=2, R=1, K=3.

In case of Figure 8 (b), four different lengths have been shown up to 1 Km. For 1 Km length, the uncorrected ACPR at 0 dBm is around -22 dBc while the linearization results in a significant reduction to -36 dBc.

It can be deduced from the trend in Figure 8 (a) and (b) that the proposed DPD technique linearizes the length cases as well. Indeed, linearizing the laser nonlinearity is primary aim of the short link lengths, however, in addition to an optical channel consisting in a fiber patch cord (few meters of length), the proposed technique works efficiently for longer fiber lengths.

It must be put into evidence that linearization method has been evaluated for conditions which are very critical such as high PAPR and I_{bias} close to the threshold with optimal values of K and Q. If these critical conditions are relaxed, this will lead to better linearization performance of the predistorter.

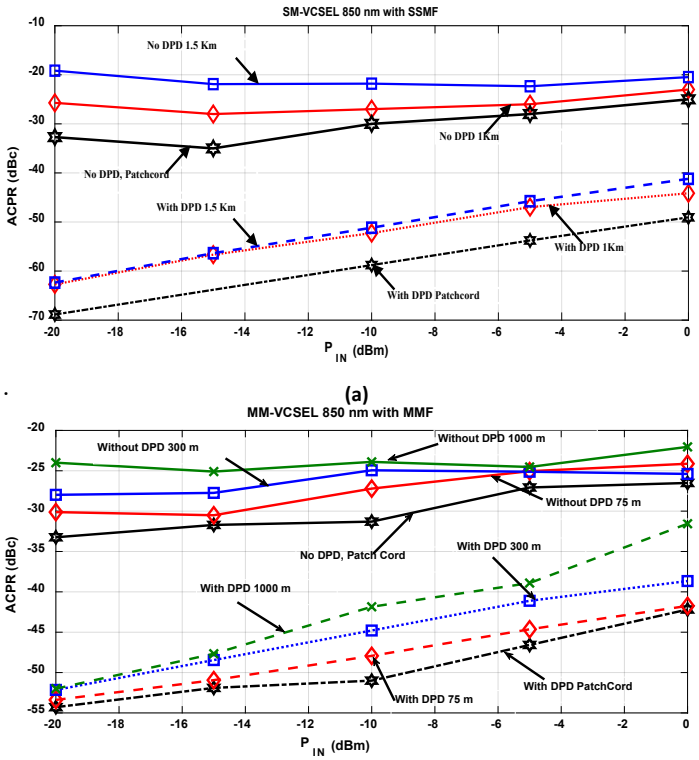
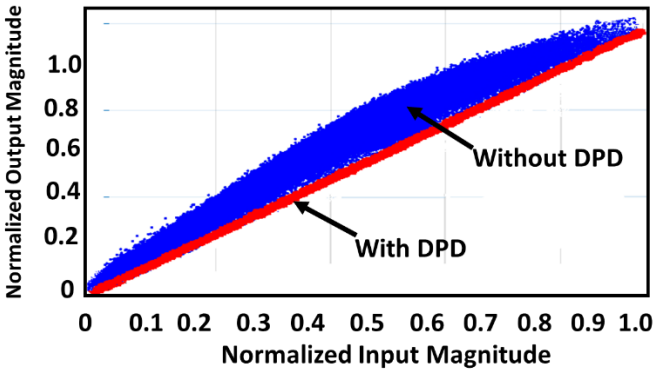


Figure 8. ACPR results vs. varying input signal power using (Q=2 and K=3 for GMP) for (a) SM-VCSEL with SSMF and (b) MM-VSEL with MMF.

Figure 9 represents the AM-AM and AM-PM statistics with and without the linearization employed for the RoF link consisting of 1 Km SSMF and SM-VCSEL. It can be seen that after DPD, the AM/AM is now a straight line and the phase difference is reduced to within 2 degrees.



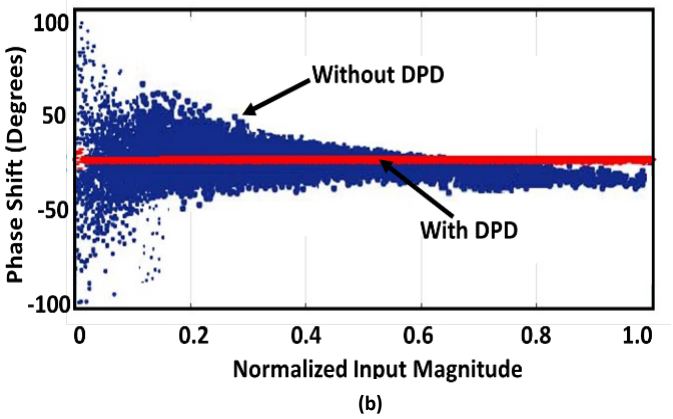


Figure 9. (a) AM-AM and (b) AM-PM plots for SM-VCSEL at 1 Km of SSMF for 0 *dBm* of input RF power with and without DPD.

Table 2 reports the results of the proposed technique, referred to 1 Km of link length with 0 *dBm*. The results are summarized for MP and GMP with K=3 and Q=2 in terms of ACPR and NMSE. The suggested technique linearizes the SM-VCSEL link and MM-VCSEL.

Table 2. Linearization performance for proposed DPD (P_{IN}= 0 *dBm, Link length= 1 Km*)

Model	*SM-VCSEL*		*MM-VCSEL*	
	NMSE (dB)	*ACPR (dBc)*	*NMSE (dB)*	*ACPR (dBc)*
Without DPD	-19.86	-20	-17.63	-22.04
With GMP (K=3, Q=2)	-41.548	-44	-33.18	-36
With MP (K=3, Q=2)	-31.25	-35	-30.14	-29

Indeed, Table 2 suggests that linearization with GMP is better than MP for link length of 1 Km (SM for SM-VCSEL and MM for MM-VCSEL). The reduction in ACPR and NMSE for SM-VCSEL is 22 and 20 dB respectively. While, for MM-VCSEL the improvement in ACPR and NMSE is 14 and 16 dB respectively. This suggests that for SM-VCSEL, the linearization is better than MM-VCSEL for same parameters of data. The improvement for MM-VCSEL is different from SM-VCSEL because the nonlinearities of two lasers are different. It should be noted that MM-VCSEL can achieve higher linearization by selecting appropriate sets of coefficients.

Note that in an applicative scenario, the process of DPD may be positioned at the Central Office (CO), e.g. where BBUs are placed (see Fig. 1) and compensate the RoF Downlink nonlinearities. A periodical re-training of the Digital Predistortion system is in this case necessary, requiring however a negligible time with respect to the time of normal operation of the RoF system. Various methods can be validated for enforcing the periodical training phase. Namely, if a RoF Uplink is applied, the nonlinearities in this case be compensated by a digital post distorter block located at a correspondent CO [38]. By utilizing this method, the RoF downlink nonlinearities can be accumulated at BBU which will decrease the cost and complexity of remote antenna unit.

The time devoted to the periodical re-training of the Digital Predistortion system would in any case remain negligible with respect to the time of normal operation of the RoF system, still allowing its adoption in real applicative scenarios. Similarly, it is possible to utilize an additional photodiode in a base station and feedback the laser nonlinearities and approximating that laser is the main cause of nonlinearity in the RoF link.

Note also that the proposed predistortion method can be applied also to LTE signals of larger bandwidth. The polynomial models nonetheless would require higher values of K & Q with respect to the case presented in the submitted work.

Considering the LTE signal bandwidth and higher modulation format, they would result in a higher PAPR of the transmitted RF signal [39]. At the same time, the increase in bandwidth also determines a correspondent increase in the overall base-band memory of the system to be taken into account by the model.
The polynomial models proposed can still be applied in these operating conditions. They nonetheless would require higher values of the K & Q with respect to the case presented in the submitted work.
These requirements would impact the cost of the DPD implementation, which would be higher due to the higher sampling rate of ADCs, and higher computing capabilities of FPGAs.

7. Conclusions

This article proposes a Digital Pre-Distortion mechanism for linearizing VCSEL based RoF links with different characteristics specifically link lengths. The proposed technique demonstrates a digital predistorter based on MP and GMP. The experiments have been prosecuted for systems based on SM-VCSEL followed by SSMF and on MM-VCSEL followed by MMF. The signal transmitted was a 5-MHz Bandwidth 64 QAM LTE signals and different link lengths have been considered. The performance has been explored in terms of ACPR and NMSE showing that for a link length up to 1 Km, both SM-VCSEL and MM-VCSEL can be linearized in good proportion. Particularly, utilizing SM-VCSEL, GMP results in 22 dBs of reduction in ACPR while MM-VCSEL results in 14 dBs of reduction in spectral regrowth. The results testify an encouraging link performance with low complexity of the predistorter model. It has been demonstrated for the first time that GMPs achieves superior linearization as compared to MP for link lengths up to 1.5 Km.

References

[1] M. U. Hadi et al., "Experimental evaluation of digital predistortion for VCSEL-SSMF-based Radio-over-Fiber link," 2018 International Topical Meeting on Microwave Photonics (MWP), Toulouse, France, 2018, pp. 1-4.
[2] A. Gupta et al., "A Survey of 5G Network: Architecture and Emerging Technologies," IEEE Access, Vol. 3, p. 1206-1232 (2015).
[3] Muhammad Usman Hadi, Hyun Jung, Salman Ghaffar, Pier Andrea Traverso, Giovanni Tartarini, Optimized digital radio over fiber system for medium range communication, Optics Communications, Volume 443, 2019, Pages 177-185, ISSN 0030-4018, https://doi.org/10.1016/j.optcom.2019.03.037.
[4] K. Y. Lau, "RF transport over optical fiber in urban wireless infrastructures," in IEEE/OSA Journal of Optical Communications and Networking, vol. 4, no. 4, pp. 326-335, April 2012.

[5] Y. Shi ; D. Visani ; C.M. Okonkwo ; H. Yang ; H.P.A. van den Boom ; G. Tartarini ; E. Tangdiongga ; A.M.J. Koonen "First demonstration of HD video distribution over large-core POF employing UWB for in-home networks,", 2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference, pp 1-3 (2011)

[6] Y. Shi et al., "Ultrawideband Signal Distribution Over Large-Core POF for In-Home Networks," in Journal of Lightwave Technology, vol. 30, no. 18, pp. 2995-3002, Sept.15, 2012.

[7] Hadi, M., Aslam, N. & Jung, H. (2019). Performance Appraisal of Sigma Delta Modulated Radio over Fiber System. Journal of Optical Communications, 0(0), pp. -Jan. 2019, from doi:10.1515/joc-2018-0227

[8] Hadi, M., Jung, H., Traverso, P., et al. (2019). Digital Radio Frequency Transport over Optical Fiber for 5G Fronthaul Links. Journal of Optical Communications, 0(0), pp. - Apr. 2019, from doi:10.1515/joc-2019-0051

[9] Hadi, M., Hadi, M., Aslam, N., et al. (2019). Experimental Demonstration of MASH Based Sigma Delta Radio over Fiber System for 5G C-RAN Downlink. Journal of Optical Communications, 0(0), pp. -. Retrieved 23 Feb. 2019, from doi:10.1515/joc-2019-0011

[10] J. Nanni et al., "VCSEL-Based Radio-Over-G652 Fiber System for Short-/Medium-Range MFH Solutions," IEEE/OSA J. Lightw. Technol, vol. 36, pp. 4430-4437, Oct.1, 2018.

[11] D. Visani, G. Tartarini, M. N. Petersen, L. Tarlazzi and P. Faccin, "Link Design Rules for Cost-Effective Short-Range Radio Over Multimode Fiber Systems," in IEEE Transactions on Microwave Theory and Techniques, vol. 58, no. 11, pp. 3144-3153, Nov. 2010

[12] D. Visani et al.: Wired and wireless multi-service transmission over 1mm-core GI-POF for in-home networks, Electron. Lett., vol. 47, pp.203-205, Feb. 2011

[13] D. Visani et al., " 3×2N-QAM Constellation Formats for DMT Over 1-mm Core Diameter Plastic Optical Fiber", IEEE Photonics Technology Letters, 23, pp. 768-770 (2011).

[14] J. A. Tatum et al., "VCSEL-based interconnects for current and future data centers," J. Lightw. Technol., vol. 33, no. 4, pp. 727–732, Feb. 15, 2015

[15] J. Nanni et al., "Modal noise in 850nm VCSEL-based radio over fiber systems for manifold applications," Proc. Fotonica AEIT Italian Conference on Photonics Technologies, Turin, 2015, pp. 1-4.

[16] M.Sauer, A. Kobyakov, and A.B. Ruffin, ``Radio-over-fiber transmission with mitigated stimulated brillouin scattering,", IEEE Photonics Technology Letters, vol.19, no.19, pp. 1487--1489, Oct 2007..

[17] G. Tartarini and P. Faccin, ``Efficient characterization of harmonic and intermodulation distortion effects in dispersive radio over fiber systems with direct laser modulation," Microwave and Optical Technology Letters, vol. 46, no. 2, pp. 114--117, 2005.

[18] P. M. Sevila, V. Almenar, and J. L. Corral, "Transmission Over SSMF at 850 nm: Bimodal Propagation and Equalization," *IEEE/OSA J. Lightw.Technol.*, vol. 35, no. 19, pp. 4125–4136, Oct 2017

[19] J. Nanni et al., "Modal Noise Mitigation in 850-nm VCSEL-Based Transmission Systems Over Single-Mode Fiber," in IEEE Transactions on Microwave Theory and Techniques, vol. 64, no. 10, pp. 3342-3350, Oct. 2016

[20] J. Weiss, "Analog Optical RF-Links for Large Radio Telescopes,"IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS), San Diego, CA, 2018, pp. 24-27. doi: 10.1109/BCICTS.2018.8551058

[21] D. Visani et al., "Wired and wireless multi-service transmission over 1mm-core GI-POF for in-home networks," in Electronics Letters, vol. 47, no. 3, pp. 203-205, 3 February 2011.

[22] C. Wang et al., "Machine Learning Detection for DMT Modulated 112-Gbps VCSEL-MMF Optical Interconnection," Asia Communications and Photonics Conference (ACP), Hangzhou, 2018, pp. 1-3.

[23] W. P. Ng, T. Kanesan, Z. Ghassemlooy and C. Lu, "Theoretical and Experimental Optimum System Design for LTE-RoF Over Varying Transmission Span and Identification of System Nonlinear Limit," in IEEE Photonics Journal, vol. 4, no. 5, pp. 1560-1571, Oct. 2012.

[24] S. F. Yu, "Nonlinear dynamics of vertical-cavity surface-emitting lasers," in *IEEE Journal of Quantum Electronics*, vol. 35, no. 3, pp. 332-341, March 1999.

[25] Roselli, L., Borgioni, V., Zepparelli, F., Ambrosi, F., Comez, M., Faccin, P., Casini, A.: Analog laser predistortion for multiservice radio-over-fiber systems. IEEE J. Lightw. Technol. 37(5), 1211-1223 (2003)

[26] Zhu, R., Zhang, X.: Broadband predistortion circuit design for electro-absorption modulator in radio over fiber system. In Proceedings of the Optical Fiber Communications Conference and Exhibition(OFC), San Francisco, CA, USA (2014).

[27] F. Fuochi, M. U. Hadi, J. Nanni, P. A. Traverso and G. Tartarini, "Digital predistortion technique for the compensation of nonlinear effects in radio over fiber links," IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI), Bologna, 2016, pp. 1-6.

[28] A. Hekkala et al., "Predistortion of Radio Over Fiber Links: Algorithms, Implementation, and Measurements," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 59, no. 3, pp. 664-672, March 2012.

[29] Mateo, C., Carro, P.L., Garca-Dcar, P., De Mingo, J., Salinas, I.: Radio-overfiber linearization with optimized genetic algorithm CPWL model. Opt. Express. 25(4). pp. 3694-3708, Feb. 2017.

[30] M. U. Hadi, P. A. Traverso, G. Tartarini, O. Venard, G. Baudoin and J. Polleux, "Digital Predistortion for Linearity Improvement of VCSEL-SSMF-Based Radio-Over-Fiber Links," in *IEEE Microwave and Wireless Components Letters*, vol. 29, no. 2, pp. 155-157, Feb. 2019. doi: 10.1109/LMWC.2018.2889004

[31] M. Abi Hussein, V. A. Bohara and O. Venard, "On the system level convergence of ILA and DLA for digital predistortion," 2012 International Symposium on Wireless Communication Systems (ISWCS), Paris, 2012, pp. 870-874.

[32] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," Signal Processing, IEEE Transactions on, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[33] L. Ding, G. Zhou, D. Morgan, Z. Ma, J. Kenney, J. Kim, and C. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," Communications, IEEE Transactions on, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[34] Hekkala, A., Lasanen, M.: 'Performance of adaptive algorithms for compensation of radio over fiber links'. Proc. IEEE Wireless Telecommunication Symp., Prague, Czech Republic, April 2009, pp. 1–5

[35] D. Morgan, Z. Ma, J. Kim, M. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of rf power amplifiers," Signal Processing, IEEE Transactions on, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[36] L. Guan, and A. Zhu, "Optimized Low-Complexity Implementation of Least Squares Based Model Extraction for Digital Predistortion of RF Power Amplifiers," *IEEE Trans. Microw. Theory Tech*., vol. 60, no. 3, pp. 594-603, Mar. 2012.

[37] 3GPP TS36.104 V15.2.0, Evolved Universal Terrestrial Radio Access (E-UTRA) User Equipment (UE) Radio Transmission and Reception (Release 15), March 2018.

[38] L. C. Vieira and N. J. Gomes, "Experimental demonstration of digital predistortion for orthogonal frequency-division multiplexing-radio over fibre links near laser resonance," in IET Optoelectronics, vol. 9, no. 6, pp. 310-316, 12 2015. doi: 10.1049/iet-opt.2014.0160

[39] M. U. Hadi, "Extending the benefits of LTE to unlicensed spectrum," *2015 International Conference on Information and Communication Technologies (ICICT)*, Karachi, 2015, pp. 1-3. doi: 10.1109/ICICT.2015.7469592

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 164-170 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

An Adaptive Fuzzy-Sliding Mode Controller for The Bridge Crane

Trinh Luong Mien*

Department of cybernetics, Faculty of electrical and electronic engineering, University of transport and communications,

No.03, Caugiay, Langthuong, Dongda, Hanoi, Vietnam.

ARTICLE INFO

Article history:
Received:27 April, 2019
Accepted:27 May, 2019
Online: 10 June, 2019

Keywords:
Crane
Anti-swing
Fuzzy logic
Adaptive control
Sliding mode control

ABSTRACT

This article presents a bridge crane nonlinear dynamic model in 2-dimensional space, and then given a novel adaptive fuzzy-sliding mode controller based on combining sliding mode control with fuzzy logic and Lyapunov function. Firstly, the article proposes an intermediate variable to link signal between two slide surfaces, related to trolley movement and payload swing. Then the fuzzy controller, compensative controller and parameter adaptive update law for the bridge crane are defined. The asymptotic stability of the proposed bridge crane control system is proven based on Lyapunov stability theory. Simulation results show that the adaptive fuzzy-sliding mode controller ensures the trolley follows the input reference with the short settling time, eliminating steady error, and anti-payload swing, anti-disturbance.

1. Introduction

The bridge crane is a lifting machinery which commonly used in industry and transportation to move goods from one location to another. The bridge crane consists of three main components: hoist, trolley and beam, corresponding to three movements: lifting and lowering the cargo of the hoist (mounted on the trolley), horizontal movement of the trolley (on the beam), and vertical movement of the beam (on the factory frame). Goods (refer to as payload) are linked to cargo hoist by the cable and hanger. This connection causes the vibration of the payload when the crane moves. This is undesirable - causing unsafe operation of the crane, reducing the accuracy in controlling the cargo transport of crane, reducing the crane operation.

In some cases of actual crane operation, it is often focused on controlling the exact position of the trolley, after lifting the payload. In this case, the bridge crane is studied as a trolley movement in the 2D space, created by the horizontal movement of the trolley and the verticality of the cable. Common works on position control of the trolley are often based on state feedback control, PID control [1, 2], LQR control [3], PID combined with fuzzy [4-7], PID adjusting parameters [8] and obtain certain results. In [1], the state feedback controller is combined with the integral part to eliminate the steady error. The work [4] presents a controller PID combined with the fuzzy-sliding mode control. These controllers are designed based on the linear model of the crane and the existing of payload oscillation.

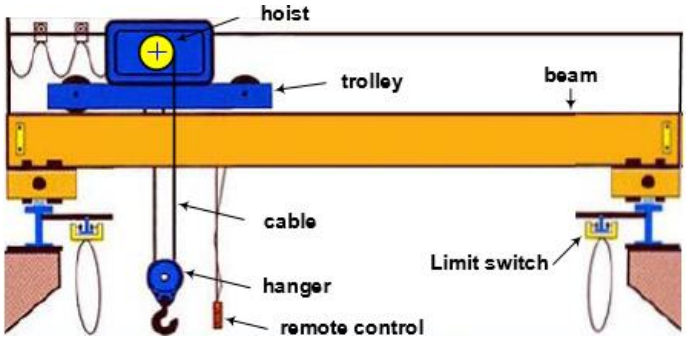


Figure 1. Bridge crane movement in 3D: trolley-0x, beam-0y, hoist-0z.

In recent times, the researches have taken into account the nonlinearity of the crane based on back stepping [9], nonlinear control [10, 11], sliding mode control [12], sliding mode control combined with fuzzy logic [13], partial linear feedback control [14], sliding mode control combined with partial linear feedback [15], sliding mode control combined with adaptive fuzzy control [16], or neural network [17]. These controllers have enriched the crane control strategy, ensuring that trolley follows the reference trajectory with small static error and small payload oscillation. However, these control algorithms are quite complicated, require a large amount of computation, the deployment of them on hardware devices is difficult and not considering the impact of

* Trinh Luong Mien. Tel: +84 904 684 595. Email: mientl@utc.edu.vn

noise in the working environment of the crane, such as the impact of the wind.

This article presents a nonlinear mathematical model of a crane in two-dimensional space (2D bridge crane model). On that basis, an adaptive fuzzy-sliding mode controller is designed for 2D crane model based on the sliding mode control principle combined with fuzzy logic and Lyapunov function, ensuring that the trolley moves quickly, following set reference and eliminating payload oscillation, anti-disturbance. The rest of this article is organized as follows. Section 2 introduces the bridge crane dynamic model. Section 3 deals with the design of controllers for the bridge crane 2D. Section 4 presents the simulation results. Finally, section 5 presents conclusions.

2. Bride crane dynamic model

The motion of the bridge crane in the two-dimensional space *Oxz* as shown in Figure 2, where: *x(t)* is the position of the trolley moving in the *Ox* direction, *α(t)* is the oscillation angle of the payload. Assuming that the cable is a rigid rod and connected with a cargo hanger, the cable length is constant and the cable mass is negligible; ignore the trolley friction when moving; consider payload as the point $P(x_p,z_p)$ with mass m_p; trolley with mass m_t; F_x is the force driving a trolley in the *Ox*-direction, g is the gravitational acceleration.

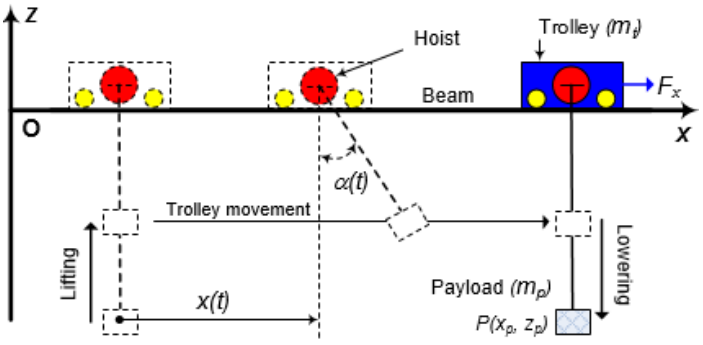


Figure 2. Bridge crane trolley movement in 2D

The dynamic model of bridge crane is defined based on energy conservation law. The total kinetic energy K and the potential energy T of the system are determined as follows:

$$K = K_{trolley} + K_{payload} = \frac{1}{2} m_t \dot{x}^2 + \frac{1}{2} m_p (\dot{x}_p^2 + \dot{z}_p^2)$$
$$; \; x_p = x + l\sin\alpha \; , \; z_p = -l\cos\alpha \qquad (1)$$
$$T = T_{trolley} + T_{payload} = -m_p g l \cos\alpha$$

Lagrange function is determined by the following formula:

$$L = K - T \qquad (2)$$

The mathematical equation describing crane dynamics is determined from the Euler-Lagrange equation, with $q = [x, \alpha]^T$, $\tau = [F_x, 0]^T$, as shown below

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q})\right) - \frac{\partial L}{\partial q}(q, \dot{q}) = \tau \qquad (3)$$

After transforming, we obtain the mathematical equation describing the dynamics of crane:

$$\begin{cases} (m_t + m_p)\ddot{x} + m_p l(\ddot{\alpha}\cos\alpha - \dot{\alpha}^2 \sin\alpha) = F_x \\ \ddot{x}\cos\alpha + l\ddot{\alpha} + g\sin\alpha = 0 \end{cases} \qquad (4)$$

The system of equations (4) is separated into 2 subsystems, describing 2D bridge crane dynamics in the state space, including: system A describes the trolley movement; system B describes the payload oscillation.

$$A: \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f_1(x) + g_1(x)u \end{cases} \qquad (5)$$

$$B: \begin{cases} \dot{x}_3 = x_4 \\ \dot{x}_4 = f_2(x) + g_2(x)u \end{cases} \qquad (6)$$

where:

$$x_1 = x(t); x_2 = \dot{x}(t); x_3 = \alpha(t); x_4 = \dot{\alpha}(t); u = F_x$$

$$f_1(x) = \frac{m_p l x_4^2 \sin x_3 + m_p g \sin x_3 \cos x_3}{m_t + m_p \sin^2 x_3}$$

$$g_1(x) = \frac{1}{m_t + m_p \sin^2 x_3}$$

$$f_2(x) = -\frac{(m_t + m_p) g \sin x_3 + m_p l x_4^2 \sin x_3 \cos x_3}{(m_t + m_p \sin^2 x_3) l}$$

$$g_2(x) = -\frac{1}{(m_t + m_p \sin^2 x_3) l}$$

The 2D bridge crane model is a nonlinear system. When the trolley moves, the payload oscillation will appear, so that as the proposed controller not only ensures the trolley follows the reference trajectory but also quenching payload oscillation, in order to ensure safe operation and accurate cargo transportation.

3. Adaptive fuzzy-sliding mode controller design for the bridge crane

3.1. Sliding mode controller

In this research, mathematical model of 2D bridge crane is described by equations (5) and (6), corresponding to system A, system B.

The goal here is to synthesize a controller $u = u(\underline{x})$ so that the outputs $\underline{x}(t) = [x, \alpha]^T$ of the crane (5) and (6) follows on the input reference $\underline{x}_d(t) = [x_d, \alpha_d]^T$ i.e. the steady error:

$$\underline{e}(t) = [e_1(t), e_2(t)]^T \to \underline{0}$$

with $e_1 = x_1 - x_{1d}$;

$$e_2 = x_3 - x_{3d};\ x_{1d} = x_d; x_{3d} = \alpha_d, \alpha_d = 0.$$

According to the sliding control principle, we determine the general sliding surface for systems A and B:

$$s_{1b} = \dot{e}_1 + \lambda_1 e_1 = \dot{x}_1 - \dot{x}_{1d} + \lambda_1 (x_1 - x_{1d}) = x_2 + \lambda_1 x_1 - (\dot{x}_{1d} - \lambda_1 x_{1d} \quad (7)$$

$$s_{2b} = \dot{e}_2 + \lambda_2 e_2 = \dot{x}_3 - \dot{x}_{3d} + \lambda_2 (x_3 - x_{3d}) = x_4 + \lambda_2 x_3 (8) \dot{x}_{3d} - \lambda_2 x_{3d}$$

where: λ_1, λ_2 are positive constants.

To ensure the control target for bridge crane (5) and (6), we define the sliding surface s_1 for system A to ensure precise control of trolley position and sliding surface s_2 for system B to ensure the general control target for bridge crane, that is: controlling the trajectory of the trolley, quenching the payload oscillation of the payload, the anti-disturbance. Therefore, we need to use the intermediate variable z to link the signal between the sliding surface s_1 and the sliding surface s_2. On that basis, the author defines the sliding surface s_1, s_2 as follows:

$$s_1 = s_{1b} = \dot{e}_1 + \lambda_1 e_1 = \dot{x}_1 - \dot{x}_{1d} + \lambda_1 (x_1 - x_{1d}) = x_2 + \lambda_1 x_1 (9) - \dot{x}_{1d} - \lambda_1 x_{1d}$$

$$s_2 = \dot{e}_2 + \lambda_2 (e_2 - z) = \dot{x}_3 - \dot{x}_{3d} + \lambda_2 (x_3 - x_{3d} - z) = x_4 (10) + \lambda_2 (x_3 - z) - \dot{x}_{3d} - \lambda_2 x_{3d}$$

Intermediate variable z are defined as follows:

$$z = \bar{z}.sat(s_1/\Phi_z) \quad (11)$$

where: $|z| \le \bar{z};\ 0 < \bar{z} < 1$, $\bar{z}$ is the upper limit of |z|, Φ_z is the boundary limit s_1.

$$sat(s_1/\Phi_z) = \begin{cases} sign(s_1/\Phi_z) & if \quad |s_1/\Phi_z| \ge 1 \\ s_1/\Phi_z & if \quad |s_1/\Phi_z| < 1 \end{cases}$$

Because of $\bar{z} < 1$ it can be effective in reducing. When s_1 reduced z is also decreases. When $s_1 \to 0,\ z \to 0,\ e_2 \to 0$ then $s_2 \to 0$ and achieves the control target for crane (5) and (6). Therefore, by adding intermediate variables z now s_1 and s_2 at the same time tends to 0.

According to the sliding control principle, when $\dot{s}_2 = 0$ we get the sliding control law for bridge crane, it looks like this:

$$u_{eq} = \frac{-\lambda_2 x_4 + \lambda_2 \dot{z} - f_2 + \ddot{x}_{3d} + \lambda_2 \dot{x}_{3d}}{g_2} \quad (12)$$

Control law (12) for bridge crane is a complex nonlinear function, for which $\dot{z}$ is not defined. Therefore, in this article, the author proposes adaptive fuzzy-sliding mode control law to imitate the sliding control law (12) based on fuzzy logic control and Lyapunov function.

3.2. Fuzzy controller based on sliding mode control principle

Define fuzzy controller with input s_2 and output u_f with fuzzy control rules *i-th as below*:

Rule i: IF s_2 is F_i THEN u_f is θ_i, i=1,2..n (13)

where: F_i is the input fuzzy set with the membership function $\mu_{Fi}(.)$ and θ_i are the changing singleton output values.

Defuzzification with centroid method, we obtain:

$$u_f(s_2, \theta) = \frac{\sum_{i=0}^{n} \xi_i.\theta_i}{\sum_{i=0}^{n} \xi_i} = \theta^T \xi \quad (14)$$

where: $\xi_i = \mu_{Fi}(s_2)$ is the reliability of the *i-th* fuzzy control rule.

$$\theta = [\theta_1, \theta_2, ..., \theta_n]^T;$$

$$\xi = [\xi_1 / \sum_{i=1}^{n} \xi_i, \xi_2 / \sum_{i=1}^{n} \xi_i, ..., \xi_n / \sum_{i=1}^{n} \xi_i]^T \quad (15)$$

According to the fuzzy approximation theory, there is an optimal fuzzy controller u_f^* of the form:

$$u_f^* = (\theta^*)^T \xi \quad (16)$$

where: θ^* is the optimal parameter set as follows:

$$\theta^* = \arg\min_{|\theta| \le M_\theta} \left\{ \sup_{|s_2| \le M_{s2}} |u_f - u_{eq}| \right\}, \forall t \quad (17)$$

where: M_θ, M_{s2} are the appropriate values when designing.

The control signal (16) is closest to (12), in other words, the deviation *d(t)* is the smallest:

$$d(t) = u_f^* - u_{eq} \text{ với } 0 \le d(t) \le D \quad (18)$$

Here the limit *D* is a positive number, but it is uncertain in practice. Called $\hat{D}$ the estimated value of the limit D. Then the estimate of the limit *D* is $\tilde{D}$ is defined as follows: $\tilde{D} = D - \hat{D}$

Therefore, based on (12), the author proposed a new control law for bridge cranes (5) and (6) with the following form:

$$u = u(s_2, \theta, \hat{D}) = u_f(s_2, \theta) + u_c(s_2, \hat{D}) \quad (19)$$

Here $u_f(s_2, \theta)$ the fuzzy controller (14) has the closest value to (12); and $u_c(s_2, \hat{D})$ is the compensated controller to compensate for the deviation (if any) between control laws (14) and (12). The compensation controller $u_c(s_2, \hat{D})$ is determined based on the Lyapunov stability theory, ensuring the stable asymptotic bridge crane control system.

3.3. Adaptive laws based on Lyapunov function

Next, the author summarizes the compensation controller $u_c(s_2,\hat{D})$ and determines the adaptive law to update the parameters θ, $\hat{D}$ so that (19) optimal operation is closest to (12), which means $\theta(t)\to\theta^*$ and $\hat{D}(t)\to D$. Set $\tilde{\theta}=\theta^*-\theta$ and select Lyapynov function with the following form:

$$V(s_2,\tilde{\theta},\tilde{D})=\frac{1}{2}s_2^2+\frac{1}{2\eta_1}\tilde{\theta}\tilde{\theta}+\frac{1}{2\eta_2}\tilde{D}^2 \quad (20)$$

According to (10) sliding surface derivative s_2 and from (19) combined with (12), (15), (18), we obtain:

$$\begin{aligned}\dot{s}_2&=\dot{x}_4+\lambda_2(\dot{x}_3-\dot{x}_{3d})-\ddot{x}_{3d}-\lambda_2\dot{x}_{3d}\\&=\lambda_2x_4-\lambda_2\dot{x}_{3d}+f_2+g_2u-\ddot{x}_{3d}-\lambda_2\dot{x}_{3d}\\&=\lambda_2x_4-\lambda_2\dot{x}_{3d}+f_2+g_2(u_c+u_f-u_f^*+u_f^*)-\ddot{x}_{3d}-\lambda_2\dot{x}_{3d}\\&=g_2(u_f-u_f^*+u_c+d)\end{aligned} \quad (21)$$

The derivative (20) over time, combined with (14), (18), (21) we have:

$$\begin{aligned}\dot{V}(s_2,\tilde{\theta},\tilde{D})&=s_2\dot{s}_2+\frac{1}{\eta_1}\tilde{\theta}\dot{\tilde{\theta}}+\frac{1}{\eta_2}\tilde{D}\dot{\tilde{D}}\\&=s_2g_2(u_f-u_f^*+u_c+d)+\frac{1}{\eta_1}\tilde{\theta}\dot{\tilde{\theta}}+\frac{1}{\eta_2}(D-\hat{D})(-\dot{\hat{D}})\\&=s_2g_2(-\tilde{\theta}\xi+u_c+d)+\frac{1}{\eta_1}\tilde{\theta}\dot{\tilde{\theta}}-\frac{1}{\eta_2}\dot{\hat{D}}(D-\hat{D})\\&=\frac{1}{\eta_1}\tilde{\theta}(\dot{\tilde{\theta}}-\eta_1g_2s_2\xi)+s_2g_2u_c+\frac{1}{\eta_2}\hat{D}\dot{\hat{D}}+s_2g_2d-\frac{1}{\eta_2}D\dot{\hat{D}}\end{aligned} \quad (22)$$

According to Lyapunov stability theory, the bridge crane control system is stable, the necessary condition is $\dot{V}\le 0$. From there according to (22), we determine the compensation controller $u_c(s_2,\hat{D})$ according to (23) and the parameter $\hat{D}$ update law according to (24), the parameter θ update law as follows (25).

$$u_c=-\hat{D}sign(s_2g_2) \quad (23)$$

$$\dot{\hat{D}}=-\dot{\tilde{D}}=\eta_2|s_2g_2| \quad (24)$$

$$\dot{\tilde{\theta}}=-\dot{\hat{\theta}}=\eta_1g_2s_2\xi \quad (25)$$

3.4. Prove asymptotic stability of bridge crane control system

Replacing (23,24,25) into (22) we obtain:

$$\dot{V}(s_2,\tilde{\theta},\tilde{D})=s_2g_2d-|s_2g_2|D\le 0 \quad (24)$$

According to (24), Lyapunov (20) function, $\dot{V}\le 0$, which ensures that $s_2,\hat{\theta},\hat{D}$ is limited. Thus, when time reaches infinity function *V* goes to 0:

$$\lim_{t\to\infty}V=0 \quad (25)$$

This also ensures that when $t\to\infty$ then $s_1\to 0, s_2\to 0$, $\theta\to\theta^*,\hat{D}\to D$, i.e. the crane control systems (5) and (6) are asymptotically stable, the output is tracked to the input reference.

Thus, the crane control system described by (5) and (6) uses the new control law proposed in the article, determined by (19), (14), (23), (24), (25) based on combination of sliding mode control, fuzzy logic and adaptive law based on Lyapunov function, ensuring precise control of the trolley position according to the set reference with eliminated payload oscillation, anti-disturbance.

The block diagram of adaptive fuzzy-sliding control system for 2D crane is shown as Figure 3, including: sliding surface block, fuzzy control block and compensation control block, parameters adaptive update block.

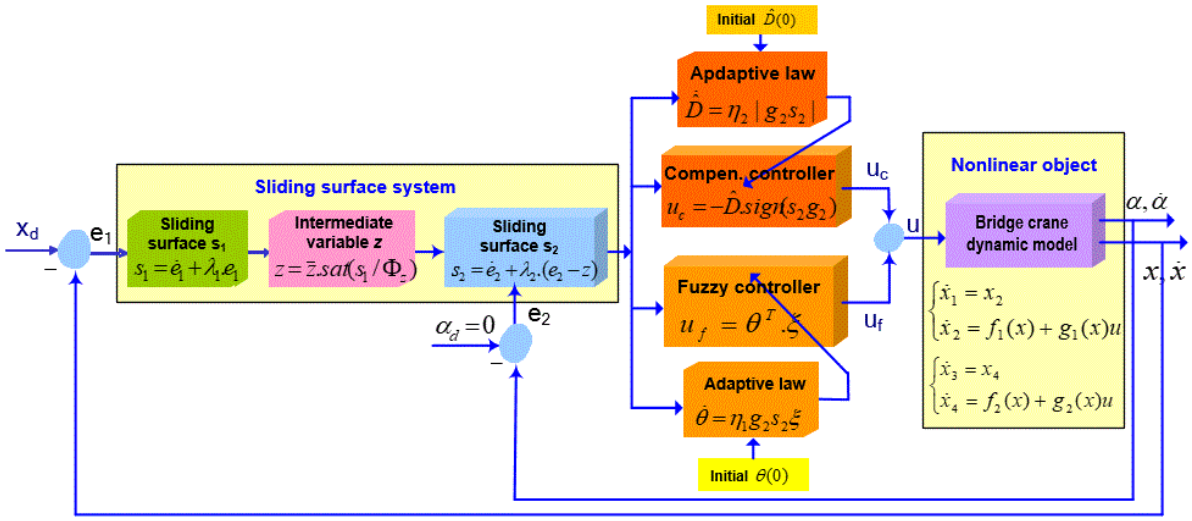


Figure 3. The block diagram of adaptive fuzzy-sliding control system for 2D crane

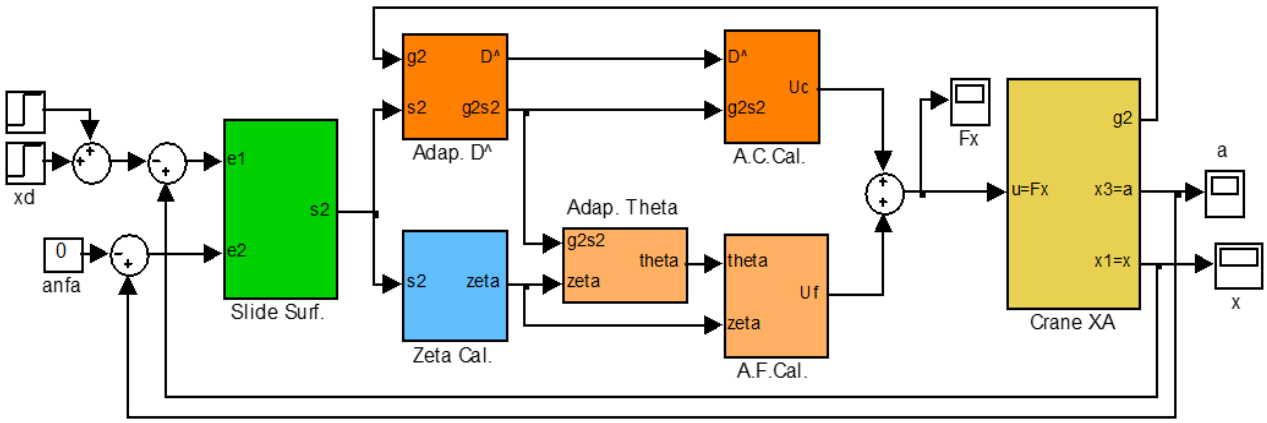


Figure 4. Simulation diagram of the bridge crane adaptive fuzzy-sliding mode control system

4. Simulation results

Simulation parameters of the crane are selected based on physical crane model at UTC laboratory [1].

According to the principle of sliding mode control [12] and crane parameters [1], we choose the sliding surface parameters as follows:

$$\lambda_1 = 0.6, \lambda_2 = 100, \eta_1 = 50, \eta_2 = 0.001, \bar{z} = 0.99, \Phi_z = 5.$$

Building a 2D crane control system on Matlab, using the adaptive fuzzy-sliding control law (AFS) proposed in the paper, we obtained the simulated system diagram as shown in Figure 4, including function blocks: Slide Surf block: calculate the sliding surface of the system; ZetaCal block: calculate ξ; Adap D^ block: calculate the update value $\hat{D}$ for the offset compensation controller; Adap. Theta block: calculate updated value θ for fuzzy controller; AF Cal. block: output of fuzzy controller; AC Cal. block: the output of the controller compensates for deviation; Crane XA block for bridge crane dynamic model.

Based on experience in crane operation and fuzzy control principle, with the above crane parameter [5], the author performs fuzzy input s_2 of the controller (14) by {NB, NM, NS, ZE, PS, PM, PB} having triangular integrated functions with domain $s_2 \in$[-20; +20] and fuzzy output u_f with {UN, UM, US, UZ, FS, FM, FB} having singleton functions with the physical value domain $u_f \in$[0;100].

Simulation scenario of the bridge crane adaptive fuzzy-sliding control system includes the following cases:

+ Case 1: Let the trolley run to position 0.5m and then the trolley runs to position 1.0m and changes coefficients λ_1, λ_2;

+ Case 2: Let the trolley run to position 0.5m and then the trolley runs to position 1.0m and changes coefficients η_1, η_2;

+ Case 3: Let the trolley run to position 0.5m and then the trolley runs to position 1.0m and changes coefficients $\bar{z}, \Phi_z$;

+ Case 4: The input reference is a slope;

+ Case 5: There is impact disturbance;

+ Case 6: Use AFS controller and PID controller [1].

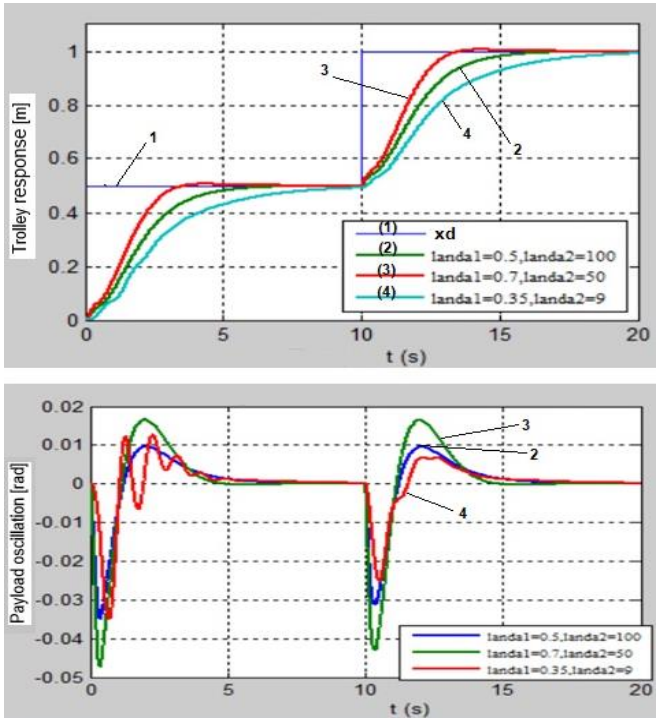


Figure 5. Crane response with different λ_1, λ_2

Table 1. The proposed controllers' performance

Index \ Controller	PID	AFS
Rise time	Small, ~2.5s	Small, ~2.5s
Over shoot	Large, 20.5%	Not
Steady time	Large, ~7s	Small, ~5s
Steady error	<5%, but large with disturbance	~0, or very small
Anti-swing payload	<10°, but large with disturbance	~0, or very small
Anti-disturbance	<10°, but large with loud disturbance	~0, or very small

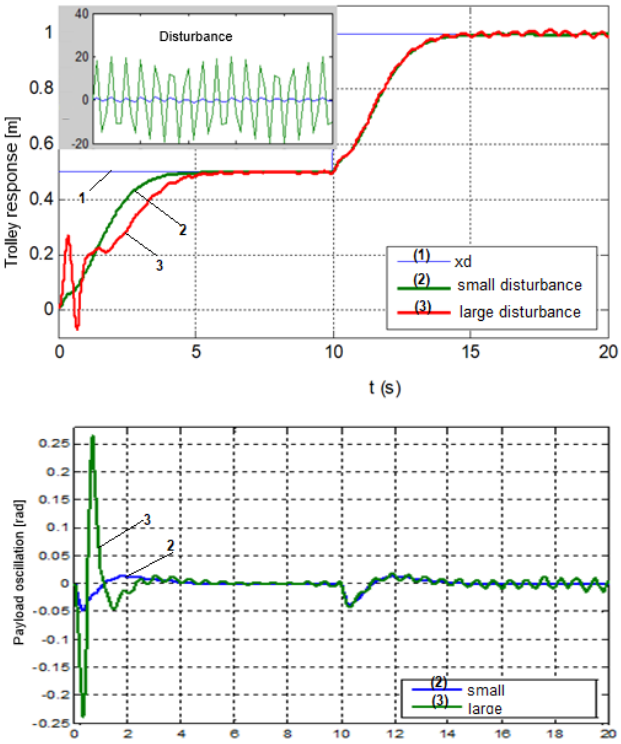


Figure 6. Crane response as affected by disturbance

Simulation results show that the quality of the bridge crane control system as using AFS controller is better than as using PID controller. The application of the AFS controller to the crane obtained the following results: quick trolley's position response, short settling time, eliminating steady error, anti-swing (or very small) payload, anti-disturbances (or very small).

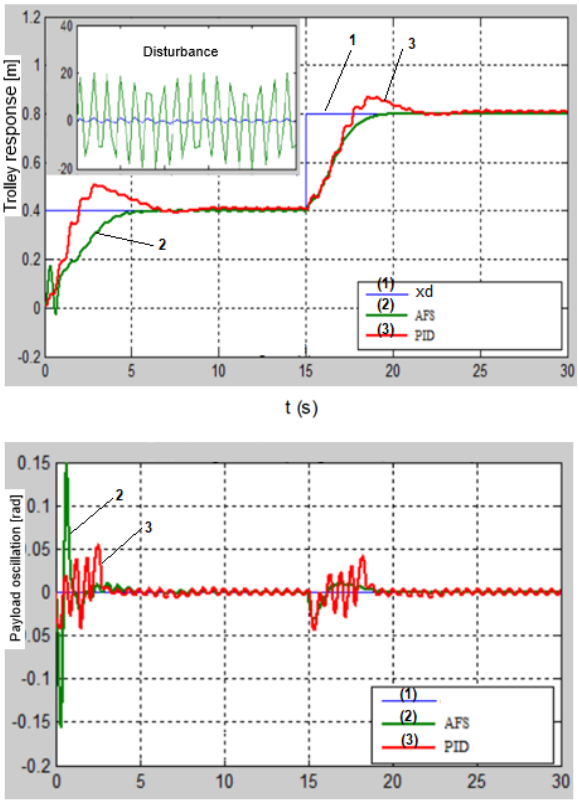


Figure 7. Crane response with AFS and PID controller

The simulation results also show that the selection of the coefficient of the sliding surface affects the quality of the adaptive control system of fuzzy-sliding crane, specifically as follows: when increasing λ_1, the trolley position response is better but greater is the payload oscillation; when decreasing λ_2 or increasing it too large, the payload swings strongly; when η_1 is too small or when η_2 is too large, the trolley position response has vibration and greater payload oscillation; when Φ_z decreases or $\overline{z}$ decreases too small, the trolley position response has vibration and the payload Will swing larger.

5. Conclusion

The article has proposed a new adaptive fuzzy-sliding mode controller (AFS) for 2D bridge crane based on combining sliding mode control with fuzzy logic and Lyapunov function. The AFS controller ensures the objective of controlling the position of the trolley tracking according to the reference trajectory, eliminating the payload oscillation and anti-disturbance (small amplitude). The quality of the bridge crane AFS control system is better than PID controller when crane is affected by large disturbance. The adaptive fuzzy-sliding mode control algorithm allows simple installation on hardware control devices with a normal amount of calculation. The AFS controller allows the crane to operate safely and reliably in harsh environments, such as the harbor.

The success of the AFS controller for 2D bridge crane will continuously study to be applied for the object in real time and for 3D crane with the random disturbances.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Trinh Luong Mien, Nguyen Van Tiem, *Design the state feedback controller for bridge crane*, The 5th national metrology conference, 20-21/05/2010, 720-725, 2010.

[2] Khalid L. Sorensen, William Singhose, *A controller enabling precise positioning and sway reduction in bridge and gantry cranes*, Control Engineering Practice 15, 2007.

[3] Jafari J., Ghazal M., Nazemizadeh M., *A LQR Optimal Method to Control the Position of an Overhead Crane*, International Journal of Robotics and Automation, 2014.

[4] Le Xuan Hai, Thai Huu Nguyen, et al., *Anti-sway tracking control of overhead crane system based on pid and fuzzy sliding mode control*, Journal of Science and Technology 55(1), 116-127, 2017.

[5] Amanpreet Kaur, Priyahansha, Shashiprabha Kumari, *Tanvi Singh, Position Control of Overhead Cranes Using fuzzy Controller*, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2014.

[6] Pal A.K., Mudi R.K., *An Adaptive PD-Type FLC and Its Real Time Implementation to Overhead Crane Control*, International Association of Scientific Innovation and Research, 2013.

[7] Shebel Asad, Maazouz Salahat, *Design of Fuzzy PD-Controlled Overhead Crane System with Anti-Swing Compensation*, S. ASAD ET AL Engineering, 2011.

[8] M. Nazemizadeh, *A PID Tuning Method for Tracking Control of an Underactuated Gantry Crane*, Universal Journal of Engineering Mechanics, 2013.

[9] Yonggang Li, Shuqing Zhou, Hongqiu Zhu, *A backstepping controller design for underactuated crane system*, The 30th Chinese Control and Decision Conference (2018 CCDC), 2018.

[10] Mahan Mahrueyan and Hamid Khaloozadeh, *Designing a Nonlinear Optimal Anti-Sway Controller for Container Crane Systems*, International Conference on Circuits, System and Simulation, 2011.
[11] Fang Y., Zergeroglu E., *Nonlinear coupling control laws for overhead crane system*, IEEE/ASME Transaction on Mechatronics, 2003.
[12] Shibly Ahmed AL-Samarraie, Bashar Fateh Midhat, *Sliding Mode Controller Design for a Crane Container System*, IJCCCE, 2014.
[13] Diantong Liu, Jianqiang Yi, Dongbin Zhao, Wei Wang, *Adaptive sliding mode fuzzy control for a two-dimensional overhead crane*, Mechatronics 15, 2005.
[14] Hahn Park, Dongkyoung Chwa, and Keum-Shik Hong, *A Feedback Linearization Control of Container Cranes: varying rope length*, International Journal of Control Automation and Systems, 2007.
[15] Tuan Anh Le, Soon-Geul Lee, Sang-Chan Moon, *Partial feedback linearization and sliding mode techniques for 2D crane control*, Transactions of the Institute of Measurement and Control 2014, Vol. 36(1) 78–87, 2014.
[16] Trinh Luong Mien, *Adaptive fuzzy sliding mode control for gantry crane as varying rope length*, International Journal of Engineering and Technology, Vol. 8 No. 4 Aug-Sep2016, 1784-1791, 2016.
[17] Chen Zhi-mei, Meng Wen-jun, Zhang Jing-gang, *Intelligent NN anti-swing control for bridge crane*, J. Cent. South Univ., 2012.

ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 4, No. 3, 171-182 (2019)
www.astesj.com
Special Issue on Advancement in Engineering and Computer Science

ASTES Journal
ISSN: 2415-6698

An Efficient Automotive Paint Defect Detection System

Sohail Akhtar[*,1], Adarsh Tandiya[2], Medhat Moussa[1], Cole Tarry[1]

[1]*The Robotic Institute, School of Engineering, University of Guelph, Guelph, ON, N1G 2W1, Canada*

[2]*Praemo Inc., Kitchener, ON, N2G 2Z3, Canada*

ARTICLEINFO

Article history:
Received: 25 March, 2019
Accepted: 23 May, 2019
Online: 12 June, 2019

Keywords:
Defect detection
Deflectometry
Profilometry
Automatic inspection
Painted surface inspection

ABSTRACT

Vision-based defect detection techniques are widely used for quality control purposes. In this work, an efficient deflectometry based detection system is developed for semi-specular/painted surface defect detection. This system consists of a robotic arm that carries a screen/camera setup and can detect defects on large surfaces with different topologies, such as a car bumper, by traversing its profile. A hybrid pipeline is designed that utilizes multi-threading for optimal resource utilization and process speed. Specific filters are also designed to remove spurious defects introduced by acute curvature changes and part edges. The system was successful in consistently detecting various defects on small test samples as well as on large bumper parts with varying topology and color and can accommodate inherent ambient lighting and vibration issues.

1 Introduction

Quality control is a crucial factor in manufacturing industry as it affects customer satisfaction, reduces production cost and increases profitability. In the automotive industry, vehicles are usually assembled from parts shipped by various original equipment manufacturers (OEM) to the assembly plant. Example of these parts includes front and back bumper covers, side fenders, and other exterior parts which are normally manufactured and painted to specific colour before being shipped to another plant for assembly. These outer body parts are made by Thermoplastic PolyOlefin (TPO) injection moulding process. This process consists of three main steps: moulding, cleaning, and painting. Defects may induce during any of these processes, which results in part rejection or rework, causing loss of revenue. As such, it is essential to perform a full inspection of every part. This inspection process is usually carried out by human inspectors. It is a costly and a labour intensive job which requires multiple inspection lines for high volume yield. Further, the defect judgement is very subjective, which results in inconsistencies. As a result, overall productivity and quality are diminished. This paper presents a system for inspecting automotive painted semi-specular exterior body parts and is an extension of work originally presented in 15th Conference on Computer and Robot Vision (CRV 2018) [1]. This extended version includes expended testing and analysis.

There are a series of challenges that impact the development of an inspection system for this task, including:

1. Parts are in motion while they are inspected on the production line. The vibrations induced from this motion makes profiling the surface harder.
2. The inspection process must fit within the existing production cycle time.
3. The part being inspected vary in curvature, size, shape and material, with different specular characteristics.
4. The visibility characteristic of the defect depends on the external lighting conditions. Designing a proper lighting environment for such a system is a challenging task due to high reflection coefficient of the test surface [2].
5. Each type of defect varies in shape and size, and experienced inspectors even miss some defects.

In recent years, vision-based surface inspection systems have found a burst of application in areas such as defect detection in aluminium sheets [2], locomotive

[*]Corresponding Author: Sohail Akhtar, University of Guelph, 50 Stone Rd. E, Guelph, ON, soakhtar@uoguelph.ca

https://dx.doi.org/10.25046/aj040323

rail tracks [3], liquid crystal displays (LCDs) [4], plastic objects [5], sphere parts [6], agricultural food produce [7], fabric and textile industry [8] etc. However, defect detection and surface profiling of a specular surface remains a challenging problem due to the reflective nature of the surface. Profilometry is a widely used approach for characterization, reconstruction and inspection of such surfaces. Profilometry approaches can be generally divided into contact and non-contact methods. Contact-based profilometry, such as stylus profilometer, scanning tunnelling microscopy, scanning force microscopy, coordinate measuring machine (CMM), etc., is a high-resolution method. The measurement accuracy is in the order of nanometers and is considered as a gold standard for surface finish measurement. This approach is independent of surface characteristics or shape and also works well in dirty environments. As a result, it can be used for both specular and non-specular surface measurement. However, being a direct contact method, it may damage high finish inspection surfaces such as painted automotive parts and its operating speed is also sluggish.

Non-contact profilometry methods are used for high-speed surface scanning and 3D reconstruction [9]. Over the years, many different non-contact profilometry techniques are developed, such as common structured light projection, phase shifting interferometry, deflectometry, digital holographic microscopy, etc. These approaches are based on the inhomogeneous reflection of light from defects. The visibility of defects can be enhanced by the use of a structured light source. In common structured light projection method, the specular surface is coated with a thin layer of powder to make it behave like a diffused surface [10] and then fringe pattern profilometry is used for analysis. This additional coating changes the surface geometry so cannot be reliably used for surface defect detection. On the other hand, interferometry is highly accurate but requires a reference and is not suitable for complicated free-form surfaces. Further, the field of view is limited, so it cannot be used for large surfaces. Hence, deflectometry is recently used for measuring a large object with varied surface topology. A sinusoidal fringe pattern is widely used as a structured light source in deflectometry analysis [11]. In smooth defect-free regions, imaging process obeys specular reflection while the incident rays are dispersed in defected areas. As a result, an acute gradient variation is observed in defect bearing regions, which is used for defect segmentation. These approaches have a lower measurement accuracy (in micrometres range) compared to contact profilometry approaches. Since these approaches are surface dependent, so are affected by the testing environment as any dirt or external dust particles will affect the obtained results.

A new defect detection system is presented in this paper, which is based on the deflectometry principle. The developed system can overcome the practical challenges of a production line. Controlled lighting and camera configurations are designed to counterpoise vibration effects. Various filters are used to segment the region of interest and eliminate false defects due to edge effect and abrupt curvature changes. Finally, to track parts while going through the production line, the system is designed to be mounted at the end of a robotic arm. The proposed system is validated by field testing in an operational automotive painting booth.

The remainder of this paper is organized as follows: Section 2 provides a brief review of the deflectometry principle as it is the foundation of the defect detection system. Section 3 describes the proposed system and explains the working of its different components. Section 4 depicts the system setup and presents implementation details. Section 5 archives the obtained experimental results with analysis. Finally, Section 6 summarizes the findings of this research and highlights some areas of future development.

2 Deflectometry Principle

In this section, we will provide a brief overview of the deflectometry principle, which forms the basis of the proposed system's working principle. The use of deflectometry is actively explored by researchers for the measurement of objects with an abrupt slope change or large size. Many different deflectometry methods are proposed for surface profilometry and defect detection. Some of these include Moier deflectometry [12–16], Ronchi method [17, 18], laser scanning deflectometry [19–21], and phase measuring deflectometry (PMD) [11, 22–24].

PMD technique is used in this paper as it is highly accurate and provides continuous data for surface profiling. It has a large dynamic range and can give full-field measurements. This technique is first developed by Horneber et al. [25]. Over the years, researchers have modified PMD to measure different specular objects [11, 26, 27]. PMD is also successfully applied for specular and semi-specular surface defect detection [28–33]. However, various challenges arise while implementing deflectometry to inspect a larger part like a car bumper which are successfully addressed in the developed system.

2.1 Phase-shifting Deflectometry

In deflectometry, the topographical information of a specular surface is obtained by analysing the reflections of a structured light source. A sinusoidal fringe pattern, displayed on an LCD screen located at some distance from the specular test surface, acts as the light source. The attached camera setup captures the reflections of the deformed pattern from the test surface. By applying phase-shifting and phase-unwrapping algorithms, useful information is extracted from these deformed fringe patterns which are used for surface profile construction. A phase varying single frequency pattern is generally used for deflectometry analysis. A pixelated version of such a multi-phase single frequency pattern is defined as [34]:

$$f(x,y) = \frac{G}{2}\left[1 + \sin(2\pi f \frac{x}{p} + n\frac{2\pi}{7})\right] n = 1..7 \quad (1)$$

where, $f(x,y)$ is the displayed pattern intensity at (x,y) pixel location, G is the maximum of the image intensity range, f is the frequency of the displayed pattern, p is the total pixels in a time period and n is the number of phase-shifted patterns selected.

These phase-shifted patterns are one-by-one displayed on the LCD screen, and their reflections are captured from the specular surface. The intensity of the captured pattern is given by (2) [35].

$$I(x,y) = A(x,y) + B(x,y)cos(\phi(x,y)) \quad (2)$$

where $I(x,y)$ is the captured image intensity, $A(x,y)$ is the ambient light intensity, $B(x,y)$ is the amplitude of the modulated fringe pattern and $\phi(x,y)$ is the desired phase measurement.

The measured phase is directly related to the surface topology. Computing this phase at each pixel location will create a surface phase map. Captured surface reflections of the phase-shifted patterns are used to calculate this phase map by applying Windowed Discrete-Fourier Transform (WDFT). For seven phase-shifted single frequency patterns, the closed-form solution is represented by the following relation [36]:

$$\phi = \tan^{-1}\left[\frac{(I_1 - I_7) - 3(I_3 - I_5)}{4(I_4) - 2(I_2 + I_6)}\right] \quad (3)$$

where $I_{1,\ldots,7}$ correspond to the captured reflections of the seven phase-shifted patterns displayed on the LCD screen. The phase map ($\phi(x,y)$) is in $0 - 2\pi$ range, so a phase unwrapping step is needed to recover the actual phase with the inclusion of appropriate multiples of 2π [37]. Then the absolute derivative of the phase map is computed with respect to pattern variation axis (x-axis) as defined by the following equation [33]:

$$D_y(x) = \left|\frac{\mathrm{d}\left|\phi(x)\right|}{\mathrm{d}x}\right| \quad (4)$$

here, $D_y(x)$ is a continuous function which represents line topography for the corresponding y^{th}-axis.

The effect of this differential operation is the attenuation of low-frequency signal information and accentuation of high-frequency contents induced due to the presence of defects in the specular/semi-specular surface. As a result, $D_y(x)$ remains almost constant in defect-free regions, whereas an acute variation is observed in defect-bearing areas.

3 Defect Detection System

Use of deflectometry for big automobile bumper inspection brings its own set of challenges. Its impractical to examine the entire part surface in one scan. Therefore, it is divided into smaller segments based on the reflection of the displayed pattern. Since external light sources impact the detection results, controlled lighting setup is constructed to minimize the effect of these external factors. Figure 1 shows the image processing pipeline used. Captured surface reflections are pre-processed before deflectometric defect detection operation. Several false positives are observed due to the topographical variation in the test surface. A set of spurious defect removal filters are designed to eliminate these false positives. Camera shutter speed, aperture and robotic motion are adjusted to reduce vibration effects and eradicate motion blur. Since the detection operation is carried out in small segments, the results are then combined and localized on test parts. Finally, defect characteristics are gathered and archived in the designed database which can later be used for display or higher level decision making for process improvement. The following sections provide additional details on the functionality of different components.

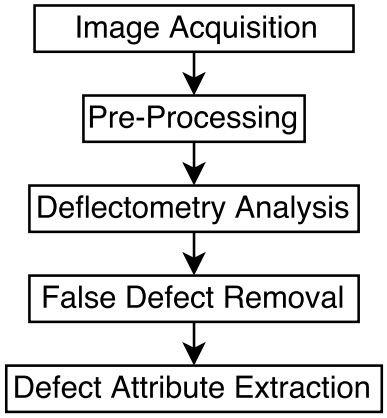


Figure 1: Defect detection scheme

3.1 Capturing and Pre-processing

Figure 2 and Figure 3 show the process involved in capturing and pre-processing of images. OpenGL library is used for the rendering of seven phase-shifted sinusoidal patterns. The patterns are generated only once during the run-time using pattern screen resolution and then stored in memory as 2D matrices. Later on, these stored patterns are displayed one-by-one for deflectometry analysis. The reflection of these patterns is captured from the part surface by an attached camera. Pattern display and capture of surface reflections are carried out sequentially. Once a set of seven surface reflections is captured, a pre-processing thread is initiated to compute the phase map and find its derivative. It is noted that a single phase-map-derivative image is created for each segment. While the pre-processing step is running, the next set of seven patterns are ready to be displayed on the screen and captured by the camera. This multi-threading approach reduces cycle time through parallel computation.

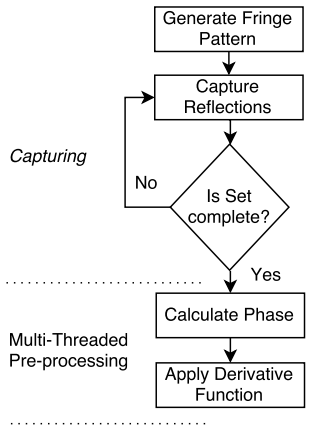


Figure 2: Sequential image capture and mutli-threaded pre-processing steps

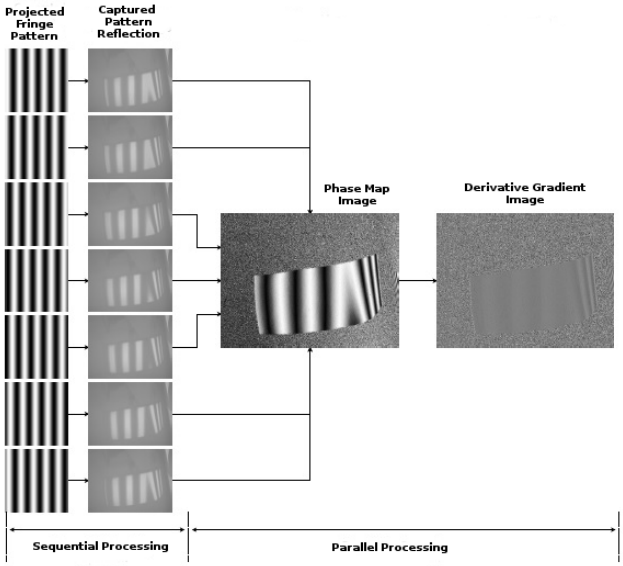


Figure 3: Surface reflection capturing and phase-map-derivative image generation

Once the Capturing and Pre-processing steps are completed, a defect detection thread is initiated, which extracts defect information from the phase-map-derivative image. Defect detection process can be divided into three main stages; Fringe region segmentation, Defect pool and spurious defect elimination and Defect attributes extraction and registration.

3.2 Fringe Region Segmentation

A fixed focus camera at a defined distance from the test surface is used in this study. As a result, the camera captures a fixed size image that may contain both the fringe projected region and non-fringe region. Further, the size of the reflected pattern varies with the surface curvature; the reflected pattern shrinks as the surface become convex. As a result, the first step in the defect detection process is to crop the fringe pattern projected region from the phase-map-derivative image.

The obtained phase-map-derivative image in the previous step is thresholded and smoothed with a Gaussian kernel. The resulted blobs are eroded to remove uneven corners and then selected based on size to create a mask. This mask is used to extract the fringe projected region as shown in Figure 4, which is later used for defect detection.

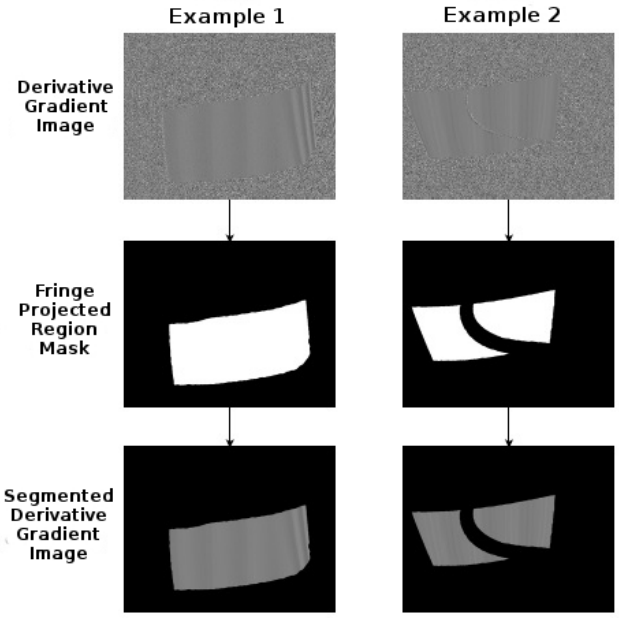


Figure 4: Fringe region extraction from the generated phase-map-derivative image

3.3 Defect Pool and Spurious Defect Elimination

Once the fringe projected region is extracted, thresholding and edge detection are used to create a defect pool. This defect pool consists of possible true and false defects. False defects are a combination of tiny defects below the accepted size criterion, edge defects and defects due to abrupt curvature changes. Three different filters; Size/Noise filter, Surface curvature filter and Edge filter are designed to remove these false defects. These filters work in parallel to remove these defects as depicted in Figure 5.

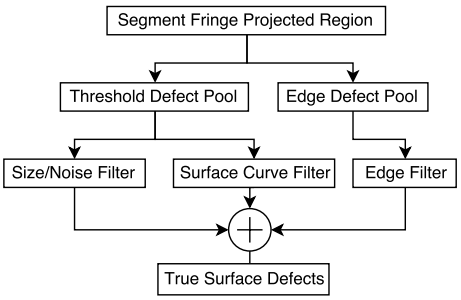


Figure 5: Processing and false defect removal

A size based filter is designed to remove tiny noise like defects. It is observed that after thresholding, several very small defects are detected in the binary image. This filter removes all those defects which are below the given size criterion. The second filter removes the false defects from the curved region. This filter uses parallel lines to determine the curved region in

the captured surface reflection. A Canny edge detection operation followed by Hough transform is used for these lines detection. The angle variation of the detected lines is exploited to determine the curved region from where the defects are removed. The last filter removes spurious defects in the edge region, which are the result of the application of WDFT. It makes use of both curved region information and defect pool generated by Canny edge detection operation to segregate defects produced due to edge effect.

The application of these filters results in three separate images that contain only legitimate defects. Later on, all these defects are aggregated together to form an image which shows all the valid defects in the inspected region as depicted in Figure 5. Application of these filters on a typical gradient image is shown in Figure 6.

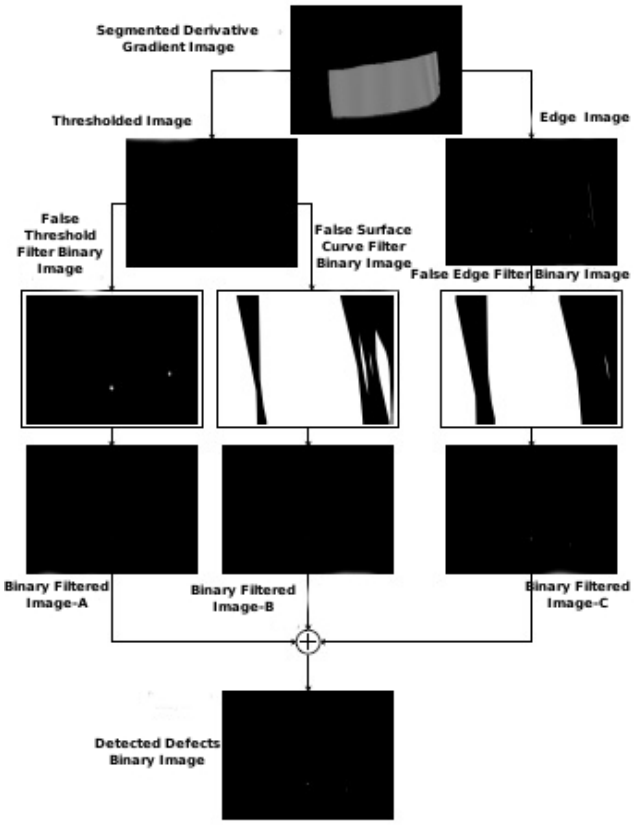


Figure 6: Application of false defect filters on defect pool

3.4 Defect Attributes Extraction and Registration

Once all the true defects are found, defect attributes such as size (in terms of pixels), bounding region and centroid location are determined. To extract these characteristics, close by defects are first fused by a morphological dilation operation. Later on, a contour-finding operation is performed to separate different defects and then above mentioned attributes are computed for each found contour.

These attributes are also recorded in the appended database, which can be used in future for tracking, decision making or other process adjustments. The results are also overlaid on a part image for physical defect localization which is saved in a directory for immediate use. All these different steps are delineated in Figure 7.

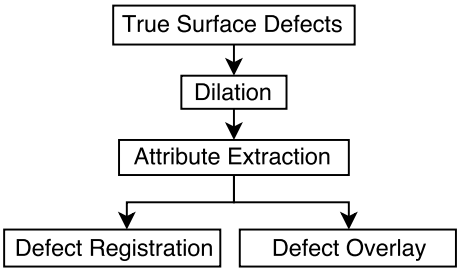


Figure 7: Defect attributes extraction and display

4 Implementation Details and Experimental Setup

4.1 System Architecture

Figure 8 shows the entire system implementation. Since the system is expected to inspect large parts, a robotic arm is used where the pattern screen and camera are mounted on the end-effector. The system includes several additional components to enable robot movement for full part inspection such as positioning of the robot end effector, generation of inspection patterns and administrative operations such as results registration and display.

The defect detection process consists of both sequential and parallel operations. As a result, a hybrid computational framework is designed. It is evident in Figure 8 that individual steps need to be completed to proceed to the next component for processing, so parallel operations (multi-threading) are only used for computationally expensive processes. First of all, robot positions the LCD pattern screen tangent to the surface to be inspected. Then seven phase-shifted sinusoidal fringe patterns are sequentially displayed on the screen, and the attached camera captures corresponding surface reflections. These surface reflections are pre-processed, and defect detection operation is applied to find the surface abnormalities. These two steps, pre-processing and defect detection, are computationally expensive, and multi-threading is used here. This multi-threading operation can be activated/deactivated by the frontal administrative control panel. All detected defect information is stored in the designed database system that can be accessed remotely. All these steps are controlled by a front end administrative control panel, which also displays a sample test part image on which detected results are overlaid. It also provides the option to configure features such as directory paths for registering captured and processed images, the number of segments to divide the test surface and displays the progress of the detection operation.

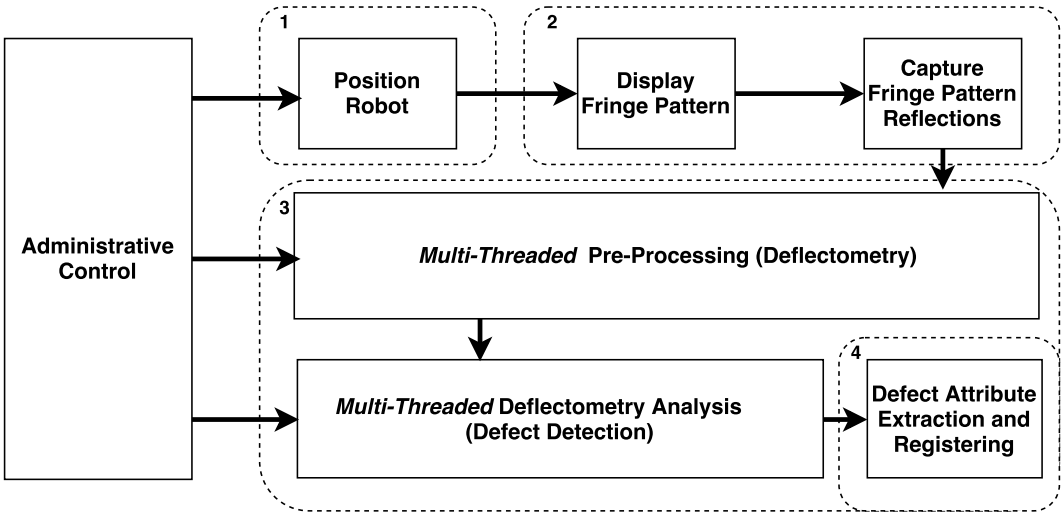


Figure 8: Schematic of implemented defect detection system

4.2 Experimental Setup

4.2.1 Overall Setup

Testing was conducted at the Robotics Institute @ Guelph lab. The experimental setup is shown in Figure 9. The test parts were placed on a carrier similar to the ones used at the part manufacturing facility. The carrier was stationary during the testing. A KUKA KR 16-2 KS industrial robotic arm is used to carry the screen/camera setup to transverse the large bumpers under test. The robotic arm was programmed to move in a predefined path to inspect each segment and hence the whole part. This operation controls the movement of the screen/camera setup. A controlled lighting environment was enforced where the pattern screen is the only source of illumination. Any direct light source incident on the part whose reflection reaches the camera will create a blind spot in the gradient image. The test surface should also be free of any dust particles as they will produce a dispersed reflection of light and will appear as a defect. A 40 inch commercial TV screen is used as the pattern screen. The computing system consists of an Intel Core i7-6700 3.40 GHz processor and has 16 GB RAM and Nvidia GTX960 graphics card running on 64-bit Windows 10 OS. The detection code is written in C++ using Qt Creator 5.6, OpenCV 3.2 and OpenGL libraries. The camera is connected via high-speed USB 3.0 connection and pattern screen is interfaced with an HDMI cable.

4.2.2 Camera Setup

A monochrome Point Grey 3.2 MP (GS3-U3-32S4M-C) camera with a 12.5 mm Fujinon lens HF12.5HSA-1 (LENS-50FS-125C) is used. The camera is attached to a specially designed fixture on the side of the pattern screen so that it looks in the middle of the projected pattern on the test sample, as seen in Figure 10. A fixed focus camera, as already mentioned, is sharply focused on the test surface, which is approximately 40 cm away from the camera. Once the robot starts traversing the surface topology, a distance of 35 cm to 45 cm is maintained so that the image remains within the camera's depth of filed and hence sharply focused.

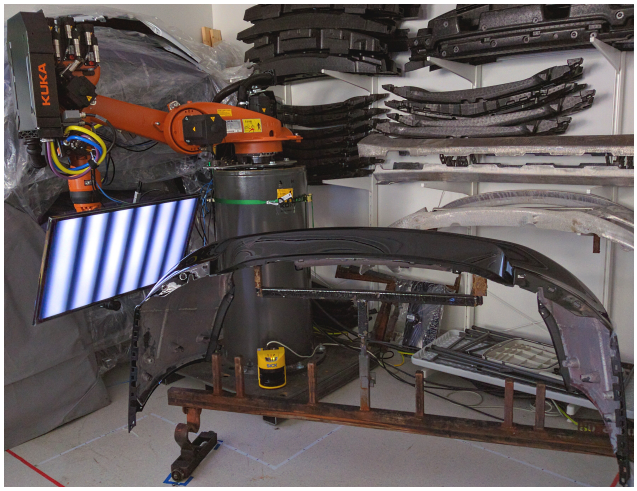


Figure 9: Experimental setup

Figure 10: Camera view

4.3 Surface Segmentation

It is virtually impossible to inspect the full test part in one capture due to its large size. There are two reasons for this. Firstly, there is a lot of variation in the surface topology of the test surface, and the pattern screen should be tangent to it to capture the displayed pattern. Secondly, only a limited payload both in terms of size and weight can be attached to the robot end effector. Hence, a huge screen which can project the pattern on the full part cannot be used. As a result, the test surface is partitioned in small regions based on the reflected pattern as shown in Figure 11.

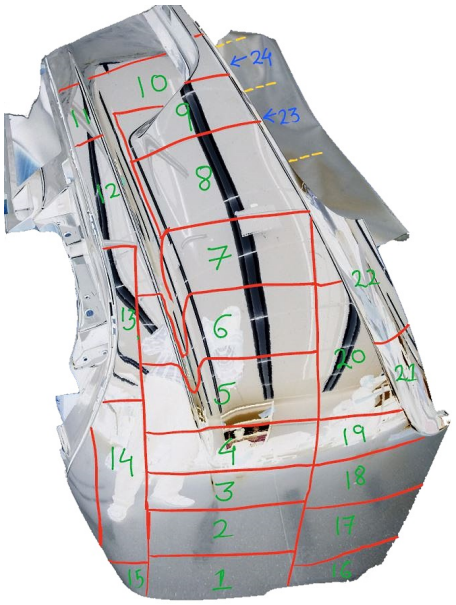


Figure 11: Bumper segmentation for inspection

Seven phase-shifted sinusoidal patterns are projected on each region, and the attached camera captures corresponding reflections. There is some overlap in the captured surface reflections among adjacent segments that accommodate for some variation in part placement. These different regions are then appended together to localize the identified defects on the entire test surface.

4.4 Test Specimen

Ten test samples are used in this study. These samples are provided by an OEM exterior body parts manufacturer. These are large SUV/Minivan bumpers of 5.5 ft long, 1.5 ft high and 2.5 ft deep, as shown in Figure 11. These bumpers are of different colours and are rejected due to the presence of various surface defects. These defects are marked on the test parts by expert human inspectors working at the manufacturing facility. In addition to these defects, some more defects are artificially created to augment the defect database used as ground truth (GT). Since the bumpers are symmetric, we only needed to inspect one half of each bumper. In real production, two robots can be used to examine the whole bumper. Each tested part is divided into 24 small segments as depicted in Figure 11.

5 Results and Analysis

Two sets of experiments are performed to evaluate the performance of the developed painted surface defect detection system. In the first study, small defect samples are used to verify the effectiveness of the deflectometry principle in detecting various defects. Once, the efficacy of the deflectometry principle is validated on these small samples; large parts are tested. Large parts have their own set of challenges in terms of size, sudden curvature change, throughput etc.

5.1 Small Test Samples Preliminary System Validation

Different type of surface defects such as dirt, cold sludge, fisheye, blister etc. occur during moulding or painting process. Small test samples bearing these defects are used to validate the effectiveness of the detection principle. Some of the sample parts are shown in Figure 12. These small parts are mounted on a tripod in front of pattern screen, as shown in Figure 13. The parts are placed 40 cm away from the camera, which captures the surface reflections. A set of seven sinusoidal fringe patterns with a phase shift of $2\pi/7$ are generated to display on the LCD screen and then their reflections are captured. These surface reflections are analysed to reconstruct the surface profile. During this testing, the fringe projected area is manually segmented for deflectometry analysis. This process is depicted in Figure 14 and the results are tabulated in Table 1. Here, detection accuracy represents whether a defect is successfully detected by the developed system.

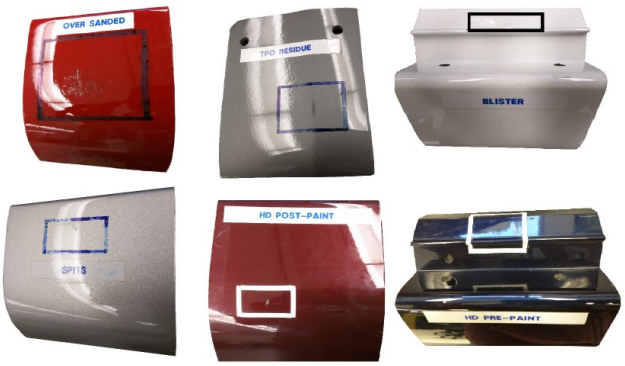


Figure 12: Some typical surface defect samples

Figure 13: Camera with fringe pattern screen and test part on tripod

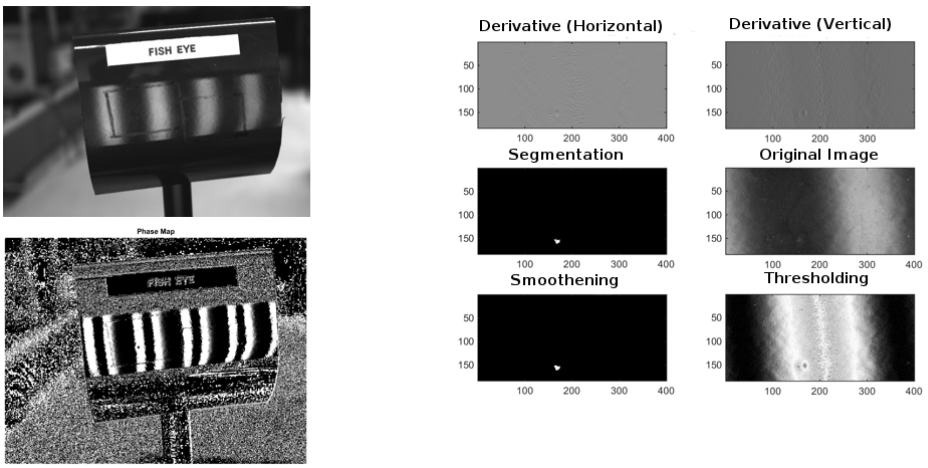


Figure 14: Fisheye defect detection

Table 1: Small sample defect detection results

Sr. No.	Defect Type	No. of Samples	Detection Accuracy
1.	Dirt	2	100%
2.	Cold Slug	4	75%
3.	Spits	1	100%
4.	Fish Eye	2	100%
5.	Flash	2	100%
6.	Blister	3	100%
7.	TPO Protusion	2	100%
8.	Over Sanded	3	100%
9.	Hot Pre-paint	1	100%
10.	Overflame Torch	2	100%
11.	Water Spots	1	100%
12.	TPO Residue	3	100%
13.	Popping	1	50%
14.	HD Post Paint	3	100%

These results indicate that the developed system mostly detects the studied defects although some low detection performance is observed for cold slug and popping defects. Popping defects are salt-and-pepper noise like small defects that occur when the solvent or air trapped in paint film escapes during the drying process. This creates a lot of tiny defects clustered in a small area. A 50% popping defect detection accuracy indicates that approximately half of these tiny defects are detected, but it is enough to declare the part defective.

5.2 Large Parts Defect Detection

Following the experimental setup outlined in section 4, various large bumpers are tested by the developed system. The developed framework was successful in detecting most of the manufacturing as well as the artificially created defects. Some of these detection results are shown in Figure 15.

Results are reported in two tables. The segment-wise results are given in Table 2 while body part color based results are reported in Table 3. These results are reported in terms of recall, precision and F-measure. Equal weighting of recall and precision is considered for computing F-measure. These metrics are defined as follows:

$$Recall = \frac{TP}{(TP+FN)} \tag{5}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{6}$$

$$F-measure = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \tag{7}$$

where TP is true positive, FP is false positive and FN is false negative. Ground truth data to compute these measures is provided by the part manufacturer.

By reviewing the results, we can deduce that the performance of the defect detection system varies depending on part colour and segment location. The performance of the algorithm is lower in segments 8, 9 and 10 where precision drops significantly to 0.46, 0.43 and 0.30, respectively. These segments contain highly curved sections which drastically distorts the projected pattern resulting in a higher number of FP. This can be improved by tuning the system differently on a per region basis to accommodate the different variations in surface curvature. The results by part colour show that the performance of the detection system is compromised on white parts. White surfaces reflect all light incident to it, so the algorithm becomes more sensitive to surrounding light. Besides, the contrast of the projected pattern is muted by white parts affecting the visibility of defects on the part surface. The combination of these factors reduces the accuracy of the algorithm on white parts.

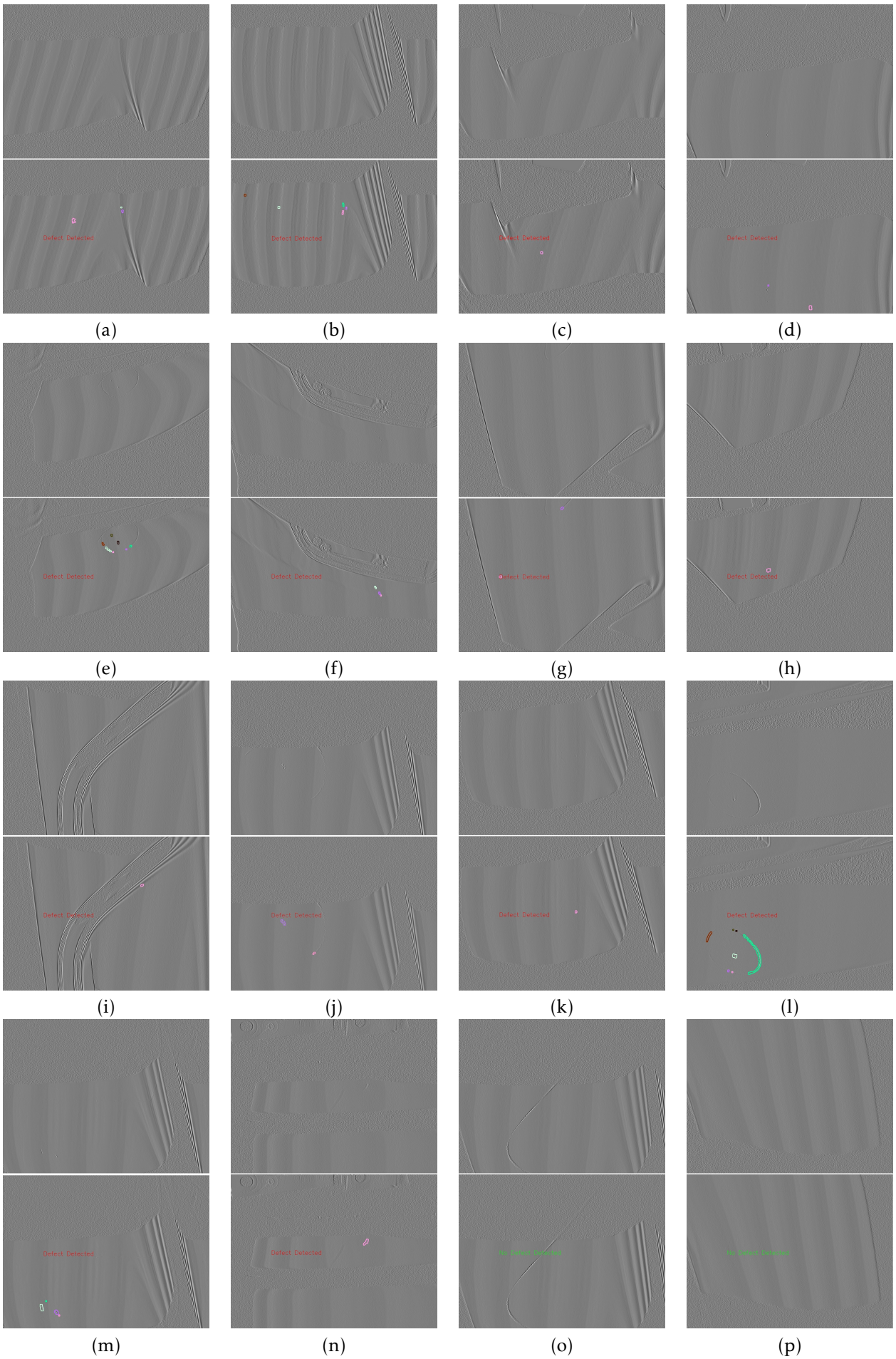


Figure 15: Some defect detection results on large bumpers First row: Gradient image, Second row: Detected defect ((o) is an example where defect is not detected as it is not enclosed by fringe region and (p) is an example of a defect free region)

Table 2: Defect detection results by segment

Seg.	Samples	*TP*	*FP*	*FN*	Recall	Precision	F-measure
1	10	6	4	0	1.00	0.60	0.75
2	10	0	0	0	NA	NA	NA
3	10	6	2	1	0.86	0.75	0.80
4	10	3	1	0	1.00	0.75	0.86
5	10	3	1	0	1.00	0.75	0.86
6	10	14	0	1	0.93	1.00	0.96
7	10	8	0	1	0.89	1.00	0.94
8	10	6	7	0	1.00	0.46	0.63
9	10	3	4	0	1.00	0.43	0.60
10	10	6	14	2	0.75	0.30	0.43
11	10	0	3	1	0.00	0.00	NA
12	10	2	0	0	1.00	1.00	1.00
13	10	2	0	0	1.00	1.00	1.00
14	10	4	0	0	1.00	1.00	1.00
15	10	3	0	0	1.00	1.00	1.00
16	10	3	0	0	1.00	1.00	1.00
17	10	0	0	2	0.00	NA	NA
18	10	1	0	0	1.00	1.00	1.00
19	10	2	1	0	1.00	0.67	0.80
20	10	3	2	1	0.75	0.60	0.67
21	10	0	0	0	NA	NA	NA
22	10	7	1	0	1.00	0.88	0.94
23	10	7	0	0	1.00	1.00	1.00
24	10	0	0	0	NA	NA	NA
Average					**0.87**	**0.76**	**0.81**

Table 3: Defect detection results by color

Part	Color	TP	FP	FN	Recall	Precision	F-measure
1	Black	14	0	0	1.00	1.00	1.00
2	Dark blue	3	1	1	0.75	0.75	0.75
3	Dark Blue	6	1	0	1.00	0.86	0.92
4	Dark silver	9	2	2	0.82	0.82	0.82
5	Dark blue	0	1	1	0.00	0.00	NA
6	Purple	16	1	2	0.89	0.94	0.91
7	Dark silver	9	2	1	0.90	0.82	0.86
8	White	5	19	2	0.71	0.21	0.32
9	White	12	9	0	1.00	0.57	0.73
10	Black	15	4	0	1.00	0.80	0.89

6 Conclusion

A deflectometry based defect detection system is presented in this research paper. The developed system has successfully detected large and subtle defects on large painted automotive parts. The system analyses gradient variation in the measured phase for defect detection and does not require an existing dataset or pre-training for its functioning. The use of robotic arm lends it the capability to inspect varying surface geometries. The system examines the whole part by operating on individual segments. These defects are then aggregated to localize them on the full part.

The system successfully detected various manufacturing defects induced during moulding and painting processes on small test samples as well as on large automotive bumpers. For large bumper testing, additional algorithms are designed and used to extract the region of interest, as well as eliminate spurious defects detected due to edge effects and abrupt curvature changes. The algorithm processing takes approximately 0.35 seconds per segment, and this doesn't include the time that it takes the robotic arm to move between segments.

It is noted that the implemented framework can only detect defects if they are enclosed by a fringe projected region from all sides. It is possible to use multiple cameras to reduce the number of segments and speed-up the inspection process. A colour fringe pattern could potentially be used to improve the system's performance on light coloured parts. Use of a phase shifted coloured pattern can be explored as it will alleviate the requirement for a stationary part-camera/screen system by eliminating the need for multiple phase-shifted patterns.

Acknowledgement This work is supported by the Natural Science and Engineering Research Council (NSERC), Canada.

References

[1] A. Tandiya, S. Akthar, M. Moussa, and C. Tarray, "Automotive semi-specular surface defect detection system," in 2018 15th Conference on Computer and Robot Vision (CRV), May 2018, 285-291. DOI: 10.1109/CRV.2018.00047

[2] C. Fernandez, C. Platero, P. Campoy, and R. Aracil, "Vision system for on-line surface inspection in aluminum casting process," in Proceedings of IECON '93 - 19th Annual Conf. of IEEE Industrial Electronics, 15-19 Nov. 1993, 1854 - 1859. DOI: 10.1109/IECON.1993.339356

[3] Z. Liu, W. Wang, X. Zhang, and W. Jia, "Inspection of rail surface defects based on image processing," CAR 2010 - 2010 2nd Int. Asia Conf. Informatics Control. Autom. Robot., **vol. 1**, pp. 472-475, 2010. DOI: 10.1109/CAR.2010.5456793

[4] Y.-G. Cen, R.-Z. Zhao, L.-H. Cen, L. Cui, Z. Miao, and Z. Wei, "Defect inspection for TFT-LCD images based on the low-rank matrix reconstruction," Neurocomputing, **vol. 149**, 1206-1215, 2015. DOI:10.1016/j.neucom.2014.09.007

[5] B. Liu, S. Wu, and S. Zou, "Automatic detection technology of surface defects on plastic products based on machine vision," 2010 International Conf. on Mechanic Automation and Control Engineering, MACE2010, 2213 - 2216, 07 2010. DOI: 10.1109/MACE.2010.5536470

[6] J. Guo and W. Shao, "Automated detection of surface defects on sphere parts using laser and CDD measurements," IECON 2011 - 37th Annu. Conf. IEEE Ind. Electron. Soc., pp. 2666-2671, 2011. DOI: 10.1109/IECON.2011.6119732

[7] S. Janardhana, J. Jaya, K. J. Sabareesaan, and J. George, "Computer aided inspection system for food products using machine vision A review," 2013 Int. Conf. Curr. Trends Eng. Technol., pp. 29 - 33, 2013. DOI: 10.1109/ICCTET.2013.6675906

[8] P. M. Cho CS, Chung BM, "Development of real-time vision based fabric inspection system," IEEE Trans. Ind. Electronics, **vol. 52**(4), 1073–9, 2005.

[9] S. V. der Jeught and J. J. Dirckx, "Real-time structured light profilometry: a review," Optics and Lasers in Engineering, **vol. 87**, 18 – 31, 2016, Digital optical & Imaging methods in structural mechanics. https://doi.org/10.1016/j.optlaseng.2016.01.011

[10] D. Palouek, M. Omasta, D. Koutny, J. Bednar, T. Kouteck, and F. Dokoupil, "Effect of matte coating on 3D optical measurement accuracy," Optical Materials, **vol. 40**, 02 2015. https://doi.org/10.1016/j.optmat.2014.11.020

[11] M. C. Knauer, J. Kaminski, and G. Hausler, "Phase measuring deflectometry: a new approach to measure specular free-form surfaces," in Optical Metrology in Production Engineering, ser. procspie, W. Osten and M. Takeda, Eds., **vol. 5457**, 366-376, Sep. 2004. https://doi.org/10.1117/12.545704

[12] I. G. O. Kafri, "Moire deflectometry: A ray deflection approach to optical testing," Optical Engineering, **vol. 24**, no. 6, 944 - 960, 1985. [Online]. Available: https://doi.org/10.1117/12.7973607

[13] M. Servn, R. Rodriguez-Vera, M. Carpio, and A. Morales, "Automatic fringe detection algorithm used for Moire deflectometry," Applied optics, **vol. 29**, 3266 - 70, 08 1990.

[14] B. Wang, X. Luo, T. Pfeifer, and H. Mischo, "Moire deflectometry based on Fourier-transform analysis," Measurement, **vol. 25**, 24 - 253, 06 1999. https://doi.org/10.1016/S0263-2241(99)00009-3

[15] A. E.-R. Ricardo Legarda-Senz, "Wavefront reconstruction using multiple directional derivatives and Fourier transform," Optical Engineering, **vol. 50**, no. 4, 1 – 4 , 2011. [Online]. Available: https://doi.org/10.1117/1.3560540

[16] D. Fontani, F. Francini, D. Jafrancesco, L. Mercatelli, and P. Sansoni, "Mirror shape detection by "Reflection Grating Moire Method" with optical design validation," Proceedings of SPIE - The International Society for Optical Engineering 5856, Jun 2005. DOI: 10.1117/12.612114

[17] S.-W. K. Ho-Jae Lee, "Precision profile measurement of aspheric surfaces by improved Ronchi test," Optical Engineering, **vol. 38**, no. 6, 1041 - 1047, 1999. [Online]. Available: https://doi.org/10.1117/1.602147

[18] G. P. Butel, G. A. Smith, and J. H. Burge, "Binary pattern deflectometry," Appl. Opt., **vol. 53**, no. 5, 923 - 930, Feb 2014. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-53-5-923

[19] J. L. H. Willem D. van Amstel, Stefan M. B. Baumer, "Optical figure testing by scanning deflectometry," Proc. SPIE 3739, Optical Fabrication and Testing, 6 September 1999. doi: 10.1117/12.360155, Available: https://doi.org/10.1117/12.369201

[20] S. Krey, W. D. van Amstel, K. Szwedowicz, J. Campos, A. Moreno, and E. J. Lous, "Fast optical scanning deflectometer for measuring the topography of large silicon wafers," Proc. SPIE 5523, Current Developments in Lens Design and Optical Engineering V, 14 October 2004. doi: 10.1117/12.559702, https://doi.org/10.1117/12.559702

[21] K. Ishikawa, T. Takamura, M. Xiao, S. Takahashi, and K. Takamasu, "Profile measurement of aspheric surfaces using scanning deflectometry and rotating autocollimator with wide measuring range," Measurement Science and Technolog, **vol. 25**, no. 6, Apr 2014. [Online]. Available: https://doi.org/10.1088%2F0957-0233%2F25%2F6%2F064008

[22] Y. Tang, X. Su, Y. Liu, and H. Jing, "3D shape measurement of the aspheric mirror by advanced phase measuring deflectometry," Opt. Express, **vol. 16**, no. 19, 15090 - 15096, Sep 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-19-15090

[23] D. Malacara, Optical Shop Testing (Wiley Series in Pure and Applied Optics), Wiley-Interscience, 2007.

[24] A. V. Maldonado, P. Su, and J. H. Burge, "Development of a portable deflectometry system for high spatial resolution surface measurements," Appl. Opt., **vol. 53**, no. 18, 4023 - 4032, Jun 2014. https://doi.org/10.1364/AO.53.004023

[25] C. Horneber, M. Knauer, and G. Husler, "Phase measuring deflectometry - a new method to measure reflecting surfaces," Annual Report Optik, 01 2000.

[26] J. Xiao, X. Wei, Z. Lu, W. Yu, and H. Wu, "A review of available methods for surface shape measurement of solar concentrator in solar thermal power applications," Renewable and Sustainable Energy Reviews, **vol. 16**, no. 5, 2539 - 2544, 2012. https://doi.org/10.1016/j.rser.2012.01.063

[27] C. E. Andraka, S. Sadlon, B. Myer, K. Trapeznikov, and C. Liebner, "Rapid reflective facet characterization using fringe reflection techniques," Proceedings of the Energy Sustainability, 19 - 23, 2009. DOI: 10.1115/1.4024250

[28] C. Devivier, F. Pierron, and M. Wisnom, "Damage detection in composite materials using deflectometry, a full-field slope measurement technique," Composites Part A: Applied Science and Manufacturing, **vol. 43**, no. 10, 1650 - 1666, 2012. https://doi.org/10.1016/j.compositesa.2011.11.009

[29] J. Geng, "Structured-light 3D surface imaging: a tutorial," Adv. Opt. Photon., **vol. 3**, no. 2, 128–160, Jun 2011. https://doi.org/10.1364/AOP.3.000128

[30] M. Kujawinska and J. Wjciak, "High accuracy Fourier transform fringe pattern analysis," Optics and Lasers in Engineering, **vol. 14**, no. 4, 325 - 339, 1991. https://doi.org/10.1016/0143-8166(91)90056-Y

[31] M. Takeda, H. Ina, and S. Kobayashi, "Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry," J. Opt. Soc. Am., **vol. 72**, no. 1, 156 - 160, Jan 1982. https://doi.org/10.1364/JOSA.72.000156

[32] L. Armesto, J. Tornero, A. Herraez, and J. Asensio, "Inspection system based on artificial vision for paint defects detection on cars bodies," in Robotics and Automation (ICRA), 2011 IEEE International Conference on, 1–4, 9-13 May 2011. DOI: 10.1109/ICRA.2011.5980570

[33] C. Tarry, M. Stachowsky, and M. Moussa, "Robust detection of paint defects in moulded plastic parts," Proc. - Conf. Comput. Robot Vision, CRV 2014, 306–312, 2014. DOI: 10.1109/CRV.2014.48

[34] T. Liu, C. Zhou, Y. Liu, S. Si, and Z. Lei, "Deflectometry for phase retrieval using a composite fringe," Optica Applicata, **vol. 44**, no. 3, 451 - 461, 2014. DOI: 10.5277/oa140309

[35] Y. Surrel, "Deflectometry: A simple and efficient noninterferometric method for slope measurement," Xth SEM Int. Congr. Exp. Mech., 2004.

[36] K. Hibino, B. F. Oreb, D. I. Farrant, and K. G. Larkin, "Phase shifting for nonsinusoidal waveforms with phase-shift errors," J. Opt. Soc. Am. A, vol. 12, no. 4, p. 761, 1995. https://doi.org/10.1364/JOSAA.12.000761

[37] E. Stoykova, J. Harizanova, and V. Sainov, "Pattern Projection Profilometry for 3D Coordinates Measurement of Dynamic Scenes," Three-Dimensional Telev., pp. 85–164, 2008. DOI: 10.1007/978-3-540-72532-9_5

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 183-189 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Quantitative Traffic Congestion Analysis Approach in Ahmedabad

Tsutomu Tsuboi*

General Manager, Global Business Development Office, Nagoya Electric Works Co., Ltd., 490-1294, Japan

ARTICLE INFO

Article history:
Received:13 April, 2019
Accepted:24 May, 2019
Online: 10 June, 2019

Keywords:
Traffic Flow
Traffic Congestion
Social Loss

ABSTRACT

This study is the extension of the previous study about "Traffic Service Quantitative Analysis Method under Developing Country" in 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). In the previous study, it is introduced how to make quantitative calculation for traffic congestion by traffic parameters and its characteristics curve such as traffic volume (q) to inverse of vehicle average speed (=1/v). In order to identify the traffic congestion condition, it is focused on vehicle speed ratio which is average speed (v_{ave}) to its free speed (v_f). And the threshold level is 2/3 (=v_{ave}/v_f). This 2/3 value comes from Viscous fluids model between parallel flat plates by using similarity of the viscous fluids flow and the traffic flow in India which is introduced at the CODATU XVII and UMI Conference 2017. This threshold value definition needs more traffic theory back up because its similarity between viscous fluids flow and traffic flow comes from the basis of traffic flow measurement results. In this extension study, it focuses on the occupancy of one of typical traffic parameter for traffic congestion. When it is compared between the traffic occupancy measurement data and Speed Ratio, the Speed Ratio 2/3 is the level of the occupancy 30%, which is used as traffic congestion condition. According to daily based traffic volume and average vehicle speed, the congestion condition is occurred by not only those traffic parameter, it is also considered about time zone traffic condition and its actual traffic condition. One of measurement point is always congested in the early evening even if its traffic volume is smaller than the morning traffic congestion. Therefore it is important to analyze traffic condition by time zone. As the result, it clarifies the relationship between Speed Ratio and occupancy. As the result, there are two types of traffic congestion in Ahmedabad city traffic.

1. Introduction

This study focuses on traffic flow and congestion analysis especially in the under developing country (India). In the previous study, it is proposed the traffic service quantitative traffic congestion analysis method in the International Conference on Advances in Computing, Communications and Informatics (ICACCI) [1]. From the actual measurement traffic data study during more than one month in Ahmedabad city of Gujarat State in India, the major traffic parameters are extracted from the traffic fundamental diagram such as traffic density (k) to vehicle average speed (v) and traffic density (k) to traffic volume (q), which introduced at the International conference CODATU 2017 [2].

As for obtaining the traffic parameters and analysis, the detail is explained in the next section. The previous study of CODATU introduces the unique traffic characteristics which is useful for using Envelopment Observation method because there is clear boundary in those characteristic and it is appropriate way to define the traffic parameters from their unique measurement data. The another phenomenon from the measurement data is vehicle speed ratio which is the ratio of average vehicle speed (v_{ave}) to free speed (v_f). The value of this speed ratio (SR) is valid for understanding congestion level of each road. When S.R. becomes lower than 0.65, the condition of road becomes congested. Author found similarity of this phenomena between the Indian traffic flow model and viscous fluids model between parallel flat plates.

In this extension of the research, more detail analysis is added between SR and occupancy of traffic parameter which is used as traffic congestion level judgement from traffic engineering. In addition, the relationship between parameter occupancy and traffic density is described in order to understand traffic congestion condition. As for the traffic congestion analysis, it also describes

* Tsutomu Tsuboi, 29-1 Mentoku Shinoda Ama, Aichi, +81-902738-3496, t_tsuboi@nagoya-denki.co.jp.

the time zone base analysis in this paper. From this time zone base analysis, we found there are two patterns for Ahmedabad traffic congestion condition. One is heavy traffic density congestion which follows the traffic flow theory and the other is unheavy traffic density congestion which is unique in India. Therefore it is important for considering what is the traffic congestion definition. when the traffic information is provided from the traffic information sign board so called VMS or Various Message Sign on the road.

2. Related Studies

2.1. Traffic Flow Theory and Measurement Analysis of Ahmedabad city

In the previous study for Ahmedabad traffic flow analysis at CODATU conference, the traffic fundamental diagram is introduced such as *k-v* curve (traffic density (*k*) to average vehicle speed (*v*)), *k-q* curve (traffic density (*k*) to traffic volume (*q*)), and *q-v* curve (traffic volume (*q*) to vehicle speed (*v*)). This analysis has been continued since Ahmedabad ITS or Intelligent Transport Systems business project from October 2014. The project is providing Ahmedabad city to provide its traffic condition for the drivers through VMS. The ITS system consist of four VMSs and 14 traffic monitoring cameras. The traffic flow condition is captured by the traffic monitoring cameras and its result of analysis is shown at the VMS for considering change route or using other transportation choice. The Ahmedabad ITS system configuration is shown Figure 1.

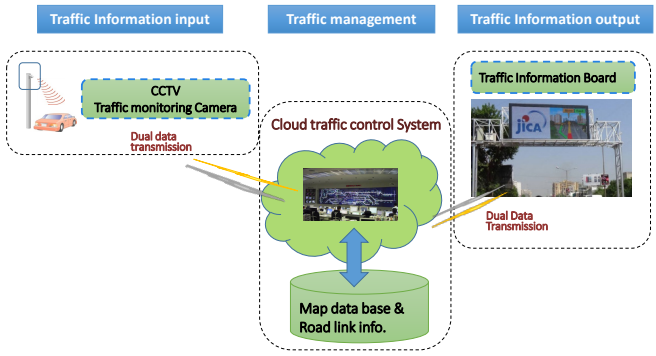


Figure 1: Ahmedabad ITS system configuration

In the CODAU study, authors found the unique characteristics of Ahmedabad traffic flow measurement data. There is clear boundary in each traffic fundamental diagram. The *k-v* curve and *k-q* curve at Camera#1 are shown in Figure 2.

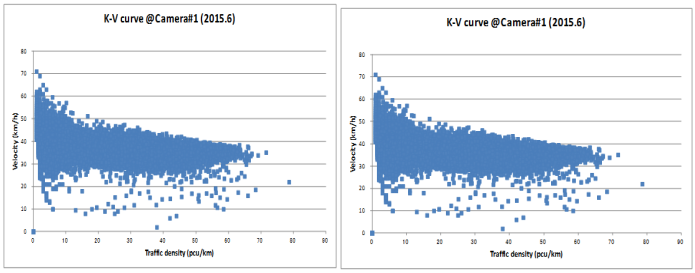


Figure 2: Traffic Flow Characteristics at Camera#1

From Figure 2, there is clear boundary in each characteristics curve by envelop observation. The other characteristics at other location are similar results which are omitted in this paper. When these boundary are taken as traffic flow characterizes, it is able to adjust traffic flow characteristics curve from traffic flow equation to the envelop line. The Figure 3 shows the example of adjustment result between traffic flow theoretical curve from its equation and the measurement data at Camera#1.

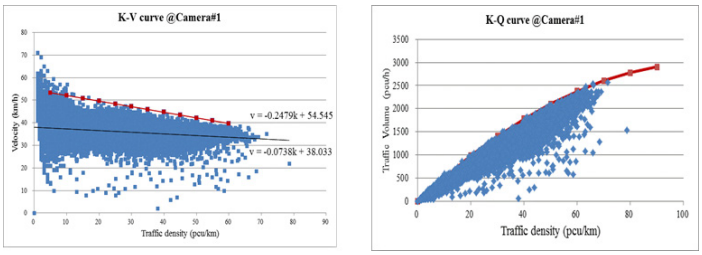


Figure 3: Traffic Flow Characteristics at Camera#1

In terms of traffic equation form the theory, traffic density (*k*) to traffic speed (*v*) equation is given by (1) and traffic density (*k*) to traffic volume (*q*) equation is given by (2). The equation (1) is so called Greenshields [3] curve.

$$v = v_f\left(1 - \frac{k}{k_j}\right) \tag{1}$$

v_f: free speed

k_j: jam traffic density

$$q = -\frac{v_f}{k_j}\left(k - \frac{k_j}{2}\right)^2 + \frac{v_f k_j}{4} \tag{2}$$

The theoretical curve illustration is shown in Figure 4.

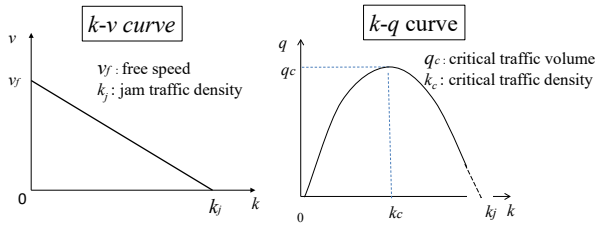


Figure 4: Theoretical Traffic Flow Curve

In the advanced country traffic flow, the actual measurement plots including Japan [4] are aligned with theoretical curve and or near boundary in Figure 4. Therefore the traffic characteristics curve from measurement in Ahmedabad is quite different from those of the advanced countries (refer to Figure 2).

2.2. Quantitative Traffic Congestion Study in Ahmedabad

The traffic congestion problem in developing countries become more serious situation these days, especially in India. Therefore it is crucial to establish for analysis method about the traffic congestion quantitatively. In the International Conference on Advances in Computing, Communications and Informatics or ICACCI, T.Tsuboi proposes quantitative traffic congestion analysis based on the actual traffic measurement in Ahmedabad. In this paper, it is introduced the traffic congestion by vehicle speed ratio (SR) which is the ratio of the average vehicle speed (v_{ave}) to the free speed (v_f). In ICACCI paper, it is defined that the

threshold level of traffic congestion is 0.65. The comparison SR at each traffic camera location is shown in Figure 5. According to Figure 5, the Camera#2 is most small value. And the actual daily traffic flow (traffic volume and average vehicle speed) at Camera#2 in June 2015 is shown in Figure 6 (a) and (b). From Figure 6(b), the average vehicle speed is lower than 20km/h, which shows the traffic congestion occurred [5].

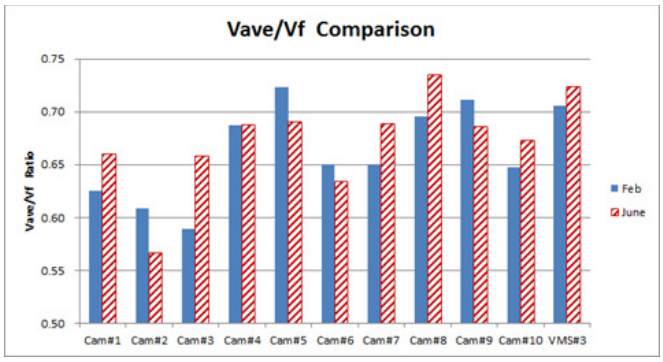
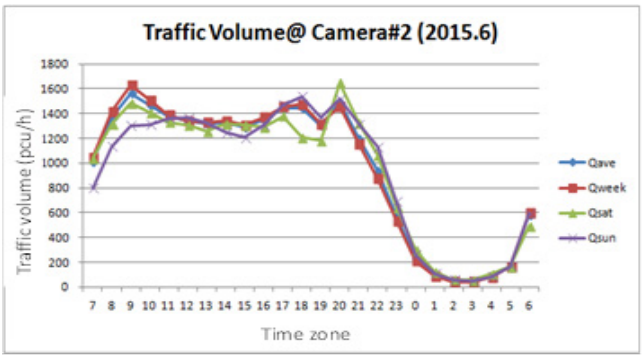
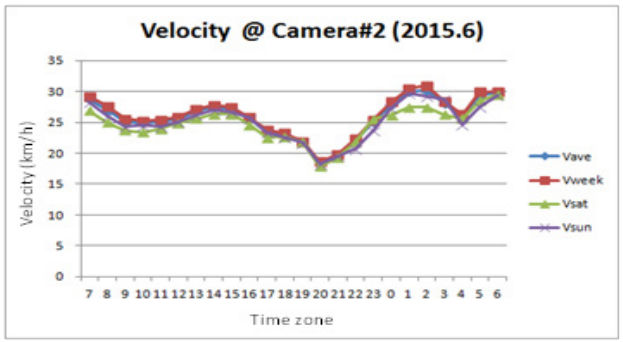


Figure 5: Traffic Speed Ratio Comparison in Ahmedabad



(a) Traffic Volume Time Zone condition at Camera#2



(b) Traffic Speed Time Zone condition at Camera#2

Figure 6: Daily Traffic Flow condition at Camera#2

From Figure 6 (b), it is a clear difference that Camera #2 vehicle velocity ratio is lower than other locations. According to traffic flow theory, most of measurement data of q–k curve is located at shadow area in Figure 7. And free traffic flow area is bottom part of q – k curve; jam traffic flow area is top of q – k curve. But our measurement data of q – k curve in Camera # 2 is located inside q – k curve (refer to Figure 7). The speed ratio of inside curve area is lower than 0.65 which means that its traffic condition becomes congested.

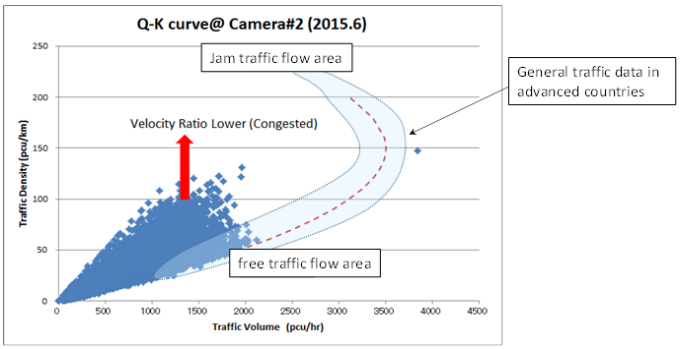


Figure 7: q – k Curve at Camera#2

In order to identify that inside curve area is under congested condition, we divide between two areas. One area has speed ratio equal and less than 065 and the other area has speed ratio more than 0.65. Figure 8 shows those two areas. According to Figure 8, the speed ratio equal and lower than 0.65 (≈ 2/3) is inside envelope line area. According to the previous traffic flow analysis, the threshold between free flow and congested flow occurs at the average vehicle velocity (v_{ave}) ≤ 2(v_f) /3 [2]. Therefore the traffic flow capacity is able to obtain by our envelope data analysis and it is able to be described by the traffic flow equation. This is the first time to show emerging countries traffic analysis and brings basic traffic flow parameters by boundary traffic analysis. And vehicle velocity ratio of average speed by free speed is good reference of explaining traffic jam condition; the critical value of speed ratio is 0.65 in this paper.

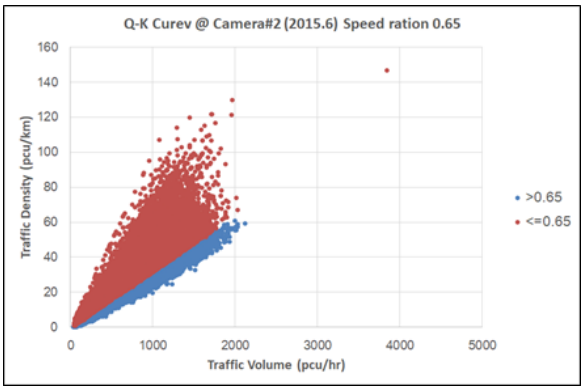


Figure 8: q – k measurement data by vehicle Speed Ratio at Camera#2

In terms of traffic congestion analysis, the congestion social loss analysis theory [6] is used. From the traffic flow theory, traffic volume (q) is mathematical product by vehicle velocity (v) multiplying to traffic density (k) shown in (3).

$$q = v \times k \tag{3}$$

Since the vehicle velocity (v) is inverse of travel time (t), then the vehicle velocity is described as $v = 1/t$. Therefore traffic volume (q) is a function of travel time (t). Assuming that the traffic user's opportunity n per unit time is constant, travel time (t) is proportional to the cost of traffic congestion. In another word, the relationship between traffic volume (q) and travel time (t) means the traffic supply curve of the relationship between traffic

service supply and cost. In order to analyze traffic volume (q), differentiating e (3) by (t) provides (4).

$$\frac{dq}{dt} = -\left(v\frac{dk}{dv} + k\right)\frac{1}{t^2} \tag{4}$$

When vehicle velocity (v) becomes large volume (travel time becomes small value), traffic density (k) becomes small value and then $dk/dv < 0$. Therefore he value in parenthesis of (4) is negative and then dq/dt becomes positive value. This relationship is illustrated in Fig.8. The supply curve line rises to the right from travel time ($1/v_f$) to ($1/v_c$), which means private cost curve for transportation. And when (v) becomes small value (travel time (t) becomes large value), the value in parenthesis of (4) becomes positive and the supply curve drops to the right, which corresponds to private cost curve drops to the right.

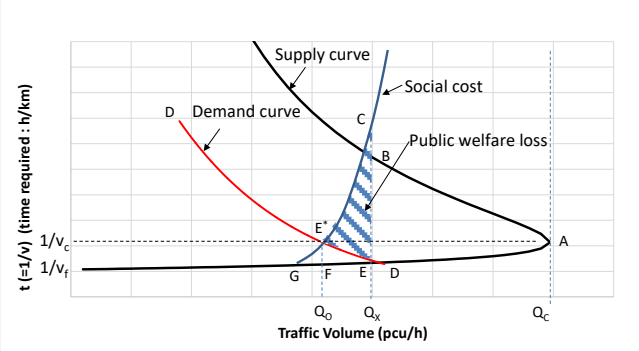


Figure 8: Road Traffic Service Market

In Figure 8, the point A is called hyper congestion condition at critical traffic volume (q_c) when traffic volume becomes larger beyond point A, travel time takes longer because of traffic congestion. There are two travel time (t) except point A. For instance, there are point E and point B at a certain traffic volume (Q_x). In case of point B, the travel time takes longer than that of at point E. If traffic demand curve D-D is given, the point E is cross point between demand curve D-D and supply curve which provides balance condition between traffic demand and supply. If social cost curve is given, the point E* becomes balance point between traffic demand and social cost. Then area CE*E provides traffic service cost loss caused by traffic congestion because infrastructure should cover at the level of traffic volume (Q_x) at the point E and social cost rises at the point C where its traffic volume (Q_x) is same as that of at the point E. Therefore area CE*E is defined as the public welfare loss which means traffic congestion social loss.

In order to calculate the public welfare loss from measurement data, it is necessary to define condition for each curve of Figure 8 from the actual measurement analyzed data. The condition of traffic congestion social loss is as follows:

- Supply curve is $q-v$ curve form envelopment observation.
- The critical speed (v_c) is used at SR=0.65 because this point is congestion threshold.
- Demand curve D-D is used from actual measurement traffic data.
- The social cost curve is unknown under this condition. Therefore BE* liner curve is taken.
- The congestion social loss is defined as area BE*E instead of area CE*E.

The Figure 9 shows the congestion social loss calculation example at Camera#2 with this quantitative social loss analysis. The result of Social loss analysis at all locations shows in Table 1. According Table 1, it is clear relationship between Speed Ratio (SR) and Social loss value.

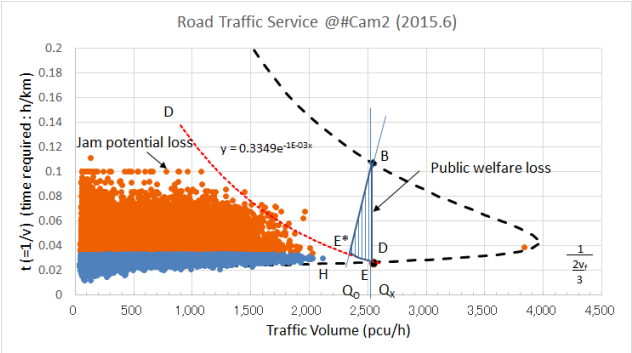


Figure 9: Road Traffic Service Market at Camera#2

Table 1: Social Loss and Speed Ratio

Location	SR(v_{ave}/v_f)	Social Loss
Camera#1	0.66	3.3
Camera#2	0.57	8.1
Camera#3	0.66	3.1
Camera#4	0.69	3.2
Camera#5	0.69	2.0
Camera#6	0.63	2.2
Camera#7	0.69	1.3
Camera#8	0.73	0.2
Camera#9	0.69	1.6
Camera#10	0.67	2.4
VMS#3	0.72	0.7

3. Traffic Congestion Analysis

In the previous section, the envelopment observation method is used for getting traffic flow equation from its measurement. Based on the each traffic flow equation, new quantitative social loss of traffic congestion analysis is introduced. In terms of traffic congestion judgement, the average vehicle speed is used. This average vehicle speed is different at each road and its condition. Therefore it is necessary to have some relative vehicle speed compared with normal driving speed which is taken as free speed in this paper. As for traffic congestion judgement, the average speed is not only parameter but also the traffic occupancy for example.

3.1. Traffic Occupancy

The traffic occupancy is used for one of parameters which shows the level of traffic congestion condition. There are two traffic occupancies in the theory, time occupancy and space occupancy. In real traffic management, time occupancy is often used for the freeway [7]. In this section, time occupancy is used.

The time occupancy is defined the percentage ratio between total measurement time (t) of the vehicles to a certain block of road section under certain time frame. The formula of time occupancy is provided by (5).

$$\mathrm{OC} = \frac{1}{T}\sum_{i} t_i \times 100\ (\%) \tag{5}$$

where T is time of measurement, (t_i) is detected time of vehicle (i)[8].

When number of existing vehicle a certain section is (N), average length of vehicle is ($\bar{l}$) in (6) is given.

$$OC = 100\frac{q}{v}\bar{l} = 100k\bar{l} \tag{6}$$

Therefore occupancy (OC) is proportional to traffic density (k) and traffic volume (q).

From the one month measurement data of Camera#1 and #2 in June 2015, traffic density (k) to occupancy (OC) relationship are shown in Figure 10. According to Figure 10, the relationship between (k) and (OC) is proportional.

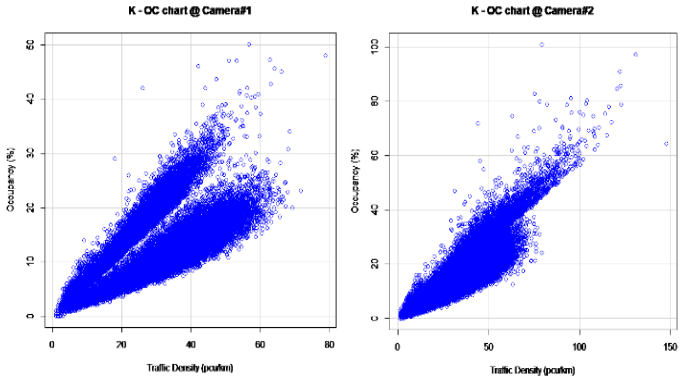
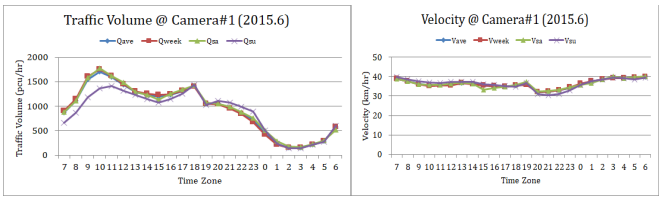


Figure 10: Example of k – q curve at Camera#1 and Camera#2

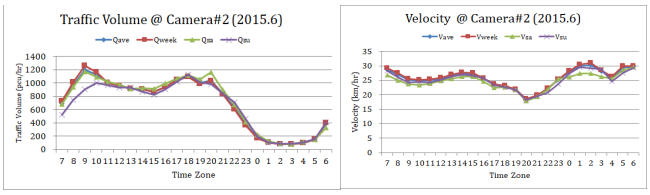
From Figure 10, it seems that the measurement data is plotted into two groups. In the section 3.3, time zone analysis is introduced because traffic congestion is dependent with time of a day (refer to Figure 6) [9].

3.2. Daily Traffic Condition

In this section, a daily traffic changes of traffic volume and velocity is described. In Figure 11, there are two cases of daily based traffic condition at Camera#1 and #2. The graph shows one month traffic volume changes from 7:00 am to next day 6:00 am and average vehicle velocity changes. There are two peaks of traffic volume in each case. One peak occurs between 7:00 am to 10:00 am and the second peak occurs between 18:00 to 20:00. When vehicle average speed is concerned, there is not so much changes of velocity in Camera#1 but different speed at Camera#2. In case of Camera#2, there is a drop during the second peak volume of traffic. And each graph shows classification by four cases—total one month average, weekday, Saturday, and Sunday. According to Figure 11, there is traffic jam in Camera#2 between 18:00 to 20:00 and the average vehicle speed goes down to less than 20 km/h which is confirmed the speed under heavy traffic congestion through the recorded traffic monitoring video data.



(a) Daily Traffic Condition at Camera #1



(b) Daily Traffic Condition at Camera#2

Figure 11: Daily Traffic Condition at Camera#1 and Camera #2

3.3. Occupancy and Traffic Density

As described in section 3.1, time zoning is adapted into Figure 10. In this section, six time zone is defined as shown in Table 2

Table 2: Social Loss and Speed Ratio

Zone Name	Time Zone
T1	7:00 — 10:59
T2	11:00 — 14:59
T3	15:00 — 18:59
T4	19:00 — 22:59
T5	23:00 — 2:59
T6	3:00 — 6:59

There are 6 time zones such as T1 is defined time frame from 7:00 am to 10:59 am which is most congested time frame as we see in Figure 11. And T4 is defined time frame from 19:00 to 22:59 which is the second congested time frame. In case of using this Time Zone classification for k - OC curve, Time Zone based k - OC curve is shown in Figure 12. In this paper, we focus on the characteristics at Camera#1 and #2.

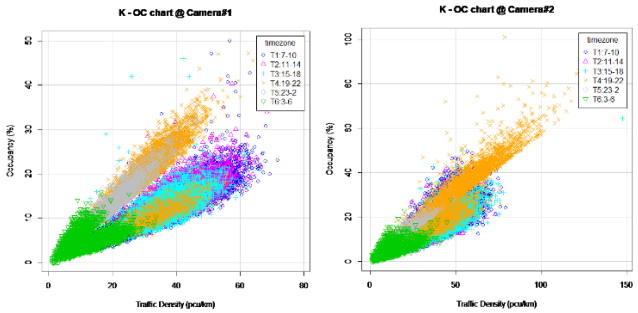


Figure 12: Time Zone based k – OC curve at Camera#1 and Camera#2

When the Time Zone based analysis, it is clear two groups data plot in Camera#1, where there is no traffic congestion condition compared with that of Camera#2. From the result of Camera#1 analysis, the upper group in k – OC curve comes from majority of

T4 (19:00-22:59). The lower group in k –OC curve comes from T1, T2, and T3. In another words, upper part of T4 is relative lower density compared with lower part of T4. This means there are two types of traffic congestion condition. Therefore one type of traffic congestion is defined as low density traffic congestion (herein after it is called as Low density Congestion) and the other one is higher traffic density congestion (herein after it is called as High density Congestion).

The next section describes traffic volume to occupancy and average vehicle speed to occupancy in order to understand the traffic congestion condition with other traffic parameters.

3.4. Occupancy and Traffic Volume and Speed

In Figure 11, it is shown about the daily traffic condition by traffic volume and average vehicle speed. In this section, it adapts Time Zone basis analysis for q-OC curve and v-OC curve in order to understand relationship between traffic volume to occupancy and average vehicle speed to occupancy.

The Traffic volume (q) to Occupancy (OC) Time Zone basis relationship at Camera#1 and #2is shown in Figure 13.

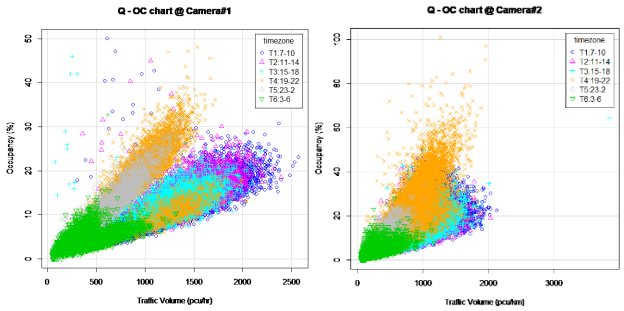


Figure 13: Time Zone based q – OC curve at Camera#1 and Camera#2

From the q – OC curve at Camera#1 in Figure 13, it shows more clear about difference between Low density congestion area and High density congestion area. The Low density congestion occurs in T4. It looks same trend of characteristics at Camera#2 but there is no clear two groups separation between Low density congestion and High density congestion.

The next Figure 14 shows average vehicle speed (v) to Occupancy (OC) relationship.

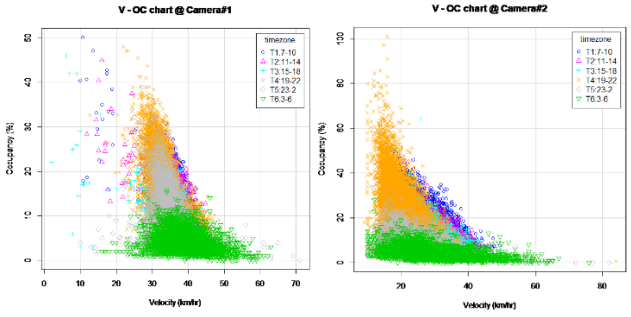


Figure 14: Time Zone based v – OC curve at Camera#1 and Camera#2

From Figure 14, it is clear that traffic congestion at lower than 30km/h occurs above 30% Occupancy condition.

Here is a question why Low density congestion occurs. It is considered at the next discussion section.

4. Discussion of Low density congestion

According to all traffic measurement location in Ahmedabad, the location Camera#2 is the most congested condition location (refer to Figure 5). And traffic condition during T4 time zone becomes most congested as described in previous section. The traffic volume at T4 of Camera#2 is the second peak of traffic volume in a daily condition as shown in Figure 6 (b) and Figure 11 (b). The highest peak of traffic volume occurs at T1 and T2. Therefore it is possible to define that High density congestion occurs at T1 and T2 and Low density congestion occurs at T4.

There is one more traffic parameter which is called Headway. The Headway is defined the distance between the front of driving vehicle group and the front of the second driving vehicle group (refer to Figure 15).

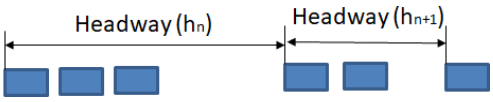


Figure 15: Headway definition

It is relative true the relationship between the small value of Headway and traffic congestion because its traffic density is large. The Headway ($\bar{h}$) and the traffic density relationship is described (7) from the theory.

$$\bar{h} = \frac{1}{k} = \frac{v}{q} \tag{7}$$

where $\bar{h}$ is average of headway. From (7) and (6), (8) is given.

$$\bar{h} = \frac{1}{k} = \frac{100\bar{l}}{OC} \tag{8}$$

From (8), the headway ($\bar{h}$) is inverse proportional to occupancy (OC) with the average of vehicle length ($\bar{l}$). The average of vehicle length is difficult to get from real measurement. Therefore the relationship between the average headway ($\bar{h}$) to occupancy (OC) is not exactly inverse relation. The Figure 16 shows $\bar{h}$ - OC curve at Camera#1 and #2. The vertical axis is legalism for clear display.

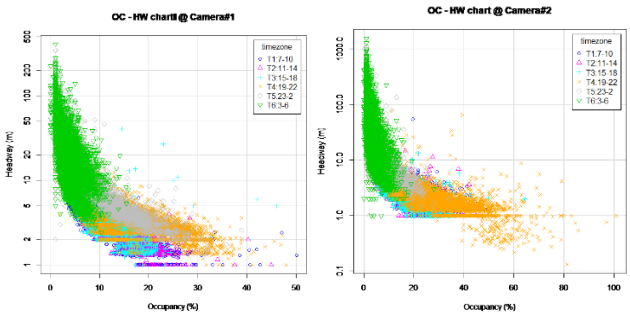


Figure 16: Time Zone based

From Figure 16, the traffic congestion occurs at T4 Time Zone and its Occupancy is above 30%. In this analysis, it is difficult to clarify between Low density congestion and High density congestion. The difference of Low density congestion and High density congestion becomes clear from traffic volume to occupancy or traffic density to occupancy relationship.

When it check the Time Zone of Low density, T4 and time period is from 18:00 to 22:59. After discussion with local traffic

police, they observe that the Low density congestion may be caused by limitation of parking space in the city. In the evening after work, the residents of Ahmedabad stop by restaurants and or shopping place and they make parking their vehicle at not permitted place such as along the street. Therefore this parking issues have potential for creating Low density congestion. If so, there is a solution for Low density congestion by preparing appropriate parking space for restaurant and shopping. The reason why the Low density congestion is detected by the traffic monitoring cameras is assumed that the collecting data is based on number of driving vehicles, not parking vehicles. For the traffic monitoring cameras, they take the vehicles as a part of side roads, which means parking vehicles make roads just narrow path and is unable to count the vehicles number as their traffic density. This situation tells that it is important to judge the traffic monitoring data by not only numerical data from the monitoring camera but also actual visual data.

In terms of quantitative social loss analysis method, the demand curve in Figure 9 is used from the measurement data, which is described as the condition of the quantitative social loss analysis method. When Figure 9 is mapped by time zone, it becomes Figure 17.

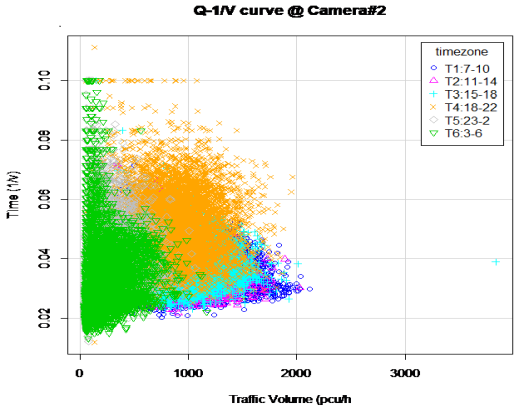


Figure 17: Time Zone based q-1/v curve at Camera#2

From Figure 17, Low density congestion area at T4 is located inside the demand curve. Therefore it is able to confirm proposed quantitative social loss analysis is valid.

5. Summary

In this study, there are three features for Indian traffic flow analysis.

- Feature 1: Introduce envelop observation method based on traffic flow data to provide traffic flow fundamental characteristics by traffic flow equation. It is able to get traffic flow parameters from this equation i.e. jam density (k_j), free speed (v_f).
- Feature 2: Propose quantitative social loss of traffic congestion method by only using traffic measurement data. As for critical speed (v_c), it is defined by Speed Ratio (SR)—average speed (v_{ave}) to free speed (v_f)—0.65 or 2/3 of (v_f), which is based on traffic flow analysis in India.
- Feature 3: There are two types traffic congestion condition in India, Low density congestion and High density congestion. It is presumed that Low density congestion mainly is caused by parking vehicle problem. Therefor Indian Traffic congestion occurs not only by large number of vehicles but also other infrastructure issues.

All this study analysis is based on one month of real traffic flow data in Ahmedabad city. It is necessary to have more different timing, location i.e. place, city, country. It is also important to have traffic video data in order to evaluate traffic congestion condition and understand congestion reason. The above all research items are future work.

Conflict of Interest

There is no conflict of interest in this study.

Acknowledgment

This study also underwent the program ID:JPMJSA1606 of the International Science and Technology Cooperation Program (SATREPS) challenges for global challenges in 2016.

Special appreciation to Mr. Kikuchi.C and Mr. Mallesh.B of Zero-Sum ITS India for providing traffic data in Ahmedabad.

References

[1] T. Tsuboi, “Traffic Service Quantitative Analysis Method under Developing Country” IEEE Xplore, 2018. DOI;10.1109/ICACCI.2018.8554495.

[2] T. Tsuboi, N. Yoshikawa, “Traffic Flow Analysis in Emerging Country (India)” in the CODATU XVII and UMI Conference, Hyderabad, India, 2017.

[3] B.D. Greenshields, A Study of Traffic Capaci, Proc.H.R.B., 14, pp.448-477, 1935.

[4] H. Kubota, T. Oguchi, K. Takahashi.” Traffic Engineering•Traffic Planning,” Rikoh Tosho Book, pp.24-25, April.2010.

[5] F. He, X. Yan, Y.L.L.Ma, “A Traffic Congestion Assessment Method for Urban Road Networks Based on Speed Performance Index”, Procedia Engineering 137 pp. 425 – 433, 2016.

[6] T. Kuroda, et al: Urban and Regional Economics, Yuhikaku Book, pp. 265-268, 2008M. H. Nehrir, C. Wang, Modeling and Control of Fuel Cells: Distributed Generation Applications, Wiley-IEEE Press, 2009.

[7] F.B. Lin, C.W. Sy, H.H. Huang, “Uniform Criteria for Level of Service analysis of Freeways.” Journal of Transportation Engineering, ASCE, 122(2), pp.123-130, 1996

[8] McMaster University, Department of Civil Engineering and Department of Geography, “TRAFFIC STREAM CHARACTERISTICS BY FRED L. HALL”, p.10

< https://www.fhwa.dot.gov/publications/research/operations/tft/chap2.pdf>

[9] T.Tsuboi, “Traffic Congestion Visualization by Traffic Parameters in India”, International Conference on Innovative Computing and Communication (ICICC), 2019

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 190-193 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Photoluminescence Investigation of Inhomogeneous Porous P-type Si

Róbert Brunner[1,*], Pavel Vojtek[2], Zuzana Zábudlá[2], and Emil Pinčík[1]

[1]*Institute of Physics SAS, Dúbravská cesta 9, 845 11 Bratislava, Slovak Republic*

[2]*Department of Experimental Physics, Faculty of Mathematics, Physics, and Informatics, Comenius University, Mlynská dolina F2, 842 28 Bratislava, Slovak Republic*

ARTICLE INFO

Article history:
Received:20 March, 2019
Accepted:22 May, 2019
Online: 18 June, 2019

Keywords:
Photoluminescence
Inhomogeneous film
Porous Si
Electrochemical etching

ABSTRACT

Photoluminescence (PL) of inhomogeneous porous silicon (PS) of p-type is investigated in this contribution. We measured the PL signal at equidistant positions separated by 0.05 mm in area localized between original crystalline Si (c-Si) wafer surface and electrochemically prepared PS layer. Two PL peaks localized at energies 1.8 and 1.9 eV were identified and their parameters were determined. The changes of PL maxima energy ("blue shift") in dependence on position were observed.

1. Introduction

Researching porous silicon (PS) structures is time-intensive. Extensive collection of results is presented e.g. in [1], but the interpretation of the origin of photoluminescence (PL) signals is still a problem.

Structural characteristics and luminescence of PS prepared by chemical etching were studied by Korsunskaya et al. in [2]. The luminescence spectrum of this type of PS was the result of superposition of two PL bands. The first band represents excitonic recombination in amorphous Si nanocluster smaller than 3 nm. The second band, which prevails at room temperature, is generated by the recombination of charge carriers mediated by the defects in silicon oxide.

Silicon nanocrystallines (NCs) are included in the PS in the form of nanowires on monocrystalline silicon surface composed of different phases of crystalline and amorphous Si. They are covered with oxides (SiO_y) and hydrides (SiH_x) (Tynyshtykbaev [3]). Authors concluded that the light emission originates in hydride coverages of nanocrystallines. In order to explain the character of PL centers, various models have been elaborated (see e.g. Lenshin [4], experimental analysis of aging effects). The model, in which the maximum of PL localized at 640 nm is related to defective levels os SiH_x and SiO_y components, is greatly accepted. The presence of groups SiO_xH_y, observed using FTIR spectroscopy in IR region, was reported in the work [5]. Only small influence of PL on surface morphology was found.

The homogeneity of effective PL from PS was studied by Kayahan [6]. Imaging spectroscopy technique was used by author in order to examine both the effect of atmospheric aging and light illumination applied during anodization process on spatial distribution of PL intensity. He concluded that chemical structure of PS surface plays role in PL and that quantum confinement effect related to surface states can be understood to be a source of luminescence.

Chan [7] observed dependence of energy of PL spectrum on size of pores in PS. Their size was controlled by concentration of HNO_3 in etching liquid. Observed "blue shift" of peaks was attributed to quantum confinement effect.

In this work are presented results of our PL experiments on PS. For evaluation of PL signal we used the decomposition procedure based on the fitting of the spectrum by the set of Gaussian peaks. The proper number of the peaks was estimated from the residuum of the fit (Brunner [8]).

2. Experimental Part

The contribution topic is to investigate the PL signal recorded on an inhomogeneous PS sample. The p-type boron doped Si wafers (100) with resistivity of 8 – 10 Ωcm and thickness 600 –

* Róbert Brunner, Email: robert.brunner@savba.sk

650 nm were used in the experiment. Standard PS was prepared by electrochemical etching process in the solution of HF and MeOH without illumination. Used apparatus is schematically depicted in the Figure 1. Electrical voltage 6V – 11V was applied through mask (Fig. 1) between the exposed area of the Si sample (positive potential) and Pt negative electrode. A constant current of 30 $mAcm^{-2}$ was maintained during the few minutes of the etching process. The PS was rinsed in ethanol and annealed in nitrogen at 250°C for 30 minutes.

Optical and electrical parameters of PS layer depend on the type and size of the formed crystallites that are controlled by etching process parameters. The mask affects current density in vicinity of their boundary. This shielding effect leads to generation of transient area between exposed and non-exposed part of surface of silicon wafer.

Measurements were performed in this transient area in 17 points separated by 0.05 mm starting from initial crystalline surface (after polishing and cleaning procedures) to prepared PS layer as shown in the Figure 2.

The PL of the sample was measured at ambient temperature. PL spectra were decomposed into Gaussian peaks (this process is explained in more details in [8, 9, 10]). Usage of Gaussian peaks is suitable for sufficiently homogeneous structures composed of domains (layers). Photon emission is considered as an event which is independent of other emission acts. This independence makes possible to use the central limit theorem, leading directly to the Gaussian profile of the PL peak. The presence of light emitting centers of different types in the sample gives the spectrum composed of a set of Gaussians.

The observed spectra were fitted in order to find parameters of the set of Gaussian functions (energy, intensity and width given as FWHM) and to calculate constant bias. This bias represents both noise and signal offset contribution. Root-mean-square (RMS) value was used in the role of the object function in the fitting process.

The calculated parameters of PL peaks (without bias) were statistically tested using bootstrap method (e.g. [11]) based on measured data set resampling. Confidence intervals of peak parameters were estimated using BCa percentile method ("bias-corrected and accelerated"). This version reduces the requirements on number of the used bootstrap data set. In our calculations we used 500 resampling procedures for each sample. Standard 5% confidence limit was used for confidence intervals.

3. Results and Discussion

Estimation of optimal number of PL peaks during fitting process was based on size of residuum. Figure 3 shows how residuum depends on the number of used peeks. Well visible "knee" indicates the optimal model. This approach is explained in more detail in [8] (but briefly speaking, usage of too small number of peaks omits relevant parts of the PL signal, whereas redundant peaks fit mainly the noise). In this case, two relevant PL peaks were found: ~1.8 eV and ~1.9 eV.

Figures 4 – 6 show dependence of parameters of PL peaks (in successive steps energy, intensity and FWHM). Signal was measured at ambient temperature. The visible dependence of parameters of peaks on the position of the measurement point is attributed to the lateral non-homogeneity of the PS layer in the transition region (Fig. 2). Calculated confidence intervals are slightly non-symmetrical. Possibility to evaluate such data sets is the main advance of the bootstrap method, because the approach based on the normal error distribution would be less effective.

Several trends were observed:

Increase of the PL peak energy in dependence on position of measurement point is visible in Figure 4. We suppose that this may be the consequence of the size effect caused by the etching of nanostructural elements. Namely it is related to the SiO_xH_y compounds covering the structure's surface. The presence of the SiO_xH_y has been confirmed in the works [5, 12] using FTIR spectrometry of the samples of this type in wavenumber range 790-1250 cm^{-1} (0.098 - 0.155 eV).

A noticeablw changes of power of PL signal in dependence on position of measured point can be seen in Figure 5. This is mainly the consequence of increased thickness of light-emitting PS layer in combination with interference of exciting light during PL experiment, but the size effect (controlling light conversion effectivity) can take place in this process also.

Observed changes of FWHM of peaks 1 and 2 on Figure 6 are close to zero and may be non-significant in comparison with the width of confidence interval. Otherwise the increase of FWHM may indicate the inhomogeneity of the sample.

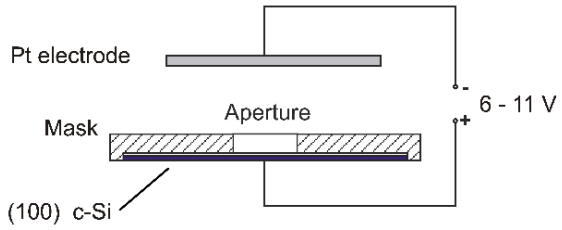


Figure 1: Experimental setup

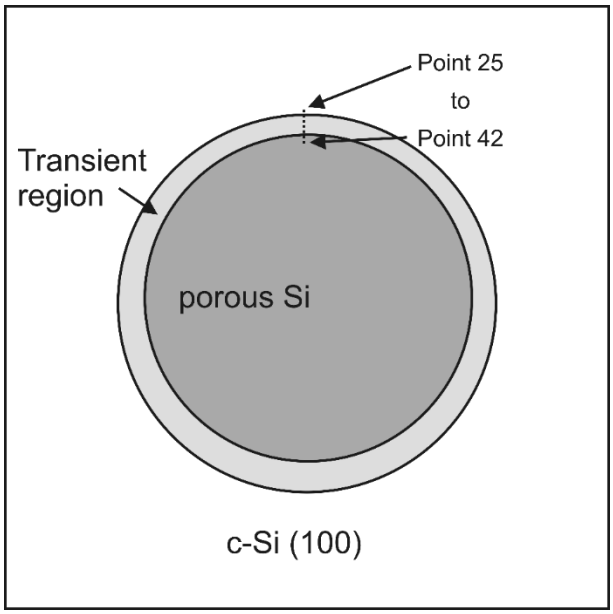


Figure 2: Sample geometry and measured area

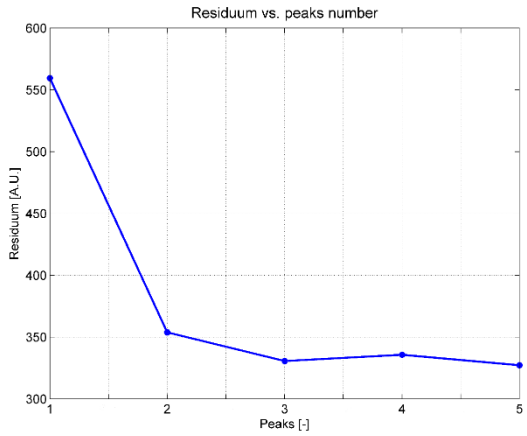


Figure 3: Residuum as a function of number of peaks

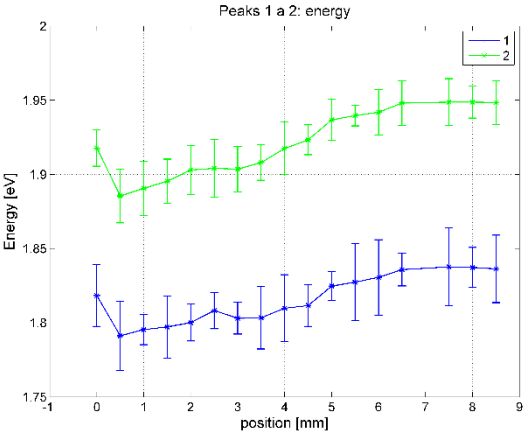


Figure 4: Dependence of energy on the position. Confidence intervals at confidence level α=0.05 are given

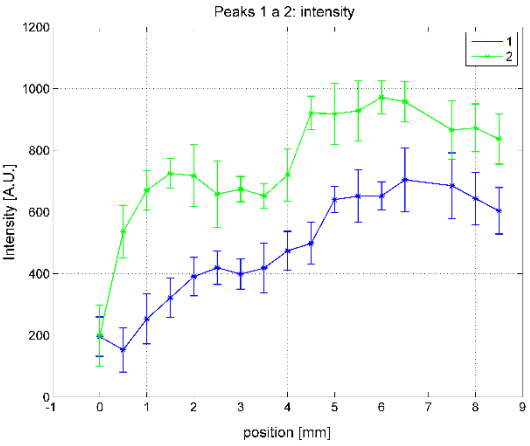


Figure 5: Dependence of intensity on the position. Confidence intervals at confidence level α=0.05 are given

4. Conclusion

Visible increased "blue shift" of both PL maxima (1: at 1.9 eV, 2: at 1.8 eV) occurred in all records in dependence of location of measuring point has been observed. We suppose that this effect is mainly related to the SiO_xH_y compounds that cover the PS structure's surface. This structure was modified during preparation of the sample by the non-homogeneous density of the current in transient region (in the vicinity of the mask boundary, see Fig. 1).

The corresponding FWHM of the peak marked as 1 is decreasing, whereas the value of the peak 2 is increasing in dependence on the position of measuring point. These effects are weak, however, the increase of the FWHM parameter may indicate the degradation of the sample homogeneity.

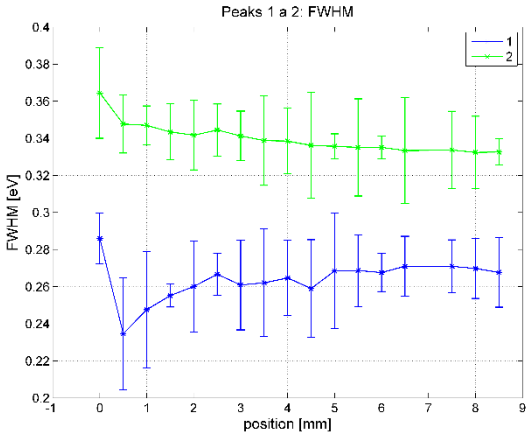


Figure 6: Dependence of width on the position. Confidence intervals at confidence level α=0.05 are given

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

The following agencies partly financially supported this contribution: APVV – project No. 15-0152 and VEGA – projects Nos. 1/0900/16, 1/0676/17, 2/1061/18, and 2/0149/18.

References

[1] Porous Silicon – from Formation to Application, Volume One, edited by Ghenadii Korotcenkov, CRC Press, Taylor and Francis Group, Boca Raton, London, New York, 2016.

[2] N.E. Korsunskaya, T.R. Stara, L.Yu. Khomenkova, K.V. Svezhentsova, N.N. Melnichenko, and F.F. Sizov, The nature of emission of porous silicon produced by chemical etching, Semiconductors, 44, 79-83, 2010.

[3] K.B. Tynyshtykbaev, V. B. Glazman, M.A. Yeleuov, A.T. Isova, B.A. Rakymetov, D. Muratov and S.Z. Tokmoldin, Photoluminescence of Por-Si with High-ordered Mosaic Structure Received at Long Anodic Etching p-Si (100) in the Electrolyte with an Internal Current Source, Optics and Photonics Journal, 3(2B), 217-221, 2013. doi: 10.4236/opj.2013.32B051.

[4] A.S. Lenshin, V.M. Kashkarov, S.Y. Turishchev, M. Smirnov, E. Domashevskaya, Influence of Natural Aging on the Photoluminescence of Porous Silicon, Technical Physics, 57(2) , 305–307, 2012.

[5] M. Kopani, M. Mikula, D. Kosnac, P. Vojtek, J. Gregus, E. Vavrinsky, M. Jergel, E. Pincik: Effect of etching time on structure of p-type porous silicon, Appl. Surf. Sci. 461, 44-47, 2018. doi: 10.1016/j.apsusc.2018.04.228.

[6] E. Kayahan, Investigation of Photoluminescence Homogeneity Distribution of Nano-Porous Silicon by Imaging Spectroscopy, Journal of Natural and Applied Sciences, 21(3), 738-742, 2017. doi: 10.19113/sdufbed.19374.

[7] K.S. Chan, Dwight T.J.E, Photoluminescence, morphological and electrical properties of porous silicon formulated with different HNO3 concentrations, Results in Physics 10, 5–9, 2018, https://doi.org/10.1016/j.rinp.2018.04.011.

[8] R. Brunner, E. Pinčík, H. Kobayashi, M. Kučera, M. Takahashi, J. Rusnák, On photoluminescence properties of a-Si:H-based structures, Appl. Surf. Sci. 256, 5596-5601, 2010. http://dx.doi.org/10.1016/j.apsusc.2010.03.036.

[9] E. Pincik, R. Brunner, H. Kobayashi, M. Mikula, P. Vojtek, J. Greguš, Z. Zabudla, K. Imamura, P. Švec Jr., The photoluminescence of multicolor

silicon, Journal of the Chinese Advanced Materials Society, 4(2), 152-171, 2015- http://dx.doi.org/10.1080/22243682.2016.1151374.
[10] M. Takahashi, T. Fukushima, Y. Seino, W.-B. Kim, K. Imamura, and H. Kobayashi, Surface Structure Chemical Transfer Method for Formation of Ultralow Reflectivity Si Surfaces, J. Electrochem. Soc. 160(8), H443-H445, 2013. doi:10.1149/2.044308jes.
[11] B. Efron, R.J. Tibshirani: An Introduction to the Bootstrap. Chapmann&Hall, New York, 1993.
[12] M. Kopani, M. Mikula, D. Kosnac, J. Gregus, and E. Pincik: Morphology and FTIR spectra of porous silicon, Journal of Electrical Engineering 68, 53-57, 2017. doi: 10.1515/jee-2017-0056.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 194-197 (2019)

www.astesj.com

Special Issue on Recent Advances in Engineering Systems

ASTESJ

ISSN: 2415-6698

Digitization Impact Assessment Model for Secondary Schools: Case of Nairobi County in Kenya

Nicodemus Kalugho Mwambela, Simon Nyaga Mwendia*

KCA University, Faculty of Computing and Information Management, 00200, Kenya

ARTICLE INFO

Article history:
Received:15 December, 2018
Accepted:25 May, 2019
Online: 18 June, 2019

Keywords:
Digitization impact index
Secondary school
Model

ABSTRACT

Digitization is viewed as the capability to utilize digital technologies for generating processing, sharing and transacting information. Studies show that digitization has now become ubiquitous. In almost all the sectors, we routinely interact with digital technologies. As a result, the impact of digitization is being observed in such sectors. For instance, the positive impact of digitization in education sector includes more efficient administration, reduction in workload, better accessibility to information, better management of school resources and enhanced quality of reports. However, in the context of secondary schools, the literature is conspicuously silent on the formal way of assessing the impact of digitization on both the academic performance and service delivery. As an attempt to address the gap, this paper describes a model that can evaluate the impact of digitization in such schools. Python programming tools were used to demonstrate implementation of the model. The model can help governments to formulate policies related digitization of secondary schools and can also be used by funding organization to demonstrate accountability when making investment decisions.

1. Introduction

According to [1] digitization refers to the capability to utilize digital technologies for generating, processing, sharing and transacting information. Previous studies show that digitization in secondary schools is realized by integrating information communication technology(ICT) tools like School Management information systems (MIS) to enhance efficiency of providing administration services. Examples of such services are formulating strategic plans, evaluating staff performance, communication and distribution of resources [1, 2].

Integration of ICT tools in education calls for the development of indicators to monitor the impact of such tools. Both funding sources and the public can use the indicators assess accountability. In addition, there is a need for tools for indicating how use of technology in education is promoting empowerment, creativity, equality, efficiency among learners. Many academic researchers have attempted to address these needs at theoretical and empirical levels but have encountered two difficulties. First, student performance is hard to measure and still there is no common understanding about its definition. Second, ICT is characterized by changing or evolving technologies and their impacts are difficult to separate from their setting. Therefore, the association between the use of ICT and educational performance is not clear, and inconsistent findings are presented in the literature [3]. To address some of these difficulties our study aims at establishing an appropriate model for assessing the digitization level and its impact on education institutions including secondary schools.

* Simon Nyaga Mwendia, Email: smwendia@kca.ac.ke

2. Related Work

2.1. Digitization

Education is one of the important mechanisms that provide the necessary skills, competencies and knowledge among people. Consequently, research on factors influencing the academic performance of students in education institutions has become a topic of increasing interest. Previous studies have reported that digitization and the integration of ICT tools is one of the factors that are attributed to improvement of academic performance in secondary schools [5]. A recent report from the Education Council of the Netherlands has revealed that the education sector is still looking for the suitable level and appropriate methods of digitization [6]. This observation has created the need for establishing an appropriate index for measuring digitization level and its impact on education. Examples of existing models of indices include digitization index [7] and digitization impact index.

a) Digitization Index (DiGiX)

According to [7], Digitization Index (DiGiX) can be described as a composite index for measuring the impact of digitization in societies such as learning institutions. This is a tool assessing the extent to which ICT is integrated in institutions[8]. Previous research work has been focused on developing indices that can assess digitization in a country. An example is Digitization Index proposed by [9] as tool for assessing the level digitization. This level can be measured using six main attributes. That is, (i) Ubiquity, which refers to the level of accessing universally digital applications and services by consumers and enterprises, (ii) affordability, which is described as the degree to which digital services are valued in a range that makes them available to a large population, (iii) reliability that is viewed as the quality of available digital services, (iv) speed, which is described as the extent to the degree to which digital services can be available in real time, (v) usability that is described as the simplicity of using digital services and the capacity of local ecosystems to improve utilization of these services and, (vi) skill that is viewed as the ability of users to incorporate digital services into their daily lives and businesses. Figure 1, illustrates how each of the metrics are related in digitization index.

c) Digitization Impact Index

Digitization Impact Index can be described as the effect of digitization of a society such as a school or a country. In the context of a school, Digitization Impact index can be studied as an impact of digitization of a school in a user satisfaction model. In this case, the user is the staff or student who is directly impacted by the utilization of digital services in the school. Key indicators to compute their level of satisfaction are derived from the rating of the efficiency of services offered in the school which would be influenced by the level of digitization. According to [10] people are used to the concept of rating things with numerical scores and these can work well in surveys. Once the respondent has been given the anchors of the scale, they can readily give a number to express their level of satisfaction. Typically, scales of 1- 5 are used where the lowest figure indicates extreme dissatisfaction and the highest shows extreme satisfaction.

Several studies have observed that digitizing schools has an impact on education that can be measured using the following indicators:

i. Improvement of students' academic performance that is generally evaluated by the mark or grade attained by the student in an examination [5]
ii. Increase of enthusiasm, interest and creativity [5].
iii. Ease of system use like explaining concepts when cast on screen [5].
iv. The speed of information access increase as a result of internet infrastructure that support pervasive, simultaneous utilization of devices for instruction, evaluation, and school operations [12].
v. The number of decisions to make increases. These include device acquisition, content options, the role of technology in supporting classroom management, security and privacy, and data management [12].

However, policy makers are confronted with two main challenges that are related to application of digital index. First, there are no standard performance indicators to measure the level at which ICT is incorporated in societies. The second challenge is that there is lack of instruments for measuring the impact of adopting mass connected digital technologies and applications in societies and economies [8]. Consequently, there is lack of a standard index for assessing the impact of digitization in secondary schools.

3. Proposed Model

As an attempt to address these challenges, we propose a model that can be used to assess the level and impact of digitization in a typical secondary school. The model was derived by combining a composite digitization index proposed by [9] with indicators of digitization impact stated by [5, 12]. This is shown in Figure1.

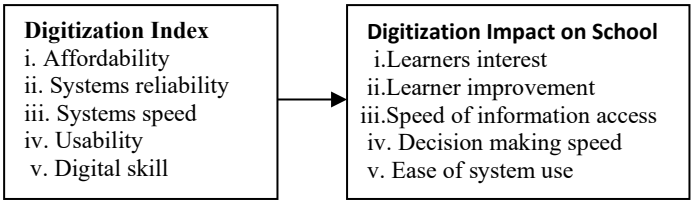


Figure1: Digitization impact conceptual model

As illustrated in Figure1, *digitization index* is representing independent variable for describing the level of digitization in a typical a secondary school. On the other hand, *digitization impact on the school* index represents the dependent variable. The proposed model can also be illustrated using simple regression equation of the form, $\mathbf{Y= a + b_1X_1}$, where Y is the *digitization impact index* on the school , a is the constant parameter, b_1 is the parameter to be estimated and variable X_1 is the *digitization index,* an independent variable.

4. Significance of the Study

There are three main benefits that are associated with the proposed model. First, it can be used by the government to formulate policies related digitization of secondary schools, particularly in the African contexts like Kenya. The second benefit is that it can be used by funding organization to demonstrate accountability when making investment decisions [3]. The third benefit is that it can be used by school managers to assess the impact of integrating ICTs in a typical secondary school.

5. Study Method

According to [13] as cited in [14], at least thirty (30) subjects required for correlational research to investigate relationship between variables and at least 10% of the accessible population is required for descriptive research. Based on this argument, the researcher collected data from 35 teachers in 11 (>50%) schools of Makadara Sub- County in Kenya.

The study sought the teachers input in assessing the reliability and speed of computer systems, adequacy of infrastructure and the impact that this had on both their students' interest and improved academic performance. The input from teachers was solicited through questionnaires and converted into numeric values by use of Likert scale.

For each of the seven (7) schools visited, individual responses obtained through questionnaires were used to generate both Digitization index and Digitization impact index. This was done by computing the percentage of arithmetic mean obtained from 5-point Likert scale. Based on the scale, the maximum rating for each indicator was five (5) and the lowest rating was one (1). Therefore, the maximum possible sum of the five identified indicators was 5X5=25. The sum of the levels was then expressed as a percentage to the maximum possible sum (25) to obtain an index for each school. For instance, if an individual respondent rated Learner interest = 3, Learner improvement =3, Speed of information access = 4, Ease of system use =3 and Decision making speed =2, then the ICT Integration index can be computed as follows:

$$X = \frac{(3+3+4+3+2)}{5X5} \times 100\ \% = \frac{15X100\%}{25} = 60$$

During data analysis, digitization index was treated as independent variable while academic performance was treated as a dependent variable. The internal consistency of the questionnaire was tested by use of Cronbach Alpha measure and found to have the acceptable value of 0.735. Correlation analysis was then carried out to assess the strength and the direction of association between variables [15] while linear regression analysis were used to develop the proposed model.

6. Results

Table1 presents shows results from the seven (7) schools, where variables X_1, represents Country's' Digitization index, and Y represents digitization impact on learners behavior.

Table1: Digitization index (X_1) and Digitization Impact index (Y)

Digitization index (X_1)	Digitization Impact index (Y)
64	44
80	80
80	76
72	92
92	60
68	72
52	52

Pearson correlation results from the seven (7) schools' sample showed that Digitization index was positively related to digitization impact with a correlation coefficient of 0.38. This is shown in Figure 2.

	Digitization_Index	Digitization_impact
Digitization_Index	1.000000	0.379779
Digitization_impact	0.379779	1.000000

Figure 2: Correlation between digitization index and digitization impact

Findings from Sklearn library that is implemented in python programming language revealed that academic slope coefficient (b_1) was approximately 0.49 while the intercept coefficient was 32.20. The two computed coefficients were then substituted in the regression equation to formulate the empirical model that is shown in Figure 3. The model means that for every 0.49 increase in the digitization index there is one unit increase in the academic slope.

$$Y = 32 + 0.49X_1$$

Figure 3: The developed empirical model

For the purpose of enhancing understanding Seaborn and Matplotlib packages were used to visualize the developed model from the seven schools sample. The results are shown as shown in Figure 4.

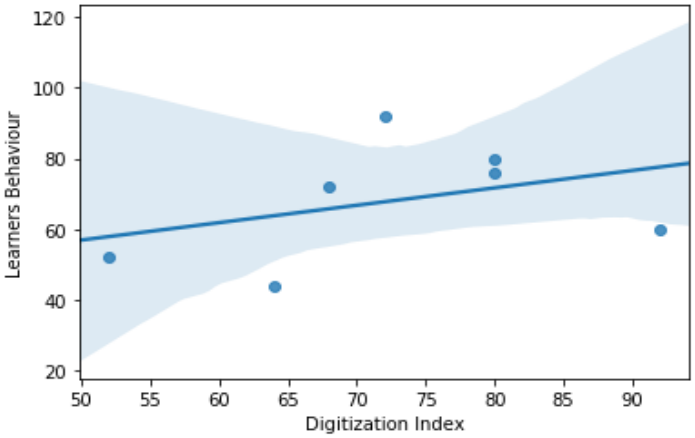


Figure 4: Visualization of the developed empirical model

7. Conclusion and Future Work

Findings from this study reveal that Digitization index of secondary schools can be measured using five indicators. That is, (i) affordability, (ii) systems reliability, (iii) systems speed, (iv) usability and (v) digital skill. In addition, linear regression equation can be used as an index for measuring the impact of digitization on both the academic performance and service delivery. Metrics for measuring such an impact include (i) Learners interest, (ii) learner improvement, (iii) speed of information access, (iv) decision making speed, and (v) ease of system use

From the reviewed literature it was observed that digitization has a positive impact on academic performance, interest and enthusiasm, ease of system use, and decision making process [5, 12]. On the other hand, the evaluation results of the proposed model showed that there is positive correlation between digitization and academic the effect on same indicators. Therefore, it can be concluded that the proposed model is consistent with findings of [5, 12].

Further research can be conducted to explore other measures for assessing digitization impact. The generalization of the proposed model can be evaluated by testing it in other education contexts like colleges and universities.

Conflict of interest

The authors declare that there is no conflict of interests in relation to the publication of this paper.

References

[1] R. Bisaso and A. Visscher, "Computerised School Information Systems Usage in an Emerging Country - Uganda," in *Information and Educational Management in the Knowledge Society*, New York,USA: Springer, 2005, pp. 81–98.

[2] M. Shah, "Impact of management information systems (MIS) on school administration: What the literature says," in *5th World Conference on Educational Sciences - WCES 2013*, Rome,Italy, 2013, pp. 2799 – 2804.

[3] A. B. Youssef and M. Dahmani, "The Impact of ICT on Student Performance in Higher Education: Direct Effects, Indirect Effects and Organisational Change," *Univ. Knowl. Soc. J. RUSC*,vol.5,no1, 2008.

[4] R. Katz, P. Koutroumpis, and F. M. Callorda, "Using a digitization index to measure the economic and social impact of digital agendas," *emeraldinsight*, vol. 6, no. 1, pp. 32–44, 2014.

[5] F. K. Kisirkoi, "Integration of ICT in Education in a Secondary School in Kenya: A Case Study," *Lit. Inf. Comput. Educ. J. LICEJ*, vol. 6, no. 2, 2015.

[6] VSNU, "Digitisation in academic education." VSNU Association of Universities,the Netherlands, 2017.

[7] N. Camara and D. Tuesta, "DiGiX: the Digitization Index." BBVA Research, 2016.

[8] R. Katz, K. Pantelis, and M. C. Fernando, "Using a digitization index to measure the economic and social impact of digital agendas," *Emeraldinsight*, vol. 16, no. 1, pp. 32–44, 2014.

[9] K. Sabbagh, B. El-Darwiche, R. Friedrich, and M. Singh, "Maximizing the impact of digitization," Booz&Company, 2012.

[10] P. Hague and N. Hague, "Customer Satisfaction Surveys & Research: How to Measure CSAT." B2B International, n.d.

[11] N. Bitner and J. Bitner, "Integrating Technology into the Classroom: Eight Keys to Success," *Jl Technol. Teach. Educ.*, vol. 10, no. 1, pp. 95–100, 2002.

[12] K. Khemani, C. Hagen, B. Ross, and A. K. Jamjourm, "Digital School." ATKearney, 2013.

[13] L. . Gay and P. . Diehl, *Research Methods for Business and Management*. Macmillan, 1992.

[14] R. Hill, "What Sample Size is 'Enough' in Internet Survey Research?," vol. 6, no. 3–4, 1998.

[15] O. Mugenda and A. Mugenda, *Research methods. Quantitative and Qualitative approaches. Nairobi.* Nairobi: Africa Centre for technology studies, 2003.

.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 198-206 (2019)

www.astesj.com

Special Issue on Recent Advances in Engineering Systems

ASTESJ

ISSN: 2415-6698

Indoor Positioning and Path Planning Platform for iRobot Create 2 Sweeping Robot

Rung-Shiang Cheng[*,1], Wei-Jun Hong[3], Jhe-Lin Li[2], Kawuu W. Lin[3]

[1]*Department of Information Technology, Overseas Chinese University, Taiwan, R.O.C*

[2]*Department of Computer and Communication, Kun Shan University, Taiwan, R.O.C*

[3]*Department of Computer Science and Information Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan, R.O.C*

ARTICLE INFO

Article history:
Received:19 December, 2018
Accepted:25 May, 2019
Online: 18 June, 2019

Keywords:
Indoor position
Path control
Sweeping robot

ABSTRACT

Many robots have been proposed for smart applications in recent years However, while these robots have good ability to move between fixed points separated by short distances, they perform less well when required to navigate complex interior spaces with an irregular layout and multiple obstacles due to their relatively crude positioning and path control capabilities. Accordingly, the present study proposes an integrated system consisting of a Raspberry Pi development platform, various sensor devices and iBeacon technology to facilitate the path control, indoor positioning, and obstacle avoidance of a programmable iRobot Create 2 sweeping robot. The capabilities of the proposed system and its various components are investigated by means of practical experiments and numerical simulations. In general, the results confirm that the proposed integrated system provides a viable platform for the future development of sophisticated indoor robots for smart indoor applications.

1. Introduction

In recent years, robot technology has been introduced into related application of smart network. However, interior space is a very complicated environment because signal transmission can be easily influenced by architectural structure. Thus, for many years, well-known services to the mess such as GPS positioning system or Google map are most applied in outdoor public infrastructure instead of interior space. This has blocked the application of robots in smart homes; at present, smart robots mostly move between fixed points, on certain tracks or only in short distance; for the movement in interior space of robots, there have not proper solutions because researches in this field is still in startup stage and robots are in lack of the support in IoT environment; plus, problems emerged when it comes to the sensing technology, spatial orientation, path planning, obstacles avoidance, robot learning and communication between robots still need to be solved and overcome by further integration.

To introduce the application of robots from factories to homes, we need to enable smart robots to arrive the right position rapidly through the indoor application service or precisely identify robots or users via indoor positioning navigation. Thus, it is necessary to conduct further research on smart robots to develop a positioning service system used indoor. In this way, smart robots can be extended from industrial application in factories to caring service in smart homes, bringing more diverse possibilities to home application service of IoT.

This study is organized as below: Section 2 introduces related work on indoor positioning, path planning and searching, avoiding obstacles and automatic following; Section 3 introduces research steps and system implementation; Section 4 represents conclusions.

2. Research on related application of home robots

How to let robots assist people to deal with daily tasks in various scenarios and precisely identify their location and arrive target areas to complete tasks will be important goals to introduce robots to smart home application. The following are some related study of home robots.

2.1. Indoor positioning

In view of this, [1] proposed to utilize RSSI signal of wireless network and triangulation algorithm as be indoor positioning

[*]Corresponding Author: Rung-Shiang Cheng, rscheng@ocu.edu.tw

mechanism for robots. The results showed the positioning precision is 0.125 m and the error is below 0.4%.

In [2], the author proposed to utilize laser distance meter to construct indoor environment model. The meter can measure its moving path and calculate via the assistance of cloud server, effectively avoiding obstacles on moving path. In [3], it proposed to utilized the information change in the sensor or the robot and calculate its moving direction and location. Through change values measured by gyro and accelerometer, the robot can calculate its X, Y and θ. After calibration, it can obtain the reference coordinate of robots and plan an effective moving path.

2.2. Path planning

Coverage path planning [4] algorithm for robots is an important efficiency indicator for home robots; random path planning is an intuitive way to move of robots. Robots forward linearly towards any direction and change path only when hitting on obstacles and repeat this process until it reaches maximum coverage scope. Such algorithms have the advantage of simplicity since very few sensors are required and there is no need to construct and maintain an indoor environment map [5]. However, they require a long time to achieve full coverage of the target environment; DmaxCoverage [6] an algorithm which can assist robots to construct environment map to reach maximum coverage scope; furthermore, in [7], this study utilized Dijkstra's algorithm to generate optimum path in environment to avoid the decrease of efficiency during random path planning.

2.3. Search, obstacle avoidance and automatic follow-up

Unlike robots, humans can use their own sense organs to create space and sense of distance. Therefore, in [8] this paper implements a mobile robot with an automatic following system. The mobile robot can perform positioning, mapping, and moving the target side. Tasks such as distance, human detection and target tracking, and can be applied in real life, such as museum, office or library guidance; in [9], this paper implements a wheeled robot that can perform search and obstacle avoidance work. The robot is equipped with a vision and distance sensor, which can match multiple objects by visual system identification objects to create multiple hypothesis blocks for the environment. These multiple hypothesis blocks can make the robot avoid obstacles and explore the target object.

3. System design and implementation

Figure 1 shows the functional block diagram of the implemented robot. The function of positioning is a crucial factor to introduce robots into interior application. To achieve this goal, this study combines Raspberry Pi with BLE Bluetooth communication technology to implement iBeacon of robots; when the robot moves, if it detects iBeacon and reports the positioning information, the system is able to know the location of the robot.

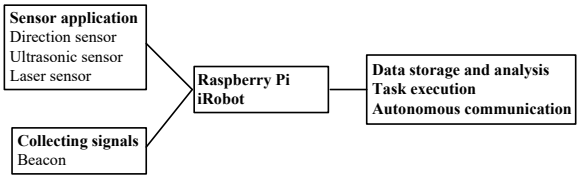


Figure 1: System block diagram of the sweeping robot

Figure 2 shows the original look of programmable sweeping robot. This study combines iRobot Create 2 programmable sweeping robot with Raspberry Pi programmable control electric circuit with I/O device to explore straight movement and rotation of robots. We also use Python to write tests of every control application. developing a smart robot which is equipped with indoor positioning function; Figure 3 is the prototype of the robot for indoor application.

Figure 2: Programmable Sweeping Robot (Original Look)

Figure 3: Smart Robot Developed by Combining iRobot Create 2 Programmable Sweeping Robot with Raspberry Pi

3.1. Composite sensing device

Linear walking and rotation are the basic movement instructions of the robot. In order to understand the difference between the expected result and the actual result after the instruction is executed, this study installed sensing devices such as distance, magnetic field, acceleration, and gyroscope on the cleaning robot. This study first measure the robot's motion data and then use the analysis results to correct the robot's motion parameters.

Show as Figure 4, take the orientation sensor as an example. After the robot adds the orientation sensor, the system developer can use the sensor's numerical change to determine the robot's movement state; because the X axis represents the robot's direction of travel, the robot can use it to determine the robot when performing a task. The forward direction; on the other hand, since the Z axis of the accelerometer is negative when the robot's operation panel is upward and downward when it is positive, it is possible to use the Z axis of the acceleration sensor to detect whether the robot is running when the robot is operated. Overturned or tilted state.

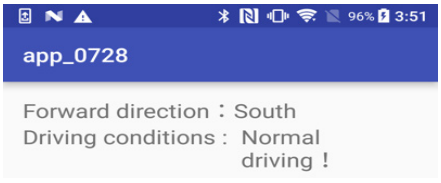


Figure 4: The result of the smart robot orientation and the tilt state of the car body

3.2. Robot moving path planning mechanism

To enhance the accuracy of control command, we need to measure reaction delay time and error when executing assigned process after robots receive commands.

3.2.1. Measure and Observe the Straight Movement

To enhance the accuracy of control command, we need to measure reaction delay time and error when executing assigned process after robots receive commands.

Assume the program execution time is 10.05 seconds to 11.00 seconds, and forwarding speed of the robot is set as 200 mm/sec; Table 1 shows the measured moving distance of the robot at different second.

As shown in Table 1, the experimental results indicate after the control device gives commands to the robot, the actual complete time of robots executing commands will have delay for about 0.5155 second (time error) and the average value of distance error is 10.1662 cm. This result can be used by the system to modify parameters so as to enhance the accuracy of positioning of the robot.

Table 1: Computed Measured Results of Experimented Robot Move Straightly for 11 Seconds

Execution Time	**Time Error**	**Computed Distance**	**Actual Distance**	**Distance Error**
10.5657	0.5157	211.3148	200.0	11.3148
10.6153	0.5153	212.3066	201.9	10.4066
10.7120	0.5120	214.2402	203.5	10.74024
10.8067	0.5067	216.1356	205.9	10.2356
10.9189	0.5189	218.3784	208.0	10.3784
11.0167	0.5167	220.3347	210.0	10.3347
11.1192	0.5192	222.3845	212.0	10.3845
11.2163	0.5163	224.3278	214.3	10.0278
11.3168	0.5168	226.3371	217.0	9.3371
11.4165	0.5165	228.3313	219.0	9.3313
11.5168	0.5168	230.3377	221.0	9.3377
	0.5155			**10.1662**

According to the above measurement it is found when speed is fixed, we can obtain stable second difference and stable distance difference. Thus, when controlling the moving distance of robots, we can do corrections by substituting values into the Eq.(1):

$$estimated\ time = \left(\frac{distance+difference}{speed\times 0.1}\right) - offset\ time \quad (1)$$

Amongst, distance indicates predetermined moving distance, difference indicates error correction distance, offset time indicates delay time of actual execution; from measured results in Table 2 it is known: let difference = 10.1662, offset time = 0.5155, assume predetermined straight moving distance of the robot is 800 cm and then the values adding into the Eq.(2):

$$estimated\ time = \left(\frac{80+10.1662}{200\times 0.1}\right) - 0.5155 = 39.9928 \quad (2)$$

Through this equation we obtain predetermined execution time of the program is 39.9928 seconds; after the program is execute, we obtain the actual execution time of the robot 40.5387 seconds, moving distance is 799 cm; the measured results show after adding it to the correction equation, the actual moving distance of the robot 799 cm is close to the predetermined value 800 cm, and difference is about 0.125%.

Table 2: Measured Results after Adding the Straight Walk to the Correction Equation

Actual Seconds	Second Difference	Actual Distance
40.5387	0.5459	799

3.2.2. Measure and Observe the Rotation Angle

To observe whether the rotation angle and direction of the robot are as what we have predicted, we combine the mobile sensor with the robot. Through function of iPhone DeviceMotion we obtain the measurement data of the robot; experimental data here will be compared with that obtained from SenseHat compound sensor; Table 3 shows the measured results by observing rotation angle per second after the robot finishes straight move and another move for 1.1~3.3 seconds at 100 mm/sec rotation speed. The testing process is shown in Figure 5.

(a) The Status of Actual Measurement

Device Motion		Magnetometer	
Pitch:	1	X:	26.18
Roll:	-2	Y:	-124.12
Yaw:	-35	Z:	-43.10
Accelemeter		**Gyroscope**	
X:	-0.006	X:	-0.049
Y:	-0.035	Y:	-0.085
Z:	-0.970	Z:	-0.029

(b) Measured Results

Figure 5: Measure Rotation Angles of Robots

From the experimental results we know when rotation speed is fixed, dividing actual angle by program seconds will obtain rotation angle per second, rotation angle per second is stable and thus we obtain the equation below:

$$excution\ time = expected \div average \quad (3)$$

Let expected denote the predetermined rotation angle, average denote the actual angle/ obtained average rotation angle according to program execution seconds (control factor) According to the measured results in Table 3, let average = 55.21061296□ /sec; during the execution, we simply need to substituting the degree of

angle to rotate into the above-mentioned equation to obtain the required seconds for rotating to the assigned angle; taking rotating 90-degree angle as example, through the equation we obtain the execution time of the program will be expressed as execution time = 90 / 55.21061296 = 1.6301 seconds.

Table 3:: Actual Measured Results of Rotation Angles of Robots

Execution Time in Seconds	Rotation Angle	Rotation Angle Per Second
1.502199	55	55
1.602238	61	55.45454545
1.702244	67	55.83333333
1.802489	72	55.38461538
1.902531	77	55
2.002577	83	55.33333333
2.102841	88	55
2.202954	93	54.70588235
2.302842	100	55.55555556
2.403040	104	54.73684211
2.503230	109	54.5
2.603253	117	55.71428571
2.703390	122	55.45454545
2.803547	126	54.7826087
2.903717	132	55
3.003546	139	55.6
3.103666	142	54.61538462
3.203762	149	55.18518519
3.303927	153	54.64285714
3.404050	161	55.51724138
3.504325	165	55
3.604233	171	55.16129032
3.704476	176	55
		55.21061296

Table 4 shows the results obtained when the robot rotated 90 degrees, 180 degrees, 270 degrees, and 360 degrees. It can be seen from the experimental results that with the increase of the program execution time, the errors generated by the robot after the execution of the rotation action will gradually appear.

Table 4: Actual results of the angles obtained after actual implementation

Estimated rotation angle	Calculated time	Actual execution time	Actual rotation angle	Angle error
90	1.6301	2.132857	91	1
		2.133154	90	0
		2.132777	91	1
		2.133019	87	3
		2.132970	91	1
		2.132866	90	0

(a) 90 degrees

Estimated rotation angle	Calculated time	Actual execution time	Actual rotation angle	Angle error
180	3.2602	3.762850	180	0
		3.764159	183	3
		3.762853	180	0
		3.763769	181	1
		3.764353	181	1
		3.764293	180	0

(b) 180 degrees

Estimated rotation angle	Calculated time	Actual execution time	Actual rotation angle	Angle error
270	4.8904	5.396417	269	1
		5.396603	272	2
		5.396394	268	2
		5.396588	268	2
		5.396419	271	1
		5.396533	270	0

(c) 270 degrees

Estimated rotation angle	Calculated time	Actual execution time	Actual rotation angle	Angle error
360	6.5205	7.028154	357	3
		7.028164	362	2
		7.028282	362	2
		7.028401	362	2
		7.028339	362	2
		7.028357	362	2

(d) 360 degrees

SenseHAT is an expansion board for the Raspberry Pi. As shown in Figure 6, since the robot's magnetic field will affect the SenseHAT's value, in order to improve the accuracy, SenseHAT will calibrate the magnetic field, and then measure the SenseHAT detection value is correct. Table 5 uses the rotation angle of 72 degrees as an example to judge the difference between the predicted angle and the actual angle. The maximum value of the experimental result drops to 5, so it needs to be corrected after each rotation command

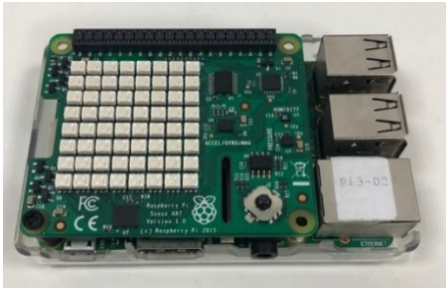


Figure 6: SenseHAT composite sensor

3.2.3. *Measure and Observe the Rotation Angle*

As mentioned above, the movement of robots is consisted of two parts including straight movement and rotation. Thus, we can write its moving path into script and then through changing

commands to make robots execute tasks. Let s denote straight moving, t denote back moving, r denote right turn and l denote left turn. Taking Figure 7 as example, the script of the moving path of the robot will be written into "s290 r90 s47 l90 s310 r30 s122." if we break down the script by rules, it can be divided into seven command list including forwarding 290 cm, rotating right for 90-degree angle, forwarding 47 cm, rotating left for 90-degree angle, forwarding 310 cm, rotating right for 30-degree angle, forwarding 122 cm and arriving the destination.

Table 5: Experiment with a rotation of 72 degrees

Initial angle	Estimated rotation angle	Actual rotation angle	Angle error
105	177	180	3
	249	251	2
	321	323	2
	33	38	5
	105	106	1
	178	181	3
	250	253	2
	322	323	1
	34	38	4
	106	106	0

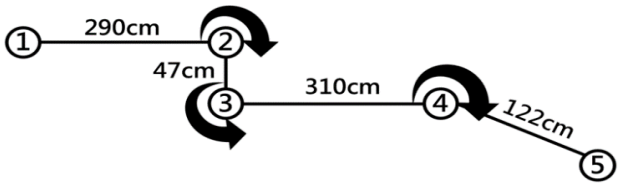


Figure 7: Moving Path Planning of Robots

3.2.4. *Production of script tools*

Although the path of the robot can use the previously written script to write the relevant parameters to the robot. However, in order to make the robot's control more flexible, the project also developed a Web control interface that allows the user to control the walking path of the robot through the Web, as shown in Figure 8:

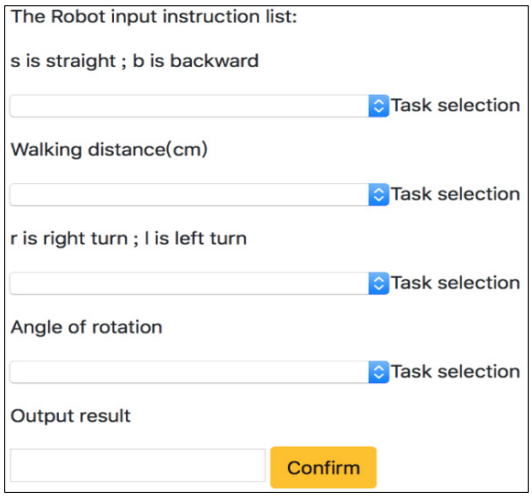


Figure 8: Manual specification of robot path

When the user issues an instruction, the system will put each value into the database and then convert it to json format output when needed, as shown in Figure 9

id	script
91	t200 r180
92	s50 l90
93	t100 r10
94	s50 l90
95	t10 r10
96	s50 l90
97	s50 l90
98	t100 r180
99	s10 l10
100	t50 r180
101	s200 r180
102	s100 l10
103	s50 r90

Figure 9: Store forward distance and turning angle

As shown in Figure 10, when robots execute Script, it would obtain forwarding distance and rotation angles from the database to execute tasks. When all assigned tasks are done, it would show the spent time for finishing all movements to notify users of time spent to complete cleaning so that users can more easily use it.

```
Today: Fri Feb  2 16:53:24 2018 16:53:24.457
Mission: t10 r10
Default speed 200mm/s

 | --- Start ---
 | - Straight walk -10 cm
 | - 0.494894 sec
 | - Turn 10 degrees
 | - 0.183419 sec
 | --- End ---

 | - It cost 0.747108 sec
Today: Fri Feb  2 16:53:25 2018 16:53:25.204
```

Figure 10: Executed Motions and Spent Time of Robots

3.3. *The Design of iBeacon Information Assist System*

Owing to smart robots can provide more human-oriented location-based service after combining with perception positioning system and have better market competitiveness in the future, this study new robot communication application service in view of the development of smart home and IoT. This study implements diverse simulation test by iRobot Create2 which is designed to control robots, applying iBeacon micro-location technology, compound multi-axis sensor technology and so on to develop smart robots for indoor positioning and path algorithm; for robots, the precision of positioning signals is a very crucial parameter.

3.3.1. *Experimental field planning and design*

To enable robots to obtain indoor coordinates and execute dynamic searching, this study constructs an IoT micro-location

system. Through the assistance of the iBeacon system, robots are able to inquire past moving track before executing tasks and transform the moving process to coordinate record to build up the spatial perception ability of robots; when robots calibrate signals and connect to system database. The robot can report its location to the system while forwarding and meanwhile report its location to the IoT micro-location system so that it can plan, correct and predict its paths.

Fig. 11 shows the experimental field. The experimental field is an 85m x 20m 2-dimensional space. A total of 14 Beacons are arranged to allow the robot to perform tasks that can be used for positioning applications.

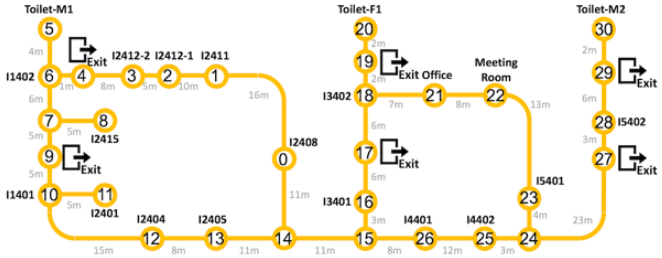


Figure 11: Experimental space for robotic indoor positioning applications

3.3.2. *Researches on Beacon Positioning*

The iBeacon messages detected by the robot contained various items of location-related information, including the UUID, the Tx Power, and so on (see Fig. 12)

```
pi@raspberrypi:~/iBeacon-Scanner- $ sudo python testblescan.py
ble thread started
----------
28:1d:2a:c4:ba:ac,5e5f9c2c626deea1145a5321c057319e,23150,1036,89,-71
4e:ab:8b:77:cd:49,0613ff4c000c0e004d219b3bf2d2bab0,21997,47890,-77,-49
f4:5c:89:c4:27:98,0100009827c4895cf40b02010607ff4c,16,523,0,-64
c8:69:cd:6c:34:64,01000064346ccd69c80b02011a07ff4c,16,513,0,-49
17:42:ec:af:bc:f8,bcafec42170f02011b0bff4c00090603,632,29288,32,-49
28:1d:2a:c4:ba:ac,5e5f9c2c626deea1145a5321c057319e,23150,1036,89,-62
4e:ab:8b:77:cd:49,0613ff4c000c0e004d219b3bf2d2bab0,21997,47890,-77,-51
f4:5c:89:c4:27:98,0100009827c4895cf40b02010607ff4c,16,523,0,-49
17:42:ec:af:bc:f8,bcafec42170f02011b0bff4c00090603,632,29288,32,-63
----------
```

Figure 12: Contents of iBeacon message

After obtaining RSSI and TxPower value between robots and Beacon, we can then use relative relation between RSSI and TxPower as the reference value of distance between the robot and Beacon; when ratio between RSSI and TxPower is less than 1.0, we use Eq. (3) to obtain distance, or we will use Eq. (4):

$$distance = \begin{cases} \left(\frac{RSSI}{TxPower}\right)^{10} & ,if\ \frac{RSSI}{TxPower} < 1.0 \quad (3) \\ \left(\frac{RSSI}{TxPower}\right)^{7.7095} \times 0.89976 + 0.111 & ,else\ (4) \end{cases}$$

Scan the iBeacon through the Raspberry Pi and calculate the distance, as shown in Figure 13, but since this method is easy to generate large fluctuations through real-time operation, it is necessary to increase the stability of the distance by accumulating balance, and plan two points iBeacon to be placed at 0.5m and 2.5m, 30 points are measured for each point, and the first 10 data are averaged, and then the last 20 data are corrected by the EWMA formula to obtain the distance. Figure 14 is the experimental results.

```
BLE thread started
MAJOR : 32 , MINOR : 14
TxPower : -59 , RSSI : -49
Ratio : 0.830508
Distance : 0.156114
--------*--------
pi@raspberrypi:~/iBeacon-Scanner- $
```

Figure 13: Raspberry Pi Scan iBeacon and Calculate Distance

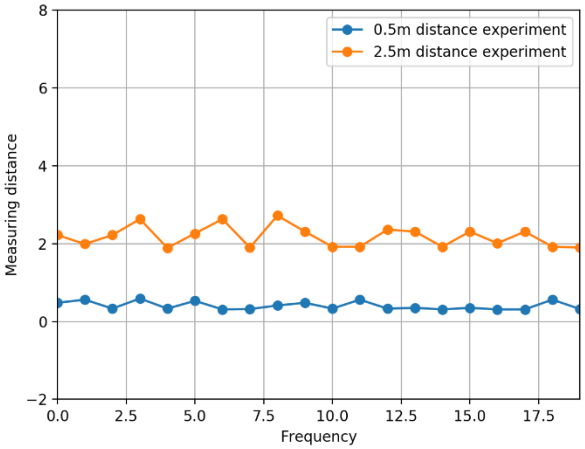


Figure 14: Revise the distance of iBeacon

3.4. *Robot positioning and map construction*

Safe walking is a challenge when robots enter an unknown environment. SLAM technology can be used to map assisted positioning by reading the value of the surrounding environment by the sensor.

3.4.1. *Robot Operating System*

With the warming of domestic robots in the past two years, autonomous positioning has become the primary threshold for robotic intelligence. To realize this function, many manufacturers have chosen to use ROS as the operating system of robots.

ROS is more like an application role as a communication intermediary software in the entire robot design architecture. At the same time, a complete system integration framework for robot application design based on an existing operating system can be developed in various programming languages such as C++ and Python. Through the functions provided by ROS, the use of SLAM can be made more convenient. Figure 15 shows the use of ROS to generate a simple picture through the virtual machine. It is necessary to constantly correct the map through the SLAM algorithm to make the map more accurate.

3.4.2. *SLAM map construction*

As shown in Fig. 16, SLAM comprises three components, namely localization, mapping and navigation. With advances in computer technology and sensing capabilities, SLAM has evolved into the method of choice for robots to detect unfamiliar environments and find their way through them. As shown in Fig. 17, various SLAM technologies are available, including VSLAM,

Wifi-SLAM and Rplidar SLAM. Among these technologies, Rplidar SLAM is one of the most commonly used and is hence also adopted in the present study.

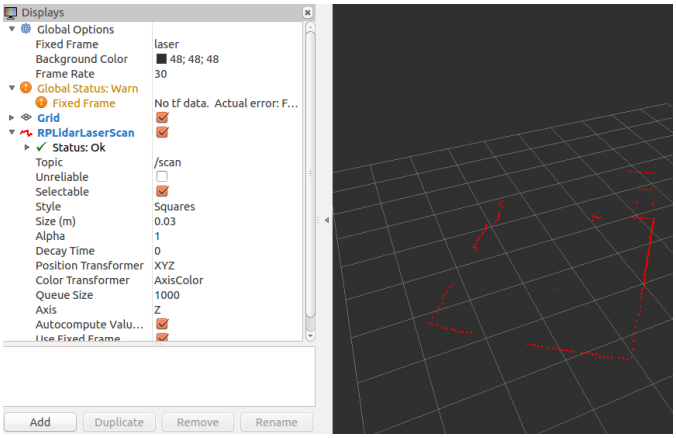


Figure 15: A Depicting map outlines through ROS

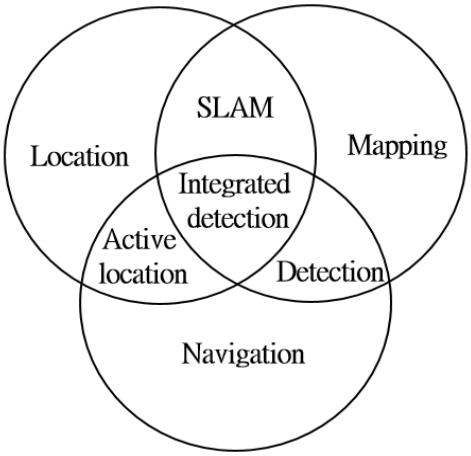


Figure 16: SLAM application diagram

VSLAM WiFi-SLAM Rplidar SLAM

Figure 17: Common SLAM technology

Rplidar Collect data repeatedly, filter noise and produce visual map, Rplidar A2 Typical rotational frequency is 10Hz (600RPM), can realize 0.9° angle resolution at typical rotational frequency, can adjust the rotation frequency in the range of 5Hz-15Hz with the user's demand.

3.4.3. *RPLIDAR Scan Value*

As shown in Figure 18, RPLIDAR is performed by Python to perform scanning, and the values of four angles are obtained, which are 0 degrees, 90 degrees, 180 degrees, and 270 degrees, and the results are shown in Table 6, can be applied to robot detection and The distance between obstacles.

Figure 18: Use Rplidar to get the value of the angle

Table 6: 4 angle values

Angle	Distance (cm)
0	17.6
90	103.3
180	13.4
270	20.0

3.4.4. *SLAM Algorithm*

In the SLAM algorithm, the Hector SLAM algorithm is used. The algorithm developed by Hector Labs in Germany was ported from the original RoboFram algorithm to the ROS package at the end of 2010. Using the 2D mesh method, the value of the laser scan is converted in the xy coordinates by the tf library, and possible position estimation and rendering obstacles. Repeated scanning uses a bilinear filtering method to continuously correct the map with low power consumption, immediacy and goodness. Figure 19 shows the working principle of the Hector SLM, and Figure 20 shows the experimental results of the Hector SLAM scanning virtual machine.

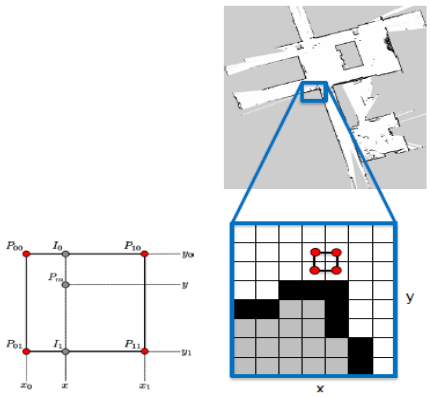


Figure 19: Hector SLAM works

3.5. *Robot Obstacle Avoidance*

To carry out research on robot obstacle avoidance, the robot is equipped with three ultrasonic sensors. As shown in Figure 21, the ultrasonic sensor can sense a distance of 2cm to 400cm, an accuracy of 0.3cm, and an angle range of 15 degrees. The robot can read the value of the sensor to determine whether it encounters an obstacle, such as Figure 22 shows.

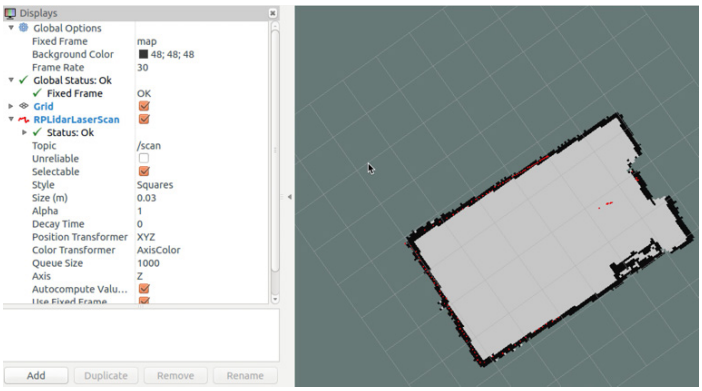


Figure 20: Logo

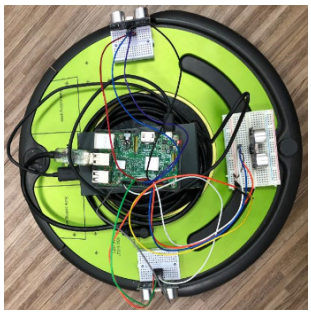


Figure 21: Install an ultrasonic sensor on the robot

```
pi@raspberrypi:~/Create2Control $ sudo python sensor_test.py
SensorNum:1 init
SensorNum:2 init
SensorNum:3 init
hinder:{'Front': 118.1, 'Right': 11.1, 'Left': 23.2}
hinder:{'Front': 118.1, 'Right': 11.1, 'Left': 23.2}
hinder:{'Front': 118.1, 'Right': 11.1, 'Left': 23.2}
```

Figure 22: Confirm that the ultrasonic sensor is working

The robot's ultrasonic sensor determines if an obstacle is encountered in front, and the robot will determine which side is longer by the left and right sensors to determine how to avoid the obstacle. As shown in Fig. 23, when the distance is less than 20 cm, the obstacle avoidance behavior is performed. When the sensor's value returns to normal, the robot will travel midway through the body to ensure that the robot completely bypasses the obstacle and rotates back to the original path.

```
hinder:{'Front': 22.4, 'Right': 115.1, 'Left': 81.9}
Car State: Stop
Car State: Turn Right
```

Figure 23: Turning with obstacles less than 20cm

To understand the execution situation of the obstacle avoidance action of the robot, Figure 24 and Figure 25 simulate and observe the execution process of the obstacle encountered by the robot, where the x-axis is time and the y-axis is the path of the robot to avoid obstacles. When the robot turns right the y axis is decremented by 1 , the y axis is incremented by 1 when it is turned left. Figure 24 simulates the robot encountering an obstacle at 2 seconds, then turning right around the obstacle and leaving the obstacle at 4 seconds. Figure 25 simulates the robot encountering an obstacle at 3 seconds, then turns left around the obstacle and leaves the obstacle at 5 seconds.

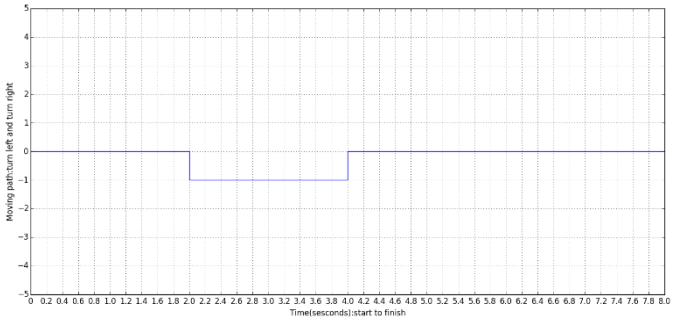


Figure 24: Simulate turn right bypass the obstacle

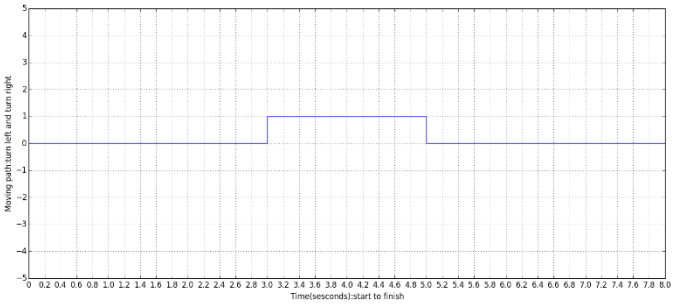


Figure 25: Simulate turn left bypass the obstacle

Let the robot perform the task of bypassing obstacles. From the experimental results, the robot encountered an obstacle within 2.4 seconds, began to perform the action of avoiding the obstacle, and left the obstacle within 3.8 seconds. Figure 26 is the result of the actual walking of the robot. The obtained time is brought into the program, and the result chart of the simulation is shown in Fig. 27, and the actual obstacle avoidance result chart is consistent.

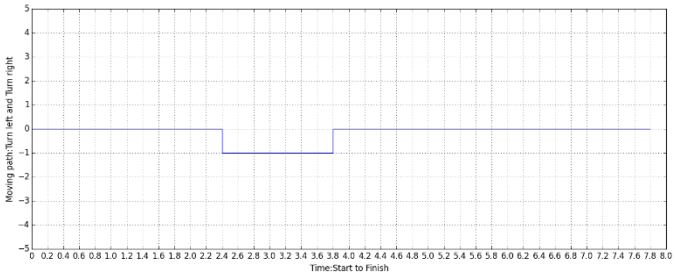


Figure 26: The actual obstacle avoidance result of the robot

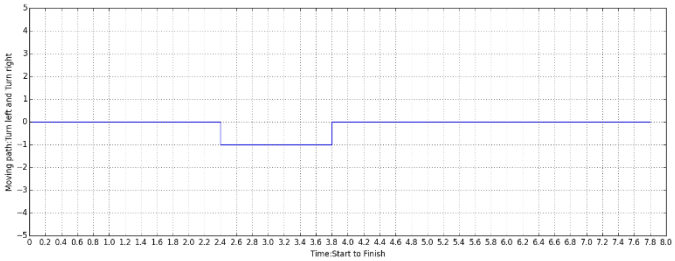


Figure 27: Simulate the results of actual obstacle avoidance

4. Conclusion

This study has developed an integrated development platform for indoor mobile robots consisting of a programmable iRobot Create 2 sweeping robot; a Raspberry Pi computer; a sensing system consisting of laser, magneto, accelerometer, gyroscope and

ultrasonic devices; and iBeacon technology. The path planning and positioning performance of the proposed platform has been evaluated experimentally and calibration equations have been proposed for both the linear and rotational motions of the robot. The use of SLAM technology to construct a map of the surrounding environment has been explained. Finally, the ability of the sweeping robot to perform automatic obstacle avoidance maneuvers has been demonstrated numerically and experimentally. In general, the results confirm the basic feasibility of the proposed framework and provide a useful basis for the future development of mobile robots for indoor smart network applications.

References

[1] Wengpeng Wang, Xin Wang "Research on robot indoor localization method based on wireless sensor network", 2015 International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2015), May 2015.

[2] Md. Masum Billah, Md. Raisuddin Khan "Smart Inertial Sensor-based Navigation System for Flexible Snake Robot", IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), pp.1-5, Nov. 2014.

[3] Nieves Pavón-Pulido, Juan Antonio López-Riquelme, Juan Jesús Pinuaga-Cascales, Joaquín Ferruz-Melero, Raul Morais dos Santos "Cybi: A Smart Companion Robot for Elderly People: Improving Teleoperation and Telepresence Skills by Combining Cloud Computing Technologies and Fuzzy Logic ", EEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), pp.198-203, April 2015.

[4] Abhishek Pandey, Anirudh Kaushik, Amit Kumar Jha,Girish Kapse,"A Technological Survey on Autonomous Home Cleaning Robots", International Journal of Scientific and Research Publications, Vol. 4, No. 4, April 2014

[5] Ercan U. Acar, Howie Choset, Yangang Zhang and Mark Schervish, "Path Planning for Robotic Demining: Robust Sensor-Based Coverage of Unstructured Environments and Probabilistic Methods", The International Journal of Robotics Research, July 2003.

[6] Feng-Min Chang and Feng-Li Lian, "Inverse observation model and multiple hypothesis tracking for indoor mobile robots", 2014 IEEE International Conference on Automation Science and Engineering (CASE), pp. 1200-1205, 18-22 Aug. 2014.

[7] Zeyn Saigol, Bram Ridder, Minlue Wang, Richard Dearden, Maria Fox, Nick Hawes, David M Lane, and Derek Long, "Efficient Search for Known Objects in Unknown Environments Using Autonomous Indoor Robots", In Workshop on Task Planning for Intelligent Robots in Service and Manufacturing, IROS 2015. October 2015.

[8] Feng-Min Chang and Feng-Li Lian, "Inverse observation model and multiple hypothesis tracking for indoor mobile robots", 2014 IEEE International Conference on Automation Science and Engineering (CASE), pp. 1200-1205, 18-22 Aug. 2014.

[9] Zeyn Saigol, Bram Ridder, Minlue Wang, Richard Dearden, Maria Fox, Nick Hawes, David M Lane, and Derek Long, "Efficient Search for Known Objects in Unknown Environments Using Autonomous Indoor Robots", In Workshop on Task Planning for Intelligent Robots in Service and Manufacturing, IROS 2015. October 2015.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 207-219 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Study of Performance of Bio- Inspired Strategies Applied to Pursuit Evasion Game Under Feedback Laws

Lairenjam Obiroy Singh[1,*], Rajagopalan Devanathan[2]

[1]*Hindustan Institute of Technology and Science, Department of Electronic and Communication Engineering, 603103, India*

[2]*Hindustan Institute of Technology and Science, Department of Electrical and Electronics Engineering, 603103, India*

ARTICLE INFO

Article history:
Received:31 March, 2019
Accepted:07 June, 2019
Online: 18 June, 2019

Keywords:
Bio- Inspired
PEG
Feedback Laws
Closed Loop Control

ABSTRACT

Pursuit Evasion Game (PEG) is an abstract model of various significant problems that appear in both civil and military applications. Bio- Inspired strategies are found to be very useful in studying the PEG. While optimal response to the pursuit strategies are available using geometric control theory, it is shown in this paper that application of linear feedback control laws can further improve the time and tracking response of these strategies in capturing the evader by the pursuer. Empirical results based on computer simulation are used to illustrate the findings. Further, considering the case of sudden turn of the evader, moving at a lower speed, it is shown that both in theory and simulation that the evader can delay the capture by pursuer and in some cases even escape from being captured. These findings are in line with what is found in nature.

1. Introduction

Pursuit evasion game (PEG) is widely prevalent in nature. It is a game that is considered to be between the pursuer and an evader. PEG can be observed among animals when they chase a victim or when they battle for territory and even when they mate. In the context of engineering, PEG finds application in missile guidance and avoidance, aircraft pursuit and evasion, maritime asset protection etc.[1].

Initial study of PEG was mainly from the point of view of game theory [2]. PEG was further studied from the point of view of geometric control theory in [3]. Pursuit manifold is defined in terms of certain criteria involving relative distance and relative velocity between the pursuer and the evader. Using a control law which enables the pursuer to reach the pursuit manifold in an optimum way and keeping the pursuit invariant on the manifold, [3] studied three strategies found in nature, viz., motion camouflage (CM), constant bearing (CB) and classical pursuit (CP). The objective of the work reported in [3] is to find a justification for the prevalent use of CM in PEG in nature. It turned out that, keeping the speed of the pursuer and the evader constant and ensuring that the pursuer moves faster than the evader, [3] has been able to show through evolutionary games that CB and CP strategies always converge to CM strategy under certain fixed assumptions on evader motion thus supporting what is seen in nature.

However, Pais [4] has argued that if the evader control law is reactive (i.e. it is in some way dependent on the baseline vector joining the pursuer and evader locations and its rate of change), then the conclusions in [3] might not hold good. Wei [5] has dropped the constant speed assumption of both the pursuer and evader and allowed the pursuer to change its acceleration as required. Wei also introduced the case of victim turning suddenly left or right when the pursuer comes too close to the evader. A related problem to PEG of confinement and escape is addressed in [6].

Our aim in this paper is this. Considering PEG of a robotic pursuit of an enemy agent, can we consider the problem as a feedback control problem introducing proportional (P),proportional plus integral (PI) and proportional plus integral plus derivative (PID) laws to improve the performance of the pursuer? The performance can be studied under different bio-inspired pursuit strategies, such as CM, CB and CP. The evader may be allowed to follow reactive and non- reactive control laws. A simpler version of this study for the CM case only has been reported earlier in [7]. Further, in view of the increased agility of the evader (i.e. ability to turn at a lower speed compared to the pursuer), we ask what is the outcome of the PEG given the chance

*Corresponding Author: Lairenjam Obiroy Singh, obi3925@gmail.com

that evader makes a sudden 90 degree turn left or right? This paper is a much expanded version of [8] especially in regard to the sudden turn strategy of the evader.

Bopardikar [9] has shown the Bio- inspired co-operative strategies of pursuer for successful confinement of evader. [9] gives the required number of pursuers and speed ratio for guaranteed confinement strategy. X Liang and Y Xiao [10] show the coalition formation of robots for intrusion detection by using the game theory approach. The paper is based on the nature's coalition formation such as predator trying to catch the prey. An analytical method is used to study the tradeoff of coalition or collaboration between the robots.

Initial study of camouflage strategies [11,12] are considered when either pursuer or evader is stationary. Srinivasan [13] analyzed and investigated the motion camouflage and highlighted many ways for future extension. The authors of [14] study the unmanned surface vehicles (USVs) for performing patrolling operations and detecting intruders for harbor protection with a view to reducing the number of humans exposed to threat. An intelligent swarm management unit (SMU) is used for the supervision of all the USVs in operation. A real time motion planner for the USVs in the presence of multiple obstacles is also presented.

[15] addresses the problem associated with the classical pursuit evasion games and mentions that their study is more difficult than the classical one. The author of [15] also explored the geometry of the problem by obtaining sufficient conditions for both pursuer and evader to win. A max- min problem is formulated, from the pursuers point of view, which is solved using outer approximation method. The solution of the max- min formulation is used to synthesize a feedback solution governing the pursuer's behavior in the form of receding horizon control. [16] discusses a harbor defense situation and the author highlights the important features where the problem is different from the classical PEG.

The main contribution of the paper can be summarized as follows: (i) PI and PID feedback laws are shown to enable the pursuer to capture the evader in a shorter time compared to P alone for different conditions of the evader following non-reactive and reactive control laws. (ii) The PI, PID feedback laws followed by the pursuer enable him to capture the evader and not allow him to escape contrary to the the case of P alone. This is termed tracking performance of the pursuer (iii) In the case the evader uses his agility to turn 90 degree left or right suddenly with respect to the baseline, it is shown that the time to capture the evader can be delayed when the pursuer follows any of the CM, CB and CP strategies. (iv) It is also shown that in certain cases of CB followed by pursuer, with the evader following the sudden turn strategy, the evader is able to totally escape from the pursuer inspite of the speed advantage enjoyed by the pursuer.

The rest of the paper is organized as follows: The following section gives the required background on modeling of PEG and the pursuit manifolds under different strategies. Section 3 presents the feedback control system configuration considered together with the derivation of expressions for PI and PID for the three strategies. Also provided in section 3 is an analysis of the effect of sudden turn of the evader in respect of the CM, CB and CP strategies of the pursuer. Section 4 gives the simulation results together with the discussion on the comparative performance of the different control strategies as well as different control modes of the evader. Section 5 concludes the paper.

2. Background

Wei and Krishnaprasad [17], [18] model the interaction in pursuit in terms of gyroscopically interacting particles. We follow a similar approach.

$$\dot{r}_p = x_p\,,\ \dot{x}_p = y_p u_{pr}\,,\ \dot{y}_p = -x_p u_{pr} \qquad (1)$$

where $r_p \in \mathbb{R}^2$is the position of the pursuer, x_p its velocity and y_p is the acceleration of the pursuer. The steering control of the pursuer is given by the scalar u_{pr}. The motion of the evader (with speed v) is given by

$$\dot{r}_e = \upsilon x_e\,,\ \dot{x}_e = \upsilon y_e u_e\,,\ \dot{y}_e = -\upsilon x_e u_e \qquad (2)$$

where $r_e \in \mathbb{R}^2$ is the position, x_e is the velocity and y_e is the acceleration of the evader. The steering control of the evader, u_e, is a scalar. We also define

$$r = r_p - r_e \qquad (3)$$

which is referred to as the 'baseline' between the pursuer and the evader.

2.1. Pursuit manifolds and cost functions

Two particle pursuer evader system is described in the Euclidean plane of two dimensions. We define the cost functions F on the pursuit manifold $G \in \mathbb{R}^2$ as $F: G \times G \to \mathbb{R}$ associated with different pursuit strategies as in [4] as follows where $(\cdot)$ represents the dot product between two unit vectors and $|\cdot|$ stands for the Euclidean norm.

$$\Gamma = \left(\frac{r}{|r|}.\frac{\dot{r}}{|\dot{r}|}\right) = \frac{\frac{d|r|}{dt}}{\left|\frac{dr}{dt}\right|} \text{ (motion camouflage)} \qquad (4)$$

and

$$\Lambda = \left(\frac{r}{|r|}\ .\ R\ x_p\right) \text{(constant bearing)} \qquad (5)$$

where

$$R = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}, \qquad (6)$$

for $\theta \in (-\pi/2, \pi/2)$.

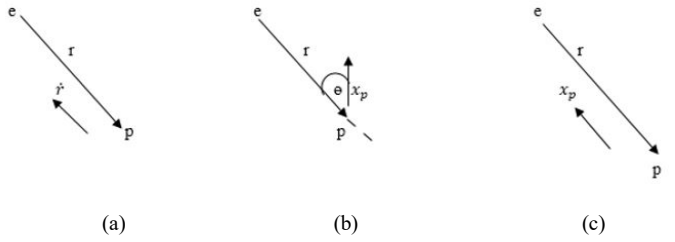
For $R = I_2$ the identity matrix of order 2, we define

$$\Lambda_0 = \left(\frac{r}{|r|}\ .\ x_p\right) \qquad (7)$$

to be the cost function associated with classical pursuit.

All three cost functions Γ, Λ and Λ_0 are well defined and that they take values in the interval [-1,1]. The cost functions Γ, Λ and

Λ_0 define the respective pursuit manifolds. Γ is seen to correspond to the cosine of the angle between r and $\dot{r}$. Camouflage pursuit manifold is defined by the condition $\Gamma = -1$, which corresponds to the case of the angle between r and $\dot{r}$ being π. (see Figure 1(a)).



(a) (b) (c)

Figure. 1 Geometric representation of pursuit manifolds: (a) motion camouflage (CM) pursuit, (b) constant bearing (CB) pursuit and (c) classical pursuit (CP).

The constant bearing pursuit manifold is represented by the condition $\Lambda = -1$. This condition is satisfied when the heading of the pursuer makes an angle θ with the baseline vector (as shown in Figure 1(b)). Similarly, the classical pursuit manifold is defined by the condition $\Lambda_0 = -1$. This condition is satisfied when the heading of the pursuer is aligned (in opposite direction) with the baseline (see Figure 1 (c)).

3. System Modelling

3.1. Pursuit-Evasion system

Representing

$$r_p = [r_{px}\ r_{py}]^t$$

$$r_e = [r_{ex}\ r_{ey}]^t$$

we write (1) and (2) in terms of state equations as follows:

$$\left.\begin{array}{ll} x_1 = r_{px}; & x_7 = r_{ex}; \\ x_2 = r_{py}; & x_8 = r_{ey}; \\ x_3 = x_{px}; & x_9 = x_{ex}; \\ x_4 = x_{py}; & x_{10} = x_{ey}; \\ x_5 = y_{px}; & x_{11} = y_{ex}; \\ x_6 = y_{py}; & x_{12} = y_{ey}; \end{array}\right\} \quad (8)$$

$$\left.\begin{array}{c} \dot{x}_1 = \dot{r}_{px} = x_{px} = x_3 \\ \dot{x}_2 = \dot{r}_{py} = x_{py} = x_4 \\ \dot{x}_3 = \dot{x}_{px} = y_{px}u_{pr} = x_5u_{pr} \\ \dot{x}_4 = \dot{x}_{py} = y_{py}u_{pr} = x_6u_{pr} \\ \dot{x}_5 = -\dot{y}_{px} = -x_{px}u_{pr} = -x_3u_{pr} \\ \dot{x}_6 = -\dot{y}_{py} = -x_{py}u_{pr} = -x_4u_{pr} \\ \dot{x}_7 = \upsilon\dot{r}_{ex} = \upsilon x_{ex} = \upsilon x_9 \\ \dot{x}_8 = \upsilon\dot{r}_{ey} = \upsilon x_{ey} = \upsilon x_{10} \\ \dot{x}_9 = \upsilon\dot{x}_{ex} = \upsilon y_{ex}u_e = \upsilon x_{11}u_e \\ \dot{x}_{10} = \upsilon\dot{x}_{ey} = \upsilon y_{ey}u_e = \upsilon x_{12}u_e \\ \dot{x}_{11} = -\upsilon\dot{y}_{ex} = -\upsilon x_{ex}u_e = -\upsilon x_9u_e \\ \dot{x}_{12} = -\upsilon\dot{y}_{ey} = -\upsilon x_{ey}u_e = -\upsilon x_{10}u_e \end{array}\right\} \quad (9)$$

$$r = r_p - r_e = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} x_1 - x_7 \\ x_2 - x_8 \end{bmatrix} \quad (10)$$

$$\dot{r} = \dot{r}_p - \dot{r}_e = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} - \begin{bmatrix} \dot{x}_7 \\ \dot{x}_8 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} - \upsilon \begin{bmatrix} x_9 \\ x_{10} \end{bmatrix} \quad (11)$$

3.2. Feedback Laws.

In this sub-section, we formulate the pursuit strategies in terms of feedback control laws. The maintenance of the cost function Γ, Λ and Λ_0 associated with a strategy at the reference value of -1 on the respective manifold is represented in the form of a feedback control system as shown in figure 2. u_{pr} will take expresson for proportional control as u_p, for proportional integral control as u_{pi} and for proportional, integral and derivative control as u_{pid}.

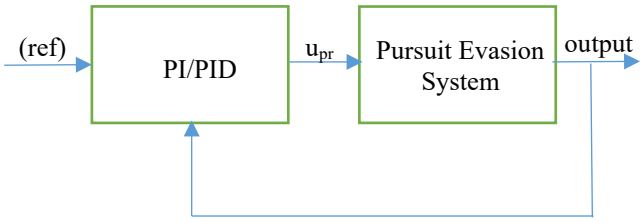


Figure. 2 Manifold Control

3.2.1. Feedback Laws for CM.

Using the results of [4] the feedback control law is defined as

$$u_p = -\mu_1 \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \quad (12)$$

where $\langle\cdot\rangle$ represents the dot product of two vectors and $x^{\perp}$ is defined as the vector x rotated counter- clockwise in the plane by an angle $\frac{\pi}{2}$. In terms of the system model, (12) can be put as

$$u_p = -\mu_1 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle \quad (13)$$

where μ_1 is the proportional (P) setting of the controller.

Proposition 1:

Using (13) under CM, PI and PID control laws can be derived as

$$u_{pi} = -\mu_1 \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle - \mu_2 \int \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle\, dt$$

$$= -\mu_1 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle - \mu_2 \int \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle \, dt \quad (14)$$

where μ_2 is the Integral (I) setting of the controller.

$$u_{pid} = -\mu_1 \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle - \mu_2 \int \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle\, dt - \mu_3 \frac{d}{dt} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle$$

$$= -\mu_1 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle - \mu_2 \int \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle \, dt$$

$$- \mu_3 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\right] . [(\upsilon^2 x_{12}\, u_e, -\upsilon^2 x_{11} u_e)] \rangle \quad (15)$$

where μ_3 is the derivative (D) setting of the controller.

Proof: - see Appendix 1

3.2.2. Feedback Laws for CB.

Considering CB, the expressions for the control law is given as follows.

$$u_p = -\mu_1 \left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right] \quad (16)$$

In terms of system model, (16) can be put as

$$u_p = -\mu_1(\eta \langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.R(x_5,x_6)^t\rangle + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle) \quad (17)$$

Proposition 2:

For CB the control laws for PI and PID can be expressed as follows.

$$u_{pi} = -\mu_1\left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right] - \mu_2 \int \left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right] dt$$

$$= -\mu_1(\eta \langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.R(x_5,x_6)^t\rangle + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle) - \mu_2 \int (\eta \langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.R(x_5,x_6)^t\rangle + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle)dt \quad (18)$$

$$u_{pid} = -\mu_1\left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right] - \mu_2 \int \left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right] dt - \mu_3 \frac{d}{dt}\left(\left[\eta \langle \frac{r}{|r|}.Ry_p\rangle + \frac{1}{|r|}\langle \frac{r}{|r|}.\dot{r}^{\perp}\rangle\right]\right)$$

$$= -\mu_1(\eta \langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.R(x_5,x_6)^t\rangle + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle) - \mu_2 \int (\eta \langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.R(x_5,x_6)^t\rangle + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle)dt - \mu_3\{(((x_3 - \upsilon x_9)\frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}} - (x_1 - x_7)\frac{1}{\left(\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}\right)^3}((x_1-x_7)(x_3-\upsilon x_9) + (x_2-x_8)(x_4 - \upsilon x_{10})), ((x_4 - \upsilon x_{10})\frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}} - (x_2 - x_8)\frac{1}{\left(\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}\right)^3}((x_1-x_7)(x_3-\upsilon x_9) + (x_2-x_8)(x_4 - \upsilon x_{10})).R(x_5,x_6)^t\rangle) + \frac{1}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}, (\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e)\rangle + \left\{-((x_1-x_7)^2+(x_2-x_8)^2)^{\frac{-3}{2}}((x_1-x_7)(x_3-\upsilon x_9) + (x_2-x_8)(x_4-\upsilon x_{10}))\right\}.\langle \frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}.(\upsilon x_{10} - x_4, x_3 - \upsilon x_9)\rangle \quad (19)$$

Proof: - See Appendix 2

3.2.3 Feedback Laws for CP.

Considering CP, the expressions for PI and PID control laws derived under proposition 2 are valid with $R = I_2$.

3.3 Control law for the Evader.

For the evader the steering control common to (CM, CB and CP of pursuer) can be as follows.

$$u_e = \cos t \quad \text{(prescribed)} \quad (20)$$

$$u_e = \langle \frac{r}{|r|}.\dot{r}_e^{\perp}\rangle \quad \text{(reactive)}$$
$$= \langle \left[\frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\right].[(-x_{10}, x_9)]\rangle \quad (21)$$

$$u_e = \langle r.\dot{r}_e\rangle \quad \text{(sudden turn)}$$
$$= \langle \left[\frac{(x_1-x_7, x_2-x_8)}{\sqrt{(x_1-x_7)^2+(x_2-x_8)^2}}\right].[x_3 - \vartheta x_9, x_4 - \vartheta x_{10}]\rangle \quad (22)$$

3.3.1 Sudden Turn of evader

Sudden turn is followed by the evader when the baseline length is shorter than a threshold. The sharp turn is again prevalent in nature where the reduced speed of the evader (victim) compared to that of the predator gives the evader the advantage of agility which enables him to turn sharply trying to escape from the pursuer. The choice of abandoning the other steering laws in favour of the sudden 90-degree turn is taken once the baseline length between the pursuer and evader is perceived to be below a certain threshold. We show below the property that once the evader takes a sudden 90 degree turn with respect to the baseline, he continues to do so in all subsequent moves since the baseline distance between the pursuer and the evader continues to be below the threshold. The property is stated and proved next for the three cases of pursuer strategy, viz, CM, CB and CP.

3.3.1.1 CM

Proposition 3:

Referring the initial baseline between the pursuer and evader as r and after a tangential move for a time dt, the baseline is denoted as r'. Assuming that the pursuer follows the CM strategy it follows that $|r'| \leq |r|$.

Proof: - See Appendix 3.

Table. 1 Time to reach $|r| = 0$ and $\Gamma = -1$ for different combinations of $P(\mu_1), I(\mu_2), D\ (\mu_3)$ for different initial values of r_p and x_p for CM.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	6.6	20.78	21.2	21.1	24.9	5.6	21.8	21.8	27.2	4.5	9.2	1
1,1,0	6.4	13.3	19	19.5	14.7	2.3	3.2	7	19.7	1.8	8.1	0.7
1,2,0	6.3	9.8	18.2	19.7	13.7	1.6	2.9	3.8	18.3	1.4	7.8	0.5
1,3,0	6.3	9.1	17.6	20.1	13.1	1.3	2.8	3.5	17.3	1.1	7.5	0.4
1,1,1	6.4	7.3	19.4	20	13.9	14.6	3.4	4.6	19.7	1.9	8.6	1.3
1,1,2	6.4	7.1	20.1	20.5	13	13.5	3.7	4.5	19.8	1.9	9.1	14.8
1,1,3	6.5	7.1	22.2	22.6	12.5	13.1	4.1	4.1	19.8	1.9	8.4	10.7
1,2,1	6.3	6.9	18.6	21.3	14.1	15.6	3	3.6	18.3	1.4	8.1	0.9
1,2,2	6.4	6.8	19.2	20.7	12.6	12.9	3.2	3.7	18.3	1.4	9	10.4
1,2,3	6.4	6.8	20.5	21.4	12.1	15.3	3.5	3.5	18.3	1.4	8.2	9.4
1,3,1	6.3	6.7	18	20.9	13.5	1.6	2.9	3.4	17.2	1.1	7.8	0.6
1,3,2	6.3	6.6	18.5	20.9	12.4	14.5	3	3.4	17.3	1.1	8.5	1.3
1,3,3	6.3	6.3	19.7	19.8	11.8	14.6	3.2	3.8	17.3	1.1	8	9.2

Table. 2 Time to reach $|r| = 0$ and $\Lambda = -1$ for different combinations of $P\ (\mu_1), I(\mu_2), D(\mu_3)$ for different initial values of r_p and x_p for CB.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	32.3	6.8	32.3	9.8	39.9	6.8	2.9	2.9	NA	6.2	25.7	0.5
1,1,0	17.4	6.6	30.2	9.3	30.6	2.7	9.3	1.7	36.1	2.4	17.8	0.4
1,2,0	17.3	6.6	29.6	9.3	28	2	9	1.5	34.1	1.8	17.8	0.4
1,3,0	16.9	6.6	29.2	9.1	27.6	1.2	8.8	1.2	33.6	2.5	17.7	0.3
1,1,1	17.1	6.5	30	9.4	30.2	2.7	5.4	1.8	36.1	2.2	17.3	0.4
1,1,2	17	6.5	29.8	9.3	30	2.7	4.5	1.9	36.1	2.6	16.9	0.4
1,1,3	16.7	6.5	29.6	9.3	29.6	2.7	3.8	1.9	36	2.7	16.6	0.4
1,2,1	17	6.5	29.5	9.2	28	2	10.2	1.4	34.1	1.8	17.6	0.4
1,2,2	16.8	6.5	29.3	9.2	27.8	2.1	9	1.5	34	1.9	17.3	0.4
1,2,3	16.6	6.5	29	9.1	27.6	2	8.1	1.5	34	2	17.1	0.3
1,3,1	16.7	6.6	29.1	9.1	27.5	1.7	9.7	1.3	33.5	1.6	17.6	0.3
1,3,2	16.5	6.5	28.9	9.1	27.4	1.7	8.7	1.3	33.4	1.6	17.3	0.3
1,3,3	16.4	6.5	28.9	9.1	27.4	1.7	8	1.3	33.4	1.6	16.5	0.3

3.3.1.2 CB

Proposition 4:

For the case of CB, considering the instantaneous case of sudden turn of 90 degrees left or right by the evader, with the pursuer following the CB law under two cases.

where *r* and *r'* are the baseline vectors initially and after the lapse of an infinitesimal time respectively.

and (ii) $\emptyset < \frac{\pi}{2}$, it follows that

$$|r| = |r'|$$

if

$$\left(\frac{\vartheta}{p} + \sin\emptyset\right)^2 + (\cos\emptyset)^2 = \frac{2|r|\cos\emptyset}{p\,dt}$$

where dt is the instantaneous period of the step considered. ϑ and p are the velocities of the evader and the pursuer respectively.

Proof: - See Appendix 4.

Remark 1: To illustrate the use of result of proposition 4(ii)

Assume, $k = \frac{|r|}{p\,dt}$.

When $\vartheta = 0.6, p = 1, |r| = 1, p\,dt = 0.1$,

$k = 10$. Then as per the result of proposition 4 (ii),

$$(0.6 + \sin\emptyset)^2 + (\cos\emptyset)^2 = 20\cos\emptyset$$

$$\emptyset \approx 82.65 \text{ degrees}$$

Table. 3 Time to reach $|r| = 0$ and $\Lambda_0 = -1$ for different combinations of $P(\mu_1), I(\mu_2), D(\mu_3)$ for different initial values of r_p and x_p for CP.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{rf}	t_r	t_{rf}	t_r	t_{rf}	t_r	t_{rf}	t_r	t_{rf}	t_r	t_{rf}
1,0,0	6.7	26.5	18.6	NA	12.6	NA	2.7	6.2	17.9	NA	8.3	0.6
1,.03,0	6.7	10.6	18.6	18.8	12.6	13.1	2.6	4.5	17.7	18.3	8.2	0.7
1,.04,0	6.7	9.4	18.5	18.6	12.6	12.7	2.6	4.2	17.7	17.9	8.1	0.6
1,.05,0	6.6	7.6	18.5	18.5	12.5	12.6	2.6	3.9	17.7	17.8	8.1	0.6
1,0,0.1	6.6	8.9										
1,0,.07			18.6	20.6								
1,0,.02					12.6	14.7			17.9	21		
1,0,.03							2.6	18.1				
1,0,.01											8.2	0.6

3.3.1.3 CP

Proposition 5:

Assume *r* and *r'* defined as in Proposition 3. When the evader takes a sudden turn left or right 90 degree, with pursuer following CP law case, it can be shown that

$$|r'| \leq |r|$$

Proof: - See Appendix 5.

4. Simulation Results.

4.1. Time response

4.1.1 Non-reactive case

In this section, we provide the results of simulation of the dynamic equations given in section 3 and discuss the same. We assume u_e given by (20) corresponds to the non-reactive case of the evader. Tables 1-3 provide results obtained through computer, simulation of the pursuit evasion game for CM, CB and CP strategies respectively under P, PI and PID control. The first columns of Tables 1-3 provide the P, I and D gains used corresponding to μ_1, μ_2, μ_3 respectively. The next six columns correspond to different initial starting coordinates for r_p and x_p, r_e and x_e are assumed to be $(0,0)^t$ and $(1,0)^t$ respectively. The two sub- columns in Tables 1-3 correspond to time (t_r) to reach zero for the magnitude of r and the time (t_{mf}) to reach manifold characterized by $\Gamma = -1$, $\Lambda = -1$ and $\Lambda_0 = -1$ corresponding to CM, CB and CP strategies respectively for the first time. In Table 3, the blank cells correspond to the case of no significant results being obtained in the simulation.

It is seen in Table 1, corresponding to CM, that PI and PID laws tend to improve the performance of the pursuer compared to using P alone, which is the existing method. For example, comparing the rows corresponding to (1, 3, 0) and (1, 3, 3) against (1, 0, 0) in the first column, it is seen that t_r and t_{mf} are reduced in the case of PI and PID settings compared to P alone. This holds for all possible starting points considered. Similarly, it is seen in Table 2, corresponding to CB that PI and PID laws tend to improve on the performance of the pursuer compared to using P alone, which can be considered as the existing method. That is, t_r and t_{mf} are much reduced in the case of PI and PID settings compared to P alone. This holds for all possible starting points considered. Similarly, it is seen in Table 3, corresponding to CP, that PI and PID laws tend to improve the performance of the pursuer compared to using P alone, which is the existing method. That is, t_r and t_{mf} are much reduced in the case of PI and PID settings compared to P alone. This holds for all possible starting points considered.

Table. 4 Time to reach $|r| = 0$ and $\Gamma = -1$ for different combination of $P(\mu_1), I(\mu_2), D(\mu_3)$ for different initial values of r_p and x_p for CM.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	5.2	5.2	10.1	16.7	10.1	16.2	13.3	21.1	13.7	25.7	8.1	0.6
1,1,0	5.0	6.4	8.8	13.5	9.1	14.1	2.5	6.8	9.4	1.9	7.4	0.5
1,2,0	4.9	5.8	8.4	11.4	7.6	1.7	2.4	6.5	8.8	1.4	7.2	0.4
1,3,0	4.9	5.5	8.2	10.5	7.2	1.3	2.3	34.6	8.6	1.2	7.1	0.3
1,1,1	5	6.3	8.8	13.5	7.6	12.6	2.6	7.4	9.7	2.7	7.3	0.5
1,1,2	4.9	6.1	8.9	13.4	7.3	12.1	2.7	8.1	10.4	14.8	7.3	0.5
1,1,3	4.9	6	9	13.4	7	11.9	2.8	2.7	9.5	14	7.3	0.4
1,2,1	4.9	5.8	8.5	11.4	7.4	10.6	2.4	19.3	8.9	1.5	7.2	0.4
1,2,2	4.9	5.7	8.6	11.4	7.1	10.2	2.5	7.8	8.9	1.5	7.1	0.4
1,2,3	4.8	5.6	8.9	11.5	6.9	9.9	2.6	2.6	9	1.6	7.1	0.4
1,3,1	4.9	5.5	8.3	10.5	7.5	9.8	2.4	11.9	8.5	1.2	7	0.4
1,3,2	4.8	5.4	8.4	10.5	7	9.3	2.4	8.7	8.5	1.2	7	0.4
1,3,3	4.8	5.4	8.7	10.7	6.8	9.1	2.5	14.4	8.6	1.2	7.1	0.3

Table. 5 Time to reach $|r| = 0$ and $\Lambda = -1$ for different combination of $P(\mu_1), I(\mu_2), D(\mu_3)$ for different initial values of r_p and x_p for CB.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	5.7(1.6)	5.8	9.7(1.7)	9.7	13.9(1.8)	4.9	2.7	10.9	15.1(1.8)	4.5	9.3(1.8)	0.4
1,1,0	23.1	5.5	25.1	9.3	25.4	2.4	22.9(.52)	1.8	26.4(.51)	2.2	24.5(.52)	0.3
1,2,0	20.3	5.4	23.3	9.1	23.2	1.9	14.2(.53)	1.4	24.5	1.7	22.5	0.3
1,3,0	19.7	5.3	22.5	9.1	22.3	1.7	11.7	1.2	23.6	1.5	21.7	0.2
1,1,1	20.7	5.4	23.3	9.2	23.7	2.4	16.3(.53)	1.8	25	2.2	22.6	0.3
1,1,2	18.8	5.4	21.8	9.1	22.2	2.3	10.7(.51)	1.9	23.8	2.3	21.1	0.3
1,1,3	17.1	5.3	20.5	9	20.8	2.3	8	1.9	22.6	2.3	19.7	0.3
1,2,1	19	5.3	22.4	9.1	22.3	1.9	11.6(.51)	1.4	23.8	1.8	21.6(.51)	0.3
1,2,2	18.5	5.3	21.7	9	21.6	1.9	9.6	1.5	23.2	1.8	20.8	0.3
1,2,3	17.6	5.3	21	8.9	20.8	1.8	8	1.5	22.6	1.8	20.1	0.4
1,3,1	18.9	5.3	22	8.9	21.6	1.6	10.3	1.3	23.1	1.5	21.1	0.2
1,3,2	18.3	5.3	21.5	8.9	21.1	1.6	8.8	1.3	22.7	1.5	20.6	0.2
1,3,3	17.7	5.3	21	8.8	20.6	1.6	8.4	1.3	22.3	1.5	20.1	0.3

Comparing the performance of CM, CB and CP under feedback laws, the initial condition column (1, 3, 5, 6) in Tables 1-3, show that the CM outperforms followed by CP and CB in terms of shorter t_r. In column 2 in Tables 1-3 the CP outperforms followed by CM and CB in terms of shorter t_r, and in column 4 in Tables 1-3 the CP outperforms followed by CM and CB in terms of shorter t_r. Similarly, in terms of shorter t_{mf} columns (1-4) in Tables 1-3, CM outperforms followed by CB and CP. For column 5 in Tables 1-3, CB outperforms followed by CM and CP in terms of shorter t_{mf}. For column 6 in Tables 1-3, CB outperforms followed by CP and CB in terms of shorter t_{mf}.

4.1.2 Reactive case

Simulation results of the dynamic equations given in section 3 by using u_e corresponding to (21) are given below.

It is shown in Tables 4-6, that PI/PID control (proposed) laws correspond to much shorter t_r and t_{mf} compared to P alone for the three strategies CM, CB and CP.

Comparison of the different strategies, under feedback laws, shows that in the column (1, 3, 6) in Tables 4-6 CM outperforms followed by CP and CB. In column (2, 4, 5) of Tables 4-6 CP outperforms followed by CM and CB in terms of shorter t_r. Similarly, in terms of shorter t_{mf} column (1, 4) in Tables 4-6 shows that CB outperforms followed by CM and CP. For column (2, 3) in Tables 4-6, CB outperforms followed by CP and CM in terms of t_r and t_{mf}. For column (5, 6) in Tables 4-6, CM outperforms followed by CB and CP in terms of t_r and t_{mf}. The controller setting of D is not included in Table 6 since no significant improvement is seen through simulation using derivative control.

4.1.3 Sudden Turn

Simulation results of the dynamic equations given in section 3 by using u_e corresponding to (22) are given below.

Tables 7-9 give the data of t_r and t_{mf} of CM, CB and CP strategies under sudden turn evader steering control laws. The empty cells in these tables mean insignificant value within the considered time frame. In CM strategies, we can see from Table 4 and Table 7 that by using sudden turn evader control laws, t_r and t_{mf} are much delayed (compare the first rows of Table 4 and

Table. 6 Time to reach $|r| = 0$ and $\Lambda_0 = -1$ for different combination of $P(\mu_1), I(\mu_2), D(\mu_3)$ for different initial values of r_p and x_p for CP.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	5	8.7	8.1	9	7	9.2	2.2	6	8.5	10.3	7.9	0.5
1,.03,0	5	7.5	8.1	9.1	7	8.4	2.2	10.1	8.5	10.1	7.8	0.5
1,.04,0	5	8	8.1	9.1	7	8.3	2.2	6.3	8.5	10.1	7.8	0.5
1,.05,0	5	8.4	8.1	9.2	7	8.2	2.2	5.1	8.5	10	7.8	0.5

Table 7) from using reactive or non-reactive evader control law for different initial conditions. In case of CB strategy, we can see from Table 5 and Table 8 that by using sudden turn evader control laws, evader escapes from the pursuer for all different initial conditions (compare the first rows of Table 5 and Table 8). This is a very significant result for the evader. For CP strategies, from Table 6 and Table 9, we can see that by using sudden turn evader control laws, capturing of evader is delayed for all initial conditions compared to the case when evader used reactive control law (compare the first rows of Table 6 and Table 9).

Table. 7 Time to reach $|r| = 0$ and $\Gamma = -1$ for P (μ_1) control for different initial values of r_p and x_p for CM.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	28.6	17.6		20.7	21.4	21.3	21.3	10.9	23.3	4.9	16.1	0.8

Table. 8 Time to reach $|r| = 0$ and $\Gamma = -1$ for P(μ_1) control for different initial values of r_p and x_p for CB.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0		4.1		7.1		5.4		3.2		4.5		0.5

Table. 9 Time to reach $|r| = 0$ and $\Gamma = -1$ for P(μ_1) control for different initial values of r_p and x_p for CP.

μ_1,μ_2,μ_3	$(r_{px},r_{py}),(x_{px},x_{py})$ (4, -6),(0.735,0.678) (1)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-4, 9),(0.327,0.945) (2)		$(r_{px},r_{py}),(x_{px},x_{py})$ (0, 8),(0.530,0.848) (3)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-1, 1),(0.632,0.775) (4)		$(r_{px},r_{py}),(x_{px},x_{py})$ (-2, 10),(0.335,0.942) (5)		$(r_{px},r_{py}),(x_{px},x_{py})$ (10,-1),(0.9,0.436) (6)	
	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}	t_r	t_{mf}
1,0,0	10.3	14.3	15.1		14.9	18.7	4	8	15.9		15.6	0.6

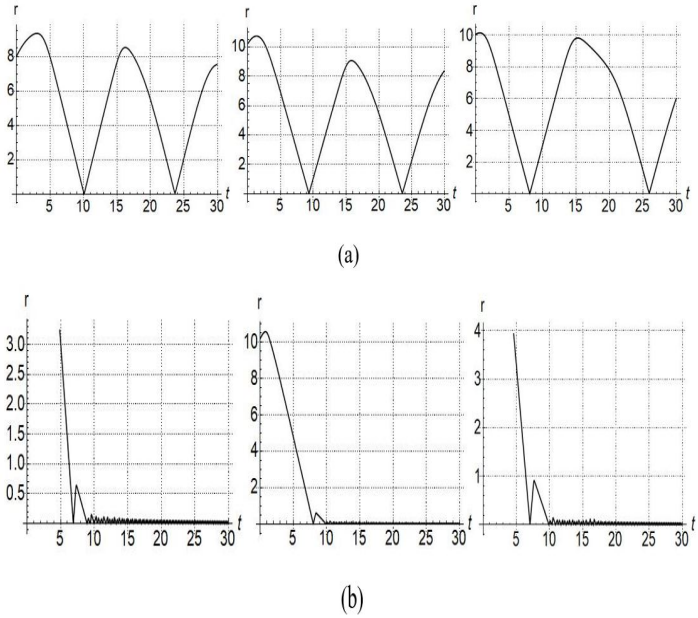


Figure. 3 r vs t plot for CM using (a) P and (b) PID.

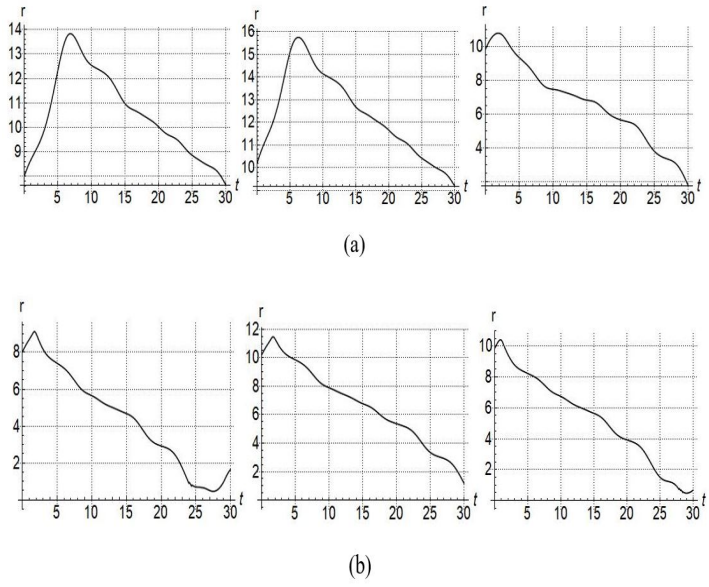
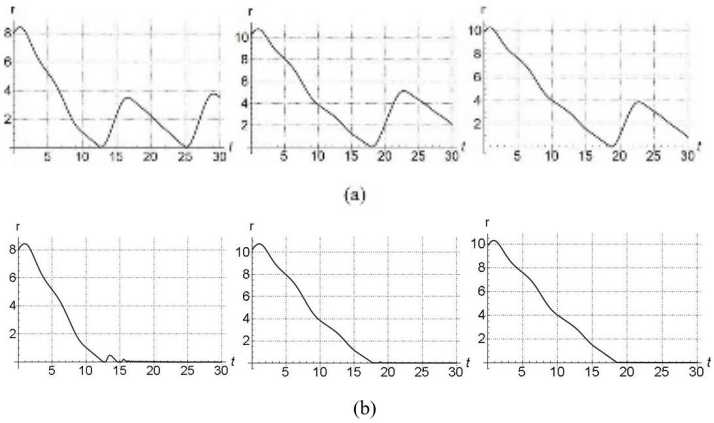


Figure. 4 r vs t plot for CB using (a) P and (b) PID.



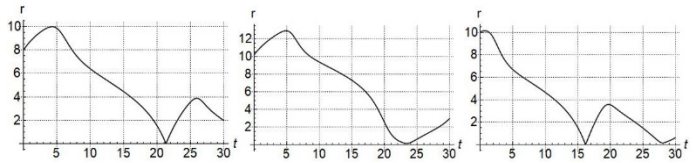
(a)

(b)

Figure. 5 r *vs* t plot for CP using (a) P and (b) PI.

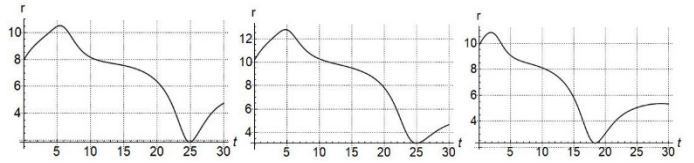
4.2. Tracking response

Figures 3-4 show the response of the baseline magnitude $|r|(t)$ as function of time for different initial settings. When |r| becomes zero, it means the distance between pursuer and evader is zero, and hence the evader is captured. Figure 3(a) and Figure 4(a) correspond to P control while Figure 3(b) and Figure 4(b) corresponds to PID control. Similarly, for Figure 5(a) and Figure 5(b) correspond to P and PI respectively for different initial conditions. It is clearly seen from Figures 3-5 that tracking response of the zero line of the baseline magnitude (i.e. the zero line of the distance between the pursuer and evader) is clearly superior for PID/ PI control, compared to P alone (top rows in Figures 3-5). In fact, P alone is not able to track the zero-baseline magnitude at all. Note that the magnitude of baseline vector going to zero corresponds to evader being captured.



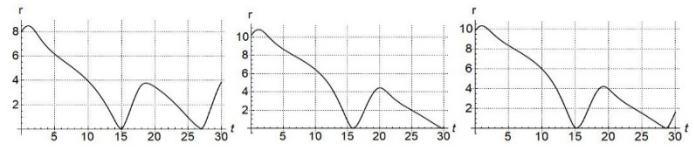
(a) (b) (c)

Figure. 6 r *vs* t plot for CM using sudden turn evader steering law.



(a) (b) (c)

Figure. 7 r *vs* t plot for CB using sudden turn evader steering law.



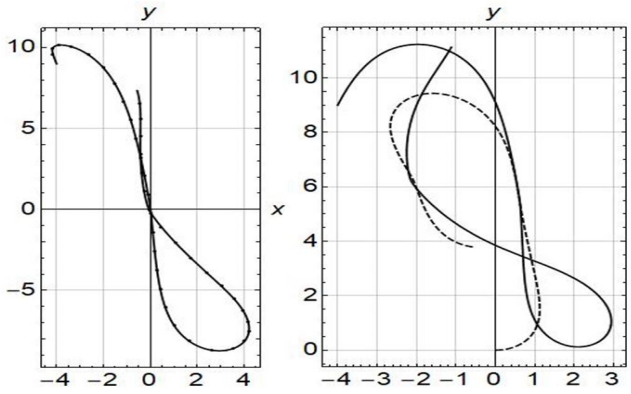
(a) (b) (c)

Figure. 8 r *vs* t plot for CP using sudden turn evader steering law.

Figures 6-8 show the response of the baseline magnitude $|r|(t)$ as function of time for different initial settings for CM, CB and CP strategies respectively. In Figure 6(a) represents the initial starting position of pursuer given as (0, 8) in x- y co-ordinate. 6(b) and 6(c) correspond to starting positions of (-2, 10) and (-4, 9) respectively. For the sake of comparison, (a), (b) and (c) in Figures 7 and 8 also correspond to the same respective starting positions. It is clearly seen from Figures 6-8 that the time to capture the evader by the pursuer is much delayed by using the sudden turn strategy compared to the case when evader used the reactive and non-reactive steering control laws. This can be seen by comparing Figure 3(a) with Figure 6, Figure 4(a) with Figure 7 and Figure 5(a) with Figure 8. In fact in Figure 4, $|r|(t)$ never reaches zero meaning that the evader is never captured in that case.

4.3. Trajectories of pursuer and evader.

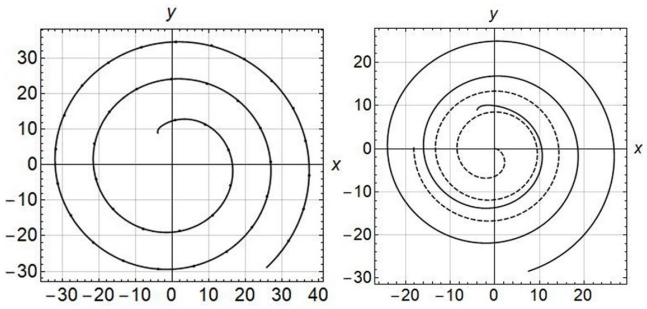
We consider in this section, the actual trajectories followed by evader and pursuer.

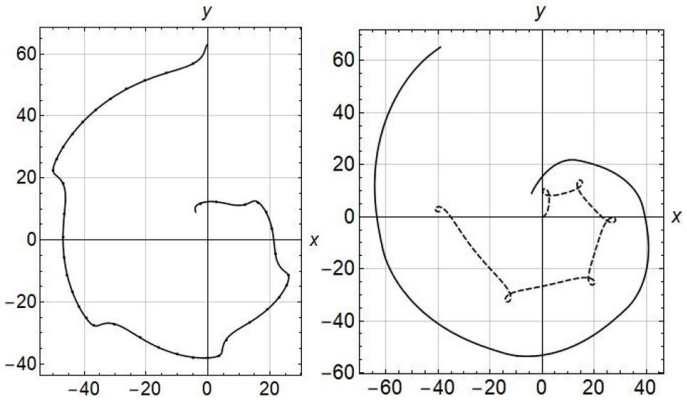


(a) (b)

Figure. 9 For normal pursuer capture evader, (a) magnitude of r, (b) trajectories of pursuer (solid line) and evader (dashed line).

Figure 9 shows evader being captured in a certain time by the pursuer. Figure 9(a) shows the magnitude of r. Figure 9(b) shows the trajectories of pursuer and evader for a 30 sec timeline. The starting point of evader is (0, 0) in x-y coordinate and the starting point for the pursuer is (-4, 9). Figure 10 shows the case when the evader escapes from the pursuer. This happens only in the case of





(a) (b)

Figure. 11 For $\emptyset = 82.65°$, (a) magnitude of r, (b) trajectories of pursuer (solid line) and evader (dashed line).

CB strategies when evader suddenly turn 90 degree left or right as discussed in Proposition 4 case (i). So, in Figure 10 (a) the value of r keeps on increasing. Figure 10 (b) shows the trajectories of pursuer and evader in 30 sec timeline. Figure 11 is for the case when $\emptyset = 82.65°$ and the baseline distance between the pursuer and the evader tend to reach a constant value.

5. Conclusions

Pursuit evasion game (PEG) has been studied in the literatiure using geometric control theory. Control laws for the pursuer have been derived so as to make the pursuer follow a certain manifold according to the different bio-inspired strategies used. The performance of the pursuer control laws is specified in terms of time to capture the evader and tarcking of pursuer path on the manifold given certain fixed evader escape strategies Though the existing pursuer control laws are optimal, it is shown in the paper, through simulations, that adding integral and derivative actions to the pursuer control laws tends to improve the performance of the pursuer further independent of the strategy used. Towards that, the pursuer path tracking on the manifold is specified in terms of a feedback loop with the reference being the cost function of the manifold used. The proportional (P), proportional- integral (PI) and proportional- integral- derivative (PID) control laws are studied on the loop with the existing control law being considered as proportional law. Due to the noninear nature of control laws in a vector setting the derivation of PI and PID control laws are non-trivial.To mimic the nature further, a sudden turn strategy is assumed to be employed by the evader in view of evader's agility advantage while moving at a lower speed compared to the pusuer. The dynamics of the PEG with the sudden turn strategy of the evader has been studied through a theoretical analysis supported by simulation. It is shown that the sudden turn strategy can help the evader delay the capture by the pursuer and in a certain case the evader can even totally excape from the pursuer. As a future work, one could consider the case when the evader moves at a varying speed (instead of the constant speed assumed in the paper) as in commonly seen in nature.

References

[1] P. Svec, A. Thakur, B. C. Shah, and S. K. Gupta, "USV Trajectory Planning for Time Varying Motion Goals in an Environment with obstacles," *Terpconnect.Umd.Edu*, pp. 1–11, 2012.

[2] R. Issacs, "Differential Games," New York: Wiley, 1965.

[3] E. Wei, E. W. Justh, and P. S. Krishnaprasad, "Pursuit and an evolutionary game," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 465, no. 2105, pp. 1539–1559, 2009.

[4] D. Pais and N. Leonard, "Pursuit and Evasion: Evolutionary Dynamics and Collective Motion," *AIAA Guid. Navig. Control Conf.*, pp. 1–14, 2010.

[5] W. Li, "A Dynamics Perspective of Pursuit-Evasion: Capturing and Escaping When the Pursuer Runs Faster Than the Agile Evader," *IEEE Trans. Automat. Contr.*, vol. 62, no. 1, pp. 451–457, Jan. 2017.

[6] W. Li, "Escape Analysis on the Confinement-Escape Problem of a Defender Against an Evader Escaping from a Circular Region," *IEEE Trans. Cybern.*, vol. 46, no. 9, pp. 2166–2172, 2016.

[7] L. Obiroy Singh and R. Devanathan, "Performance improvement of bio-inspired strategies through feedback laws," in *Advances in Intelligent Systems and Computing*, 2018, pp. 143–156.

[8] L. Obiroy Singh and R. Devanathan, "Comparative Study of the Performance of Application of Bio-Inspired Strategies to Pursuit Evasion Game Under Feedback Laws," in *The 18th International Conference on Control, Automation and Systems,South Korea*, 2018, pp. 276–281.

[9] J. H. S.D Bopardikar, F Bullo, "A Cooperative Homicidal Chauffeur Game," *Automatica*, vol. 45, pp. 1771–1777, 2009.

[10] X. Liang and Y. Xiao, "Studying Bio-Inspired Coalition Formation of Robots for Detecting Intrusion Using Game Theory," *IEEE Trans. Syst. Man Cybern.*, vol. 40, no. 3, pp. 683–693, 2010.

[11] H. . Cott, "Adaptive Coloration in animals," *London: Methues*, pp. 141–143, 1966.

[12] K. . Roeder, "Nerve Cells and insect behaviour," *Cambridge, MassachusettsHarvard Univ. Press*, pp. 160--164, 1967.

[13] M. V Srinivasan and M. Davey, "Strategies for Active Camouflage of Motion," in *Biological Sciences*, pp. 19–25.

[14] E. Simetti, A. Turetta, G. Casalino, E. Storti, and M. Cresta, "Protecting Assets within a Civilian Harbour through the use of a Team of USVs: Interception of Possible Menaces," in *IARP workshop on robots for risky interventions and environmental surveillance maintenance (RISE'10)*, 2010.

[15] J. Walrand, E. Polak, and H. Chung, "Harbor Attack: A Pursuit- Evasion Game," in *Forty- Ninth Annual Allerton Conference*, 2011, pp. 1584–1591.

[17] E. W. Justh and P. S. Krishnaprasad, "Steering laws for motion camouflage," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 462, no. 2076, pp. 3629–3643, 2006.

[18] E. W. Justh and P. S. Krishnaprasad, "Natural frames and interacting particles in three dimensions," *Proc. 44th IEEE Conf. Decis. Control. Eur. Control Conf. CDC-ECC '05*, vol. 2005, pp. 2841–2846, 2005.

Appendix 1:

Proof of Proposition 1

P and PI control:

u_p *and* u_{pi} are the control output of P and PI controller derived as given in the proposition in a straightforward way.

PID Control output u_{pid} can be derived as follows.

$$u_{pid} = -\mu_1 \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle - \mu_2 \int \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \, dt - \mu_3 \frac{d}{dt} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle$$

$$= -\mu_1 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle -$$

$$\mu_2 \int \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle \, dt$$

$$-\mu_3 \frac{d}{dt} \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle$$

Using the identity

$$\frac{d}{dt} \langle \alpha, \beta \rangle = \langle \dot{\alpha}, \beta \rangle + \langle \alpha, \dot{\beta} \rangle \quad \text{(A.1)}$$

put

$$\alpha = \frac{r}{|r|} \text{ and } \beta = \dot{r}^{\perp}$$

that is,

$$\alpha = \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] \text{ and } \beta = [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)]$$

Considering equation (A.1), since

$\dot{\alpha} = \frac{\dot{r}}{|\dot{r}|}$ is orthogonal to $\beta = \dot{r}^{\perp}$

the first term of R.H.S of equation (A.1) is zero.

Now

$$\beta = [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)]$$

from which

$$\dot{\beta} = [(\upsilon \dot{x}_{10} - \dot{x}_4), (\dot{x}_3 - \upsilon \dot{x}_9)]$$

that is,

$$\dot{\beta} = \left[\left(\upsilon^2 x_{12} u_e - x_6 u_p \right), \left(x_5 u_p - \upsilon^2 x_{11} u_e \right) \right]$$

using (9).

Approximating the term $\dot{\beta}$ as $\hat{\dot{\beta}}$ by neglecting the term associated with u_p,

$$\hat{\dot{\beta}} = [\upsilon^2 x_{12} u_e, -\upsilon^2 x_{11} u_e]$$

Therefore, equation (A.1) becomes

$$\frac{d}{dt} \langle \alpha, \beta \rangle = \langle \alpha, \hat{\dot{\beta}} \rangle$$

$$\frac{d}{dt} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle = \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon^2 x_{12} u_e, -\upsilon^2 x_{11} u_e)] \rangle$$

Therefore

$$u_{pid} = -\mu_1 \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle - \mu_2 \int \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \, dt - \mu_3 \frac{d}{dt} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle$$

$$= -\mu_1 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle -$$

$$\mu_2 \int \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)] \rangle \, dt$$

$$-\mu_3 \langle \left[\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \right] . [(\upsilon^2 x_{12} u_e, -\upsilon^2 x_{11} u_e)] \rangle$$

Hence the result.

Appendix 2:

Proof of Proposition 2:

P and PI control:

u_p *and* u_{pi} are easily derived as given in the proposition.

PID control:

$$u_{pid} = -\mu_1 \left[\eta \langle \frac{r}{|r|} . Ry_p \rangle + \frac{1}{|r|} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \right] - \mu_2 \int \left[\eta \langle \frac{r}{|r|} . Ry_p \rangle + \frac{1}{|r|} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \right] dt - \mu_3 \frac{d}{dt} \left(\left[\eta \langle \frac{r}{|r|} . Ry_p \rangle + \frac{1}{|r|} \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \right] \right) \quad \text{(A.2)}$$

Using the identity

$$\frac{d}{dt} \langle \rho, \omega \rangle = \langle \dot{\rho}, \omega \rangle + \langle \rho, \dot{\omega} \rangle \quad \text{(A.3)}$$

$$\frac{d}{dt} (\gamma \sigma) = (\dot{\gamma} \sigma) + (\gamma \dot{\sigma}) \quad \text{(A.4)}$$

put

$$\rho = \frac{r}{|r|}, \omega = Ry_p, \gamma = \frac{1}{|r|} \text{ and } \sigma = \langle \frac{r}{|r|} . \dot{r}^{\perp} \rangle \quad \text{(A.5)}$$

Now

$$\dot{\rho} = \frac{d}{dt} \frac{r}{|r|} = \begin{bmatrix} \frac{d}{dt} \frac{r_x}{|r|} \\ \frac{d}{dt} \frac{r_y}{|r|} \end{bmatrix} = \begin{bmatrix} \frac{|r| \dot{r}_x - r_x \frac{d}{dt} |r|}{|r|^2} \\ \frac{|r| \dot{r}_y - r_y \frac{d}{dt} |r|}{|r|^2} \end{bmatrix}$$

$$\dot{\rho} = \begin{bmatrix} \frac{\dot{r}_x}{|r|} - \frac{r_x \frac{d}{dt} \sqrt{r_x^2 + r_y^2}}{|r|^2} \\ \frac{\dot{r}_y}{|r|} - \frac{r_y \frac{d}{dt} \sqrt{r_x^2 + r_y^2}}{|r|^2} \end{bmatrix} = \begin{bmatrix} \frac{\dot{r}_x}{|r|} - \frac{r_x (r_x \dot{r}_x + r_y \dot{r}_y)}{|r|^3} \\ \frac{\dot{r}_y}{|r|} - \frac{r_y (r_x \dot{r}_x + r_y \dot{r}_y)}{|r|^3} \end{bmatrix} \quad \text{(A.6)}$$

$$\dot{\omega} = \frac{d}{dt} Ry_p = R \begin{bmatrix} \frac{d}{dt} y_{px} \\ \frac{d}{dt} y_{py} \end{bmatrix} = R \begin{bmatrix} x_{px} u_p \\ x_{py} u_p \end{bmatrix} \quad \text{(A.7)}$$

$\dot{\omega}$ is neglected since it is associated with the term u_p.

Therefore,

$$\frac{d}{dt} \langle \rho, \omega \rangle = \langle \dot{\rho}, \omega \rangle$$

$$= \left\langle \begin{bmatrix} \frac{\dot{r}_x}{|r|} - \frac{r_x(r_x\dot{r}_x + r_y\dot{r}_y)}{|r|^3} \\ \frac{\dot{r}_y}{|r|} - \frac{r_y(r_x\dot{r}_x + r_y\dot{r}_y)}{|r|^3} \end{bmatrix}, R\begin{bmatrix} y_{px} \\ y_{py} \end{bmatrix} \right\rangle$$

$$= \left\langle \left(\frac{x_3 - \upsilon x_9}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} - \frac{(x_1 - x_7)\big((x_1 - x_7)(x_3 - \upsilon x_9) + (x_2 - x_8)(x_4 - \upsilon x_{10})\big)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}^3} \right), R(x_5, x_6) \right\rangle$$

using equations (10) and (11).

$$\dot{\gamma} = \frac{d}{dt}\frac{1}{|r|} = \frac{d}{dt}\left(\sqrt{r_x{}^2 + r_y{}^2}\right)^{-\frac{1}{2}}$$

$$= -\left(\sqrt{r_x{}^2 + r_y{}^2}\right)^{-3}\left(r_x\dot{r}_x + r_y\dot{r}_y\right) \quad \text{(A.8)}$$

$$\dot{\sigma} = \frac{d}{dt}\langle\frac{r}{|r|}.\dot{r}^{\perp}\rangle = \langle\frac{\dot{r}}{|\dot{r}|}.\dot{r}^{\perp}\rangle + \langle\frac{r}{|r|}.\ddot{r}^{\perp}\rangle \quad \text{(A.9)}$$

$\langle\frac{\dot{r}}{|\dot{r}|}.\dot{r}^{\perp}\rangle = 0$, since $\frac{\dot{r}}{|\dot{r}|}$ is orthogonal to $\dot{r}^{\perp}$

So,

$$\dot{\sigma} = \frac{d}{dt}\langle\frac{r}{|r|}.\dot{r}^{\perp}\rangle = \langle\frac{r}{|r|}.\ddot{r}^{\perp}\rangle$$

Now,

$$\ddot{r}^{\perp} = \frac{d}{dt}\dot{r}^{\perp} = \frac{d}{dt}[\upsilon x_{10} - x_4, x_3 - \upsilon x_9]$$

$$= [\upsilon\dot{x}_{10} - \dot{x}_4, \dot{x}_3 - \upsilon\dot{x}_9] = \left[\upsilon^2 x_{12}u_e - x_6 u_p, x_5 u_p - \upsilon^2 x_{11}u_e\right]$$

using (9).

Neglecting the term associated with u_p,

$$\ddot{r}^{\perp} = [\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e] \quad \text{(A.10)}$$

Therefore,

$$\dot{\sigma} = \langle\frac{r}{|r|}.(\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e)\rangle \quad \text{(A.11)}$$

Thus,

$$\frac{d}{dt}(\gamma\sigma) = (\dot{\gamma}\sigma) + (\gamma\dot{\sigma})$$

$$= \left\{-\left(\sqrt{r_x{}^2 + r_y{}^2}\right)^{-3}\left(r_x\dot{r}_x + r_y\dot{r}_y\right).\left(\frac{\dot{r}}{|\dot{r}|}.\dot{r}^{\perp}\right)\right\} + \frac{1}{|r|}.\langle\frac{r}{|r|}.(\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e)\rangle$$

$$= \frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} \langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}, (\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e)\rangle$$

$$+ \left\{-((x_1 - x_7)^2 + (x_2 - x_8)^2)^{\frac{-3}{2}}\big((x_1 - x_7)(x_3 - \upsilon x_9) + (x_2 - x_8)(x_4 - \upsilon x_{10})\big)\right\}.\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.(\upsilon x_{10} - x_4, x_3 - \upsilon x_9)\rangle$$

Hence it follows that

$$u_{pid} = -\mu_1(\eta\,\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.R(x_5, x_6)^t\rangle + \frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle) - \mu_2\int(\eta\,\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.R(x_5, x_6)^t\rangle + \frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.(\upsilon x_{10} - x_4), (x_3 - \upsilon x_9)\rangle)dt$$

$$-\mu_3\{(((x_3 - \upsilon x_9)\frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} - (x_1 - x_7)\frac{1}{\left(\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}\right)^3}((x_1 - x_7)(x_3 - \upsilon x_9) + (x_2 - x_8)(x_4 - \upsilon x_{10}), ((x_4 - \upsilon x_{10})\frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}} - (x_2 - x_8)\frac{1}{\left(\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}\right)^3}\big((x_1 - x_7)(x_3 - \upsilon x_9) + (x_2 - x_8)(x_4 - \upsilon x_{10})\big).R(x_5, x_6)^t\rangle) + \frac{1}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}, (\upsilon^2 x_{12}u_e, -\upsilon^2 x_{11}u_e)\rangle + \left\{-((x_1 - x_7)^2 + (x_2 - x_8)^2)^{\frac{-3}{2}}\big((x_1 - x_7)(x_3 - \upsilon x_9) + (x_2 - x_8)(x_4 - \upsilon x_{10})\big)\right\}.\langle\frac{(x_1 - x_7, x_2 - x_8)}{\sqrt{(x_1 - x_7)^2 + (x_2 - x_8)^2}}.(\upsilon x_{10} - x_4, x_3 - \upsilon x_9)\rangle$$

Hence the result.

Appendix 3:

Proof of proposition 3:

Referring to Figure 12, *r* is the initial baseline while *r'* is the baseline in the next step after a lapse of time dt. p and ϑ are the velocities of the pursuer and the evader respectively. The sudden turn with respect to the baseline at every move result in a motion by the evader in a tangential direction resulting in describing the arc 'ab' in Figure 12 by the evader for arbitrarily small dt. We then have

$$|r| = |r'| + p\,dt$$

with $p > 0$, $dt > 0$ and with $dt \to 0$, It follows that $|r'| \leq |r|$. Hence the result.

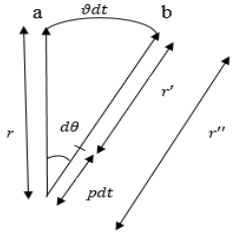


Figure. 12 Case of CM with sudden turn.

Appendix 4

Case (i):

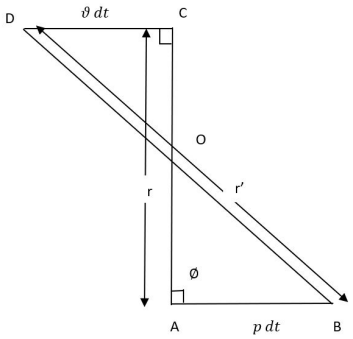


Figure. 13 Case of CB $\emptyset = \frac{\pi}{2}$ with sudden turn.

Figure 13 shows the initial baseline vector r between the pursuer (A) and evader (C). After an infinitesimally small time, with the evader taking the sudden turn of 90 degrees, let the position of evader and the pursuer be D and B respectively.

In right angle triangle AOB

$$OB > OA$$

Similarly, in right angle triangle COD

$$OD > OC$$

So,

$$OB + OD > OA + OC$$

$$|r'| > |r|$$

Case (ii):

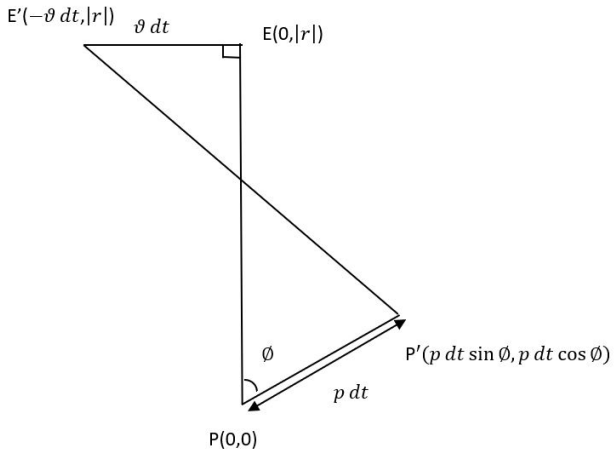


Figure. 14 Case of CB $\emptyset < \frac{\pi}{2}$ with sudden turn.

Figure 14 is redrawn from Figure 13 with $\emptyset < \frac{\pi}{2}$. Considering pursuer (P) as the origin, the coordinates of other points are easily derived. Point E denotes the evader. P' and E' are the positions of pursuer and the evader after a time dt. The baseline after time dt is given by vector

$$r' = (-\vartheta\, dt, |r|) - (pdt \sin \emptyset\,, pdt \cos \emptyset)$$

Now, to impose the condition that $|r'| = |r|$.

$$(-\vartheta\, dt - pdt \sin \emptyset)^2 + (|r| - pdt \cos \emptyset)^2 = |r|^2$$

$$\Rightarrow (\vartheta\, dt)^2 + (p\, dt)^2 (\sin \emptyset)^2 + 2\vartheta\, dt\, p\, dt \sin \emptyset + (p\, dt)^2 (\cos \emptyset)^2 - 2p\, |r| dt \cos \emptyset = 0$$

$$\Rightarrow \vartheta^2 + p^2 (\sin \emptyset)^2 + 2\vartheta p \sin \emptyset + p^2 (\cos \emptyset)^2 - \frac{2|r| p \cos \emptyset}{dt} = 0$$

$$\Rightarrow (\vartheta + p \sin \emptyset)^2 + p^2 (\cos \emptyset)^2 = \frac{2|r| p \cos \emptyset}{dt}$$

$$\Rightarrow \left(\frac{\vartheta}{p} + \sin \emptyset\right)^2 + (\cos \emptyset)^2 = \frac{2|r| \cos \emptyset}{p\, dt}$$

Hence the result.

Appendix 5:

Proof of Proposition 5:

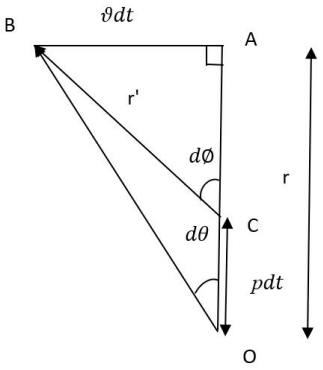


Figure. 15 Case of CP with sudden turn.

Figure 15 is similar to Figures 13 and 14 except that $\emptyset = 0$

Considering Figure 15,

$$\frac{\sin d\theta}{\sin d\emptyset} = \frac{CB}{OB} = \frac{|r'|}{\sqrt{|\vartheta\, dt|^2 + |r|^2}} \leq 1$$

$$\sqrt{|\vartheta\, dt|^2 + |r|^2} \geq |r'| \quad \forall\, dt$$

when $dt \to 0$

$$|r'| \leq |r|$$

Hence the result.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 220-226 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Brain-inspired IoT Controlled Walking Robot - Big-Foot

Anna Lekova*[,1], Ivan Chavdarov[1,2], Bozhidar Naydenov[1], Aleksandar Krastev[1], Snezhanka Kostova[1]

[1]*Bulgarian Academy of Sciences, Institute of Robotics, Str. Acad. G. Bonchev, Blk. 2, 1113, Sofia, Bulgaria*

[2]*University of Sofia, "St. Kliment Ohridski", Faculty of Mathematics and Informatics, 5 James Bourchier Blvd. 116, Sofia, Bulgaria*

ARTICLE INFO

Article history:
Received:14 April, 2019
Accepted:04 June, 2019
Online: 18 June, 2019

Keywords:
Brain-robot interaction
Walking robot
3D printing

ABSTRACT

This work presents the development of an original idea for a walking robot with a minimum number of motors, simple construction and a control system based on the brain bioelectrical activities. Described are geometric and kinematic dependencies related to the robot movement, as well as brain-inspired IoT control method. Various aspects are discussed for improving the robot's qualities, concerning the shape of the robot's feet and base in order to overcome various obstacles and maintain the static mechanical equilibrium. Improvements in the mechanical design are provided to improve reliability and enhance the scope of robot's applications. A new IoT framework for creating Human-robot interaction applications based on Node-RED "wiring" of Emotiv Brain Computer Interface (BCI) and Arduino based robot is designed, developed and tested. An educational application how to train the joint attention of children by a mind control method based on neurofeedback from beta oscillation in the right temporoparietal region is illustrated in a Node-RED flow. The neurofeedback is exposed on the walking robot.

1. Introduction

Applications of walking robots include rescue operations, work and inspection in harmful and dangerous environments, military purposes, etc. Typically, they move in an environment with obstacles whose positions and dimensions are not known in advance. Their mechanical and control systems are designed to be able to avoid or overcome obstacles. In some cases, the obstacles can dynamically change in time. This leads to more complicated designs of the walking robots compared to robots on wheels or tracks. They have more degrees of freedom and are slower. A common problem is the task of climbing and descending stairs [1-5]. Experimental robots are developed with a small number of degrees of freedom and a special shape of their feet that overcome obstacles while maintaining static stability. [7]. For these reasons, alternative solutions are investigated [5], [6] and [7]. In [7], the author presented a low-budget two-legged robot which is able to maintain static equilibrium. Other simple solutions are also investigated, such as: passive-dynamic two-legged walking [7] and different variants inspired by nature [1], [2] and [8].

Walking robots are also used for educational purposes. The results in the scientific literature show positive reactions and improvement in the attentional and positive emotional state of children with special educational needs (SEN) [9]. Nonhumanoid walking robot Big-Foot is successfully implemented to support the education of such children in two day-care centers for children with SEN in Bulgaria [10]. Some ideas have been challenged with the special educators - how the walking robot to become more intelligent and personal, in order to act as a mediator for learning and socializing because these children show deficits in early social communication skills such as Joint Attention (JA), social requesting and referencing. Joint attention is the shared focus of two individuals on an object and gaze shifting and behavioral response are the most used measures to assess the establishment of JA [11]. However, children with ASD avoid eye contact or lose focus on humans quickly. In this context, our hypothesis is that the robot has the potential to establish JA better because these children trust robots more than humans [10] and we can use this as a pure social consequence of sharing an experience. Furthermore, the JA can be made visible on a robot by exploiting a neuro physiological approach instead of gaze tracking as an indicator for a shared focus of two individuals on a same object. By integrating walking robot with a brain-aware device we propose an innovative concept for establishing and assessing JA in more objective way. For example, the Big-Foot will climb stairs only if the active brain-patterns correlated to attention system of the human brain are observed and evaluated.

*Anna Lekova, +359887435648, alekova.iser@gmail.com

The proposed nontraditional brain-inspired Robotics intervention should run anytime and anywhere with a remote supervision of the special educator over the WiFi in order to aid children not only in schools but in family environment. Therefore, the proposed framework should comply to the concept for human-robot personal communication by augmenting intelligence to robot, however not digitally by pushing a button, clicking, dragging or speaking, but biologically and continuously through emotions, mental intentions or even via chemically released by neurons rewards [12].

All this imposes technical challenges concerning robot design, proper neuroscience computing and ubiquitousness. Movement of walking robots is accomplished in two ways: motion with maintaining static stability or use of dynamic gaits. When the number of legs is small, less than 4, maintaining static stability is not an easy task as it is necessary to change the center of gravity, which requires additional degrees of freedom. Overcoming obstacles with few supports is even harder. This article discusses a robot design that has only three supports. In this case the problem of maintaining static stability is overcome by increasing the area of these supports and using suitable shapes and materials.

The technical challenges concerning ubiquitous computing are how to merge people, processes, devices and technologies with sensors and actuators. We exploited the idea behind the Internet of Things (IoT) and the innovative" Cloud Computing" infrastructure [13]. Thus, all sensing, computation, and memory can be integrated into a single standalone Socially-assistive Robotics system. Node-RED [14] is an open source development tool built by IBM, which allows to wire up IoT as nodes in flows. Node-RED is built on Node.js and can run anywhere if the applications are capable of hosting node.js, such as small single board computers like the Raspberry Pi, personal laptops or in cloud environments, such as the IBM Cloud. The Node-RED connectivity allows nodes to collect and exchange data ubiquitously and its flow-based programming is an ideal solution to wire up the biological brain intelligence to robots anytime and anywhere. Based on the idea behind IoT, that uniquely addressable "things" communicate with each other and transfer data over the existing network protocols, we propose how the information channel between the human brain and external devices to be applied for IoT brain-to-robot control. By analogy to Visual servoing [15] and Tactile servoing [16], we defined a term "Brain servoing" (or brain-based robot control) that uses EEG feedback information extracted from the brain EEG sensor to control the motion of a robot. The control tasks intend to translate a specific brain activity interaction patterns in robot commands ubiquitously. The control instructions are transferred from continuously decoded JA performance metric into robot commands and are sent to the robot actuators via a Node-RED set-up Emotiv Brain-Computer Interface (BCI) to Arduino.

In this study we illustrate a non-traditional control method where the brain electrical activity is captured by EMOTIV brain-listening headset [17] and specific spatial and temporal brain frequencies correlated to JA are translated into commands to control the walking robot Big-Foot. Because the robot Big-Foot has a simple and innovative design, children with specific needs find it attractive, and it does not create feelings of anxiety and discomfort when interacting with it. This helps the robot function as a mediator between the children and the therapist. The 3D model of the robot is cheap and easy to be controlled by children. It is extremely maneuverable and can climb stairs. These features allow it to be used in educational games for children. To the best of our knowledge we first propose an IoT framework for creating Human Robot Interaction (HRI) applications based on Node-RED "wiring" of Emotiv BCI [18] and Arduino based robot.

2. Mechanical design and improvement of the robot

Fig. 1 presents the design of the robot and Table 1 lists its main components. Two engines are mounted in the body 5. The rotor of motor 2 is connected to and rotates the circular base 1. This allows the robot to change its orientation and turn when the feet 4 are raised above the ground. The motor 3, via a connecting shaft, drives the arms 7. At the end of the arms 7, the feet 4 are mounted. The feet maintain a constant orientation with respect to the base 1 and the body 5 by means of two gear mechanisms 6 with a gear ratio of $i=1$. This design allows the robot to rotate more than 360 degrees around axis $R1$. There is also no limitation of the rotation around axis $R2$. The two rotations are reversible, making the robot extremely maneuverable.

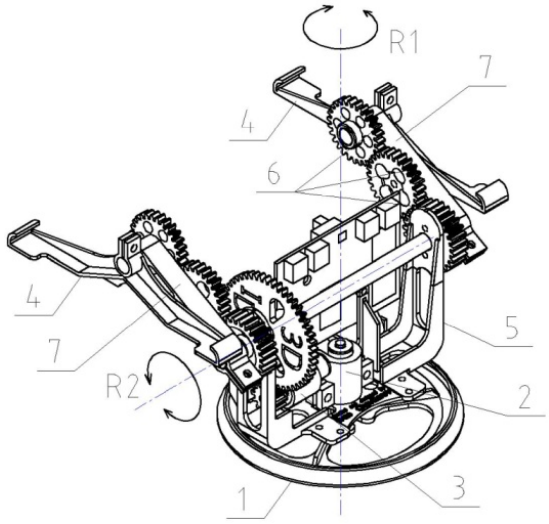


Figure 1: 3D model of the robot Big-Foot

Table 1. Part list of the main components

Position	Part description
1	Circular base
2	DC motor for rotating the base
3	DC motor for rotation of the arms
4	Feet
5	Body of the robot
6	Gear transmission mechanisms
7	Rotating arms

When the robot moves on a flat terrain it passes through two main phases. During the first phase, the feet 4 are stationary on the ground. The arms 7 move, and the body 5, together with the base 1, is moved at a distance of one step S (Fig. 2a). All points on the robot's body move along trajectories that represent arcs of circles, for example, the trajectories of points B and P on Fig. 2. The radius r_{AB} is determined by the length AB of the arms 7. During this phase, the arms 7 are rotated at an angle α

Similarly, to the first phase, the trajectories of the points of the feet are also moving along arcs of a circle with radius r_{AB}. In this case, the angle of rotation of the arms 7 is $(2\pi - \alpha)$ (Fig. 2b). During this phase, the robot does not move forward but can rotate around axis *R1* (Fig. 1). The second phase ends when the feet touch the ground. The movements during the two phases are cyclically repetitive.

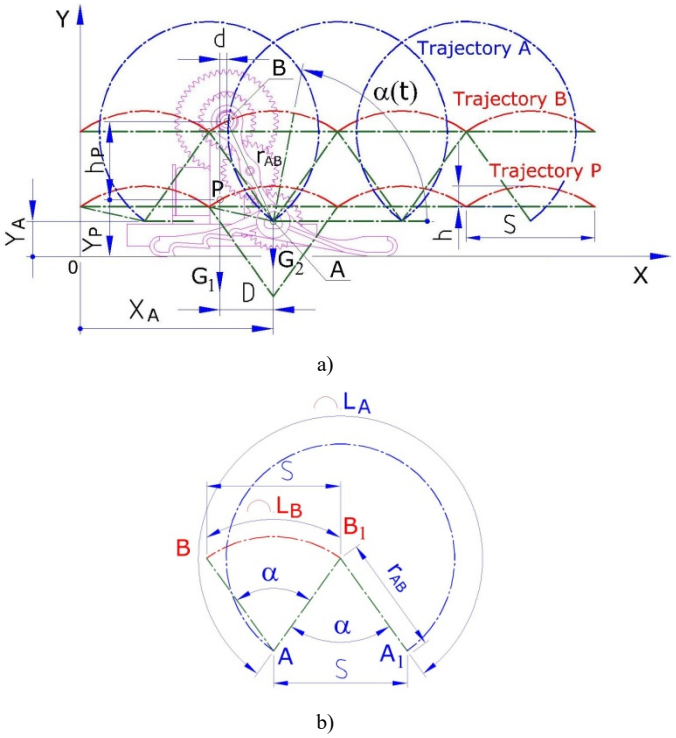
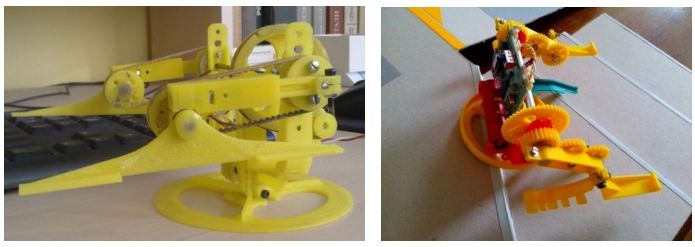


Figure 2: Trajectories of different points of the robot during the two phases of motion - a) and b) step S and angles of rotation of arm AB.

During the second phase, the robot's feet rise to a relatively high height, which helps to overcome high obstacles. The behavior of the robot has been verified through computer simulation and experiments with 3D printed prototypes.

More detailed descriptions of the kinematics and behavior of the robot in the various phases of its movement are described in [8], [19] and [20].

In the original idea described in patent [19], maintaining of parallelism between the feet and the base is achieved by a belt mechanism. The 3D printed prototype (Fig. 3a) showed that the belt drive is not suitable due to the need for considerable tensioning. The drawbacks of this model, described in more details in [20], have led to the need of improvement of the mechanical design.



a) b)

Figure 3: a) First 3D printed model, b) model with improved design climbing stairs.

Improvements have been made to the mechanical structure for a more reliable torque transmission. A new type of joint between the shaft and the feet is used (Fig. 4, position 8), which significantly reduces the stress concentration. This joint is realized by a smooth transition from a polygon to a circle, and the 3D printing technology allows its physical implementation and application in our model [20], [29].

Changing the shape of the feet and the round base makes it possible to overcome higher obstacles (Fig. 4, positions 9 and 10).

Another approach is used to improve the capabilities of the robot, namely to add additional elements. Two tails have been added (Fig. 5, position 11), which prevent the robot from rolling over when overcoming high obstacles.

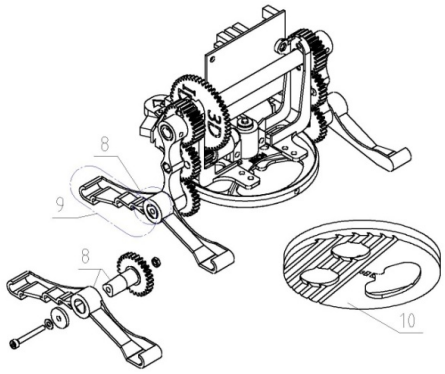


Figure 4: Improvements in the design of the Big-Foot robot

The addition a platform (Fig. 5, position 12) is used for applications of the robot for developing games for children with specific needs.

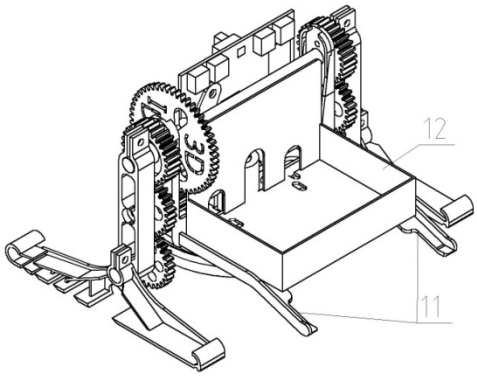


Figure 5: Walking robot with tails and a paltform.

3. Human mind control method of robot with Node-RED EmotiveBCI-to-Arduino interface

The most common way to control a robot is the joystick or mobile device. However, for some people the joystick or mobile control is difficult or impossible for many reasons and other (non-traditional) control methods have being developed based on the recent innovative sensors and technologies, such as motion sensing devices, etc. In this section, we present a non-traditional control method that rely on "Brain servoing" which uses feedback information extracted from a brain sensor (EEG feedback) to control the walking robot.

A traditional bi-directional DC control of two motors (operating at a voltage of 2.7-10.8V) is used. The current is about 1.2A per motor and the load on motor is up to 2A for few seconds. The H-bridge DC motor control is used in order to rotate motors in both directions. Four transistors in the circuit are controlled in pairs and play a role of switches to control the motor to rotate in both directions. Fig. 6 presents the connecting scheme and Fig. 7 the code uploaded on the ARDUINO MICRO. The end-user interfaces via dialog box for laptop or from mobile device are presented in Fig. 8.

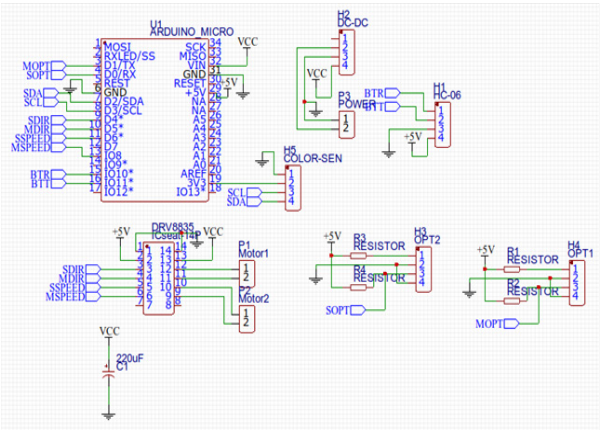


Figure 6: Connecting scheme

```
bigFoot2 | Arduino 1.8.8 (Windows Store 1.8.19.0)
File  Edit  Sketch  Tools  Help
bigFoot2 §

#define MOTOR_1_L 3
#define MOTOR_1_R 4
#define MOTOR_2_L 10
#define MOTOR_2_R 11
char  temp;
void motor_drive(byte stat_m1_r, byte stat_m1_l, byte stat_m2_f,
  // SET 1 or 0 to pins -> motors
  // L and R
  digitalWrite(MOTOR_1_R, stat_m1_l);
  digitalWrite(MOTOR_1_L, stat_m1_r);
  // F and B
  digitalWrite(MOTOR_2_R, stat_m2_f);
  digitalWrite(MOTOR_2_L, stat_m2_b);
void setup() {
 Serial.begin(9600);
  pinMode(MOTOR_1_L , OUTPUT);  pinMode(MOTOR_1_R , OUTPUT);
  pinMode(MOTOR_2_L , OUTPUT);  pinMode(MOTOR_2_R , OUTPUT);
}
void loop() {
if(Serial.available() > 0) {
        temp = Serial.read();
switch (temp) {
  case 'F':
    motor_drive(0, 0, 1, 0);
```

Figure 7: The code for ARDUINO MICRO

The nontraditional control method intends to translate in robot commands a specific brain activity interaction patterns within a framework of IoT. Fig. 9 illustrates in general this framework for Brain-robot interaction and its deployment using a Node-RED tool. An application of EmotivBCI-to-Arduino Node-RED flow is explained in Section V. The commercial EEG devices, Emotiv EPOC or Emotiv Insight [17] could be used. The placement of electrodes are shown in Fig.10. In this study EEG data are acquired and recorded with the 14-channel neuroheadset EPOC+ and the EEG signals were sampled at the rate of 256 Hz. The data are wirelessly transmitted to a host computer through Bluetooth and further processed with EmotivBCI Node-RED Toolbox [18].

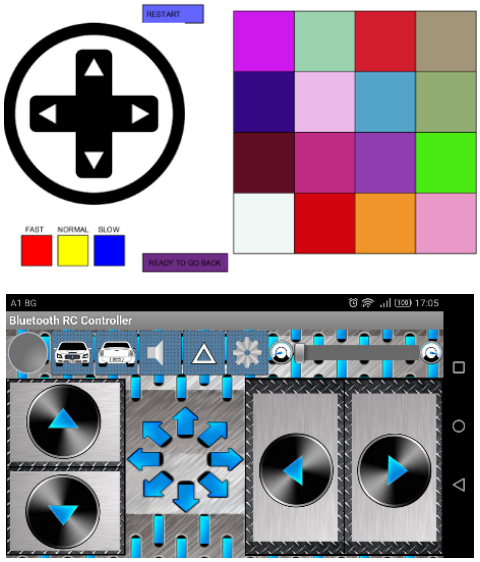


Figure 8: Traditional control methods of walking robot BigFoot

Figure 9: IoT framework for Brain-robot interaction - Node-RED EmotivBCI-to-Arduino Interface.

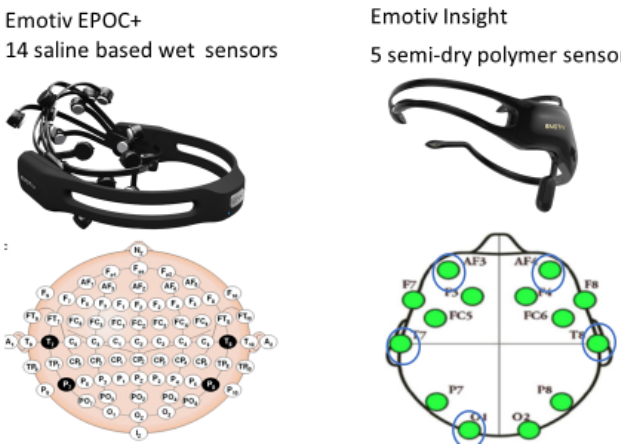


Figure 10: EMOTIV EEG headsets for EEG-based BCI - listen, record and transmit in real time the electrical activity of the brain

Node-RED uses a visual programming approach for 'wiring together' of code blocks and make up 'flows' to carry out tasks. It connects nodes as a combination of input nodes, processing nodes and output nodes in a browser-based flow editor using a wide range of nodes in the palette. The EmotivBCI Node-RED Toolbox is a custom library of input nodes for Node-RED which allow interfacing the EMOTIV technology with other Node-RED nodes and thus to create a wide variety of BCI integrations. Installation, node descriptions and use are presented in [18]. EmotivBCI Node-RED gets data from Emotiv Cortex (Cortex is built on JSON) and WebSockets for creating BCI applications and integrates data streams from the human headset with third party software or hardware. In a browser-based flow in a Web page, a user is

allowing to control the BigFoot rotation and direction of both motors using one of the following brain activity correlated patterns: in the frequency domain, correlated with human cognitive, facial and emotional states or directly by preliminary trained mental commands. EMOTIV offers the opportunity for the user to create and execute a number of Mental Commands [21], such as: push, pull, lift, drop, left, right, rotate left, rotate right, rotate forwards, rotate backwards, rotate clockwise, rotate anticlockwise, and disappear. The detected facial expressions are blink, left wink, right wink, raised eyebrows (surprise), furrowed brows (frown), smile and clenched teeth. The Emotiv technology currently offers five performance metrics detections: Engagement/Boredom, Frustration, Meditation, Instantaneous Excitement, and Long-Term Excitement, that are based on universal in nature brainwave characteristics and don't require an explicit training of the user [22].

All these correlated frequency patterns or detections are interpreted in the Node-RED function blocks and map to commands in order to control the walking robot by sending dynamically information to the serial port where the Arduino is connected. There are several ways to interact with an Arduino using Node-RED. As the Arduino appears as a Serial device, the Serial in/out nodes can be used to communicate with it after adjusting the serial port speed (baud rate) to be the same at both ends. We wired the EmotivBCI nodes to serial port and Arduino Software (IDE). We program the Arduino with the IDE, and then send and receive input over the serial port to interact with BigFoot. Two browser-based flows in a Web page, as well as Node-RED Graphical User Interface (GUI) are designed to allow the user to control the BigFoot motors dynamically by its own brain activity patterns in the frequency domain corresponding to the joint attention. Every 1 or 2 s a string for one of the following commands - 'S' for stop, 'F' for forward, 'B' for backward, 'L' for left rotation and 'R' for right rotation is sent. Except the neurofeedback exposed on the BigFoot, the GUI shows the rotation and direction as a feedback in the web page.

4. Application of the IoT framework for brain- robot control method

The walking robot is used in the development of educational play-like activities of children with SEN. Trough playing with the robot, the children improve their special orientation abilities, they learn easily and with fun by controlling the robot [11]. The interaction of children with the walking robot engages them to communicate with each other and develop their joint attention (JA). By using the non-traditional mind control methods, a neurofeedback rehabilitation is possible and will be effective for training the attention or emotion self-regulation of the brain function. We place the child's neurofeedback in the play-with-robot interventions and expose it on the walking robot.

Joint attention precedes the development of children mentalization skills [23]. Mentalization is the ability to understand the mental state, of oneself or others, that underlies evident behavior. Brain activity patterns in time and frequency domains correlated with JA are discussed in many papers [23], [24] and [25]. According to studies [25] and [26] the oscillatory brain activity in the alpha and beta ranges in the right temporoparietal region correlates with the anticipation and prediction of another person's responses and preferences. Authors in [23] have tested whether neuronal activity preceding JA correlates with mentalization in typically developing (TD) children and whether this activity is impaired in children with autistic spectrum disorder (ASD) who evidence deficits in JA and mentalization skills. TD children shown beta rhythm (15-25 Hz) in the temporoparietal region preceding the JA behavior, while ASD children did not show an increase in beta activity. In the study [23] statically significant difference in increasing the beta band power in the right parietal group of channels is found and again the data analysis suggested that the right temporoparietal region and the middle/superior frontal gyrus are the main brain regions contributing to the beta power differences between the two groups. Based on these neuroimaging findings we designed a play where a child can move the BigFoot to climb the stairs through a neurofeedback from the oscillatory brain activity in the alpha and beta ranges in the right temporoparietal region. Thus, we try to teach and train children joint attention by learning to modulate their rhythm power in the beta frequency band [15-25 Hz] in order to move the walking robot. The power spectrum of right temporoparietal lobe intensity is obtained by EmotivBCI tool and the changes in JA are calculated based on functional brain imaging in terms of event-related desynchronization (ERD) or event-related synchronization (ERS) [26].

According to Pfurtscheller [26] sensory and cognitive processing results in changes of the ongoing EEG in form of ERD/ERS that are highly frequency-band specific. For example, oscillations with 10 Hz comprise more synchronized neurons than oscillations with 40 Hz. The ERD is interpreted as a correlate of an activated cortical area with increased excitability and the ERS in the alpha and lower beta bands can be interpreted as a correlate of a deactivated cortical area [26]. Furthermore, the frequency of brain oscillations depends on the percentage of a population of neurons synchronized. With an increasing number of synchronized neurons, the average frequency becomes slower and if only 10% are synchronized, the amplitude is 10-fold the activity of the 90% of not-synchronized neurons.

In terms of information theory, a desynchronized system represents a state of maximal readiness and a maximum of information capacity [27]. The neuroscience explanation is the ERD, i.e. in the underlying neural network small areas of neurons involved in a particular neural computation (neural ensemble), work in a relative independent or desynchronized manner.

We use (8) to obtain the percentual decease (ERD) or increase (ERS) in the band power during a test (activation) interval compared with a baseline (reference) interval. The ERD/ERS index foe alpha and beta bands are the respond to different levels of JA:

$$ERD/ERS\% = (A-R) / R*100 \quad (8)$$

where A is the power within the frequency band of interest in the activity period and R is the preceding baseline or reference period. Positive numbers are obtained for ERS% and negative - for ERD% that reflect synchronization and a state of band power decrease. We tested our research hypothesis in pilot experiments with five students (right handed male and female in average age about 18 years old) by measuring whether the amount of alpha ERD increase and the amount of beta ERS increase with higher level of

JA and mentalization skills. The mean threshold of 5% for ERS in the beta rhythm was defined experimentally. We plan to perform the experiment explained in [23] with children with ASD in order to compare the establishment and assessment of JA based on the intervention and observations of the gaze shifting and behavioral response from practitioners versus establishment and assessment of JA by the brain-inspired robot control and a neurofeedback response.

The control commands are transferred from continuously decoded JA performance metric into robot commands and are sent to the robot actuators via Node-RED EmotivBCI-to-Arduino Interface. For this application, only the right temporal T8 and right parietal P8 electrodes are used for right-handed participants and their beta and alpha power are analyzed during the neutral or joint attention states. The power spectrum density for beta band is received continually from the EMOTIV Node-RED EEG streaming data. The command 'S' is sending to COM4 by default. In a function node, the ERS and ERD are calculated for each electrode and when alpha ERD significantly decrease and beta ERS increase (thresholds in %, specific for the participant), the corresponding command 'F' or 'B' is sent to COM4. The results showed that robot could be successfully navigated by the positive values of 5% for ERS in the beta rhythm, which is a correlate with joint attention on walking robot in order to complete a navigation task to climb stairs by human intention. In the future we will assign the percentual increase into fuzzy sets in order to define levels of difficulty for climbing stairs.

The Node-red flow for brain-robot interaction based on EmotivBCI toolbox is shown in Fig.11. After streaming the row data from the Emotiv node in the first Node-RED flow, the power of frequency bands of interest are stored in global variables. They are accessed in the second flow (Fig.12) to set-up the baseline settings in the frequency band of interest for 1 or 2 min. How the robot commands are mapped and sent to serial port can be seen in the first Node-RED flow.

We intend to test the proposed framework with a neurofeedback from the frontal theta for training and remembering the orientation in space and implicitly to expose it by the walking robot. The increase in theta power during successful encoding of new information is discussed from neurological point of view in [28] with the relationship with hippocampal theta induced in the cortex.

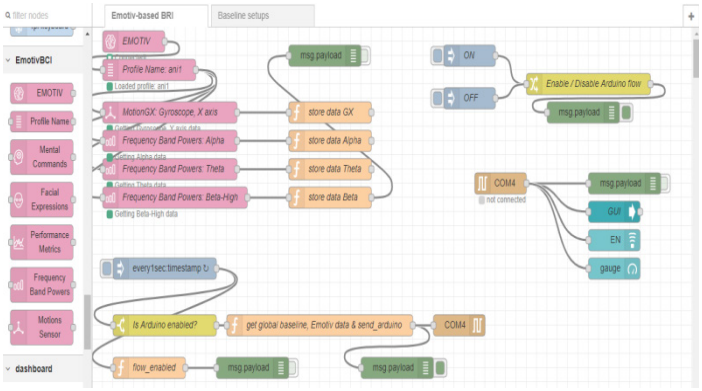


Figure 11: Node-red flow for EmotivBCI-Arduino interaction

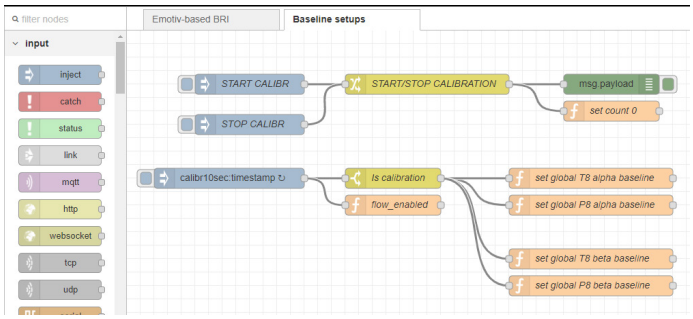


Figure 12: Node-red flow for baseline setting in the frequency band of interest

5. Conclusion

Although the proposed walking robot has a minimum number of mechanical components, it is extremely maneuverable. It has only two motors that makes it simple to control and suitable for educational or rehabilitation purposes. The robot is not expensive and easy to manufacture with a 3D printer. Several control methods are proposed, one of them is "Brain servoing" that uses EEG feedback and brain activity patterns for mind-based robot control. The proposed IoT framework for creating Human Robot Interaction (HRI) applications based on Node-RED "wiring" of Emotiv BCI and Arduino based robot has been applied for neurofeedback training. The results showed that robot could be successfully navigated by the level of human attention to complete a navigation task for climbing stairs.

In the future, the robot's design and control system will be enhanced by collecting and recording information from more sensors. Since, the proposed framework for mind-robot control is enough general, it will be easily applied for other areas of neurons involved in a particular neural computation, as well as other brain sensors and different humanoid or nonhumanoid robots in the IoT.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgment

This research is supported by the National Scientific Research Fund, Project N DH17/10. We thank to all high school students from the "National Professional High School for Computer Technologies and Systems", Pravets, Bulgaria, who participated in the experiments and implementation of the Arduino hardware.

References

[1] Ben-Tzvi P., Ito S., Goldenberg A.A. (2009), A mobile robot with autonomous climbing and descending of stairs, Robotica 27(2), pp. 171-188, https://pdfs.semanticscholar.org/0550/5591df73849ecba97e8754d518eda1bead41.pdf

[2] Poramate Manoonpong, Sakyasingha Dasgupta, Dennis Goldschmidt, Florentin Wörgötter, Reservoir-based online adaptive forward models with neural control for complex locomotion in a hexapod robot, 2014, International Joint Conference on Neural Networks (IJCNN), Pages: 3295 – 3302.

[3] Liu Juan Xiu, Chen Qing Wei, Dynamics and control study of a stair-climbing walking aid robot, 2010 International Conference on Mechanic Automation and Control Engineering, Pages: 6190 – 6194.

[4] Chih-Hsing Liu, Meng-Hsien Lin, Ying-Chia Huang, Tzu-Yang Pai, Chiu-Min Wang, The development of a multi-legged robot using eight-bar linkages as leg mechanisms with switchable modes for walking and stair climbing,

2017, 3rd International Conference on Control, Automation and Robotics (ICCAR), Pages: 103 – 108.
[5] B. Jeon, Y. Jeong, K. Kwak, S. Yeo, D. Ha, S. Kim (2009), Bio-Mimetic Articulated Mobile Robot overcoming stairs by using a slinky moving mechanism, Proceedings of ICAD2009, The Fifth International Conference on Axiomatic Design, Campus de Caparica, March 25-27, 2009, pp.173-179.
[6] Fei Sun, He Hua Ju, Ping Yuan Cui, A new 12 DOF biped robot's mechanical design and kinematic analysis, Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, 2011, Pages: 2396 – 2400.
[7] N. E. Nava Rodriguez* G. Carbone† M. Ceccarelli Design Evolution of low-cost humanoid robot CALUMA, 12th IFToMM World Congress, Besançon (France), June18-21, 2007 University Cassino, Italy, https://pdfs.semanticscholar.org/bde8/0efbb0328483ff611f82ad035981ab3a133e.pdf
[8] Chavdarov I. , Walking robot realized through 3d printing, Comptes rendus de l'Acad´emie bulgare des Sciences, Tome 69, No 8, 2016, pp. 1069-1076, http://www.proceedings.bas.bg/cgi-bin/mitko/0DOC_abs.pl?2016_8_13
[9] Kruse T., Pandey A., Alami R., Kirsch A.. Human-Aware Robot Navigation: A Survey. Robotics and Autonomous Systems, Elsevier, 2013, 61 (12), pp.1726-1743. hal-01684295
[10] Weisberg H., Jones E. Individualizing Intervention to Teach Joint Attention, Requesting, and Social Referencing to Children with Autism. Behavior Analysis in Practice (2019) 12:105–123
[11] Dimitrova, M., Lekova, A., Kostova, S., Roumenin, C., Cherneva, M., Krastev, A., Chavdarov, I. (2016) A multi-domain approach to design of CPS in special education: Issues of Evaluation and Adaptation. Proceedings of the 5th Workshop of the MPM4CPS COST Action, November 24-25, 2016, Malaga, Spain, pp.196-205.
[12] Lekova A., Pavlov V., Chavdarov I., Krastev A., Datchkinov P., Stoyanov I.. Augmented Intelligence For Teaching Robots By Imitation, International Scientific Journal "Industry 4.0" Web Issn 2534-997x; Print Issn 2543-8582, Year Ii, Issue 5, P.P. 201-204 (2017), pp. 201-204
[13] Q. Zhang, L. Cheng, and R. Boutaba, ``Cloud computing: State-of-theartand research challenges," J. Internet Services Appl., vol. 1, no. 1,pp. 718, 2010.
[14] IBM NodeRED Flow-based programming for the Internet of Things. URL: https://nodered.org/
[15] Agin, G.J., "Real Time Control of a Robot with a Mobile Camera". Technical Note 179, SRI International, Feb. 1979
[16] Q. Li, C. Sch¨urmann, R. Haschke, and H. Ritter, "A control framework for tactile servoing," in Robotics: Science and Systems, 2013
[17] Martinez-Leon, J., Cano-Izquierdo, J. & Ibarrola, J. Are low cost Brain Computer Interface headsets ready for motor imagery applications? Expert Syst Appl 49, 136 (2016).
[18] EmotivBCI Node-RED Toolbox.https://emotiv.gitbook.io/emotivbci-node-red-toolbox/node-descriptions-and-use
[19] I. Chavdarov, T. Tanev, V. Pavlov Patent application 111362. „Walking Robot", Published summary - Billetin No. 6, 30.06.2014., page 11, in Bulgarian, http://www.bpo.bg/images/stories/buletini/binder-2014-06.pdf
[20] I. Chavdarov, B. Naydenov, S. Kostova, A. Krastev, A. Lekova, "Development and Applications of a 3D Printed Walking Robot - Big-Foot", 2018 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), https://ieeexplore.ieee.org/document/8555843/references#references
[21] Create and execute a number of Mental Commands in EMOTIV. URL: https://emotiv.zendesk.com/hc/en-us/articles/201216335-Training-Mental-Commands
[22] Understanding the Performance Metrics Detection Suite. URL: https://emotiv.zendesk.com/hc/en-us/articles/201444095-Understanding-the-Performance-Metrics-Detection-Suite
[23] Soto-Icaza P., Vargas L., Aboitiz F., Billeke P, Beta oscillations precede joint attention and correlate with mentalization in typical development and autism, Cortex, Volume 113, 2019, Pages 210-228
[24] Billeke, P., Zamorano, F., Cosmelli, D., & Aboitiz, F. (2013). Oscillatory brain activity correlates with risk perception and predicts social decisions. Cerebral Cortex (New York N.Y. 1991), 23(12), pp. 2872-2883
[25] Park, J., Kim, H., Sohn, J.-W., Choi, J., & Kim, S.-P. EEG beta oscillations in the temporoparietal area related to the accuracy in estimating others' preference. Frontiers in Human Neuroscience, 12(February). 2018
[26] Gert Pfurtscheller, F. H. Lopes da Silva. Functional Brain Imaging. Hans Huber Publishers, 1988, 264 pages.
[27] Thatcher, R. W., McAlaster, R., Lester, M. L., Horst, R. L., & Cantor, D. S. (1983). Hemispheric EEG asymmetries related to cognitive functioning in children. In A. Perecuman, Cognitive Processing in the Right Hemisphere (pp. 125–145). New York: Academic Press.
[28] Klimesch W, Doppelmayr M, Russegger H, Pachinger T. Theta band power in the human scalp EEG and the encoding of new information. NeuroReport 1996;7:1235-1240.
[29] I. Chavdarov, Patent application № 112346, "3D printed joint between a shaft and a link", Published summary - Bulletin 01.2 31.01.2018., page 13-14, in Bulgarian http://www.bpo.bg/images/stories/buletini/binder-2018-01-31.pdf

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 227-234 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Multi-Band Hand of God Antenna with Functionality Variation with Use of Slots and Feed Position

Mervat Akram Madi*, Karim Youssef Kabalan

Electrical and Computer Engineering Department, American University of Beirut, 1107 2020, Lebanon

ARTICLE INFO

Article history:
Received:02 May, 2019
Accepted:25 May, 2019
Online: 18 June, 2019

Keywords:
Multiband antenna
Slotted antenna
Patch Antenna

ABSTRACT

This article investigates different versions of a new multi-band micro-strip antenna design. NASA inspired the antenna design from "Hand of God" galaxy discovery. A multi-band operation between 0.15 and 8 [GHz] is obtained and captures up to 10 frequency bands. The antenna has a small area of 4 x 5 cm^2 and is printed on 1.6 mm epoxy-FR4 substrate. Adding slots and changing the feed position, led to different designs and increase in multi-band operation. Moreover, varactors are mounted on antenna surface, which resulted in frequencies as low as 50 MHz and as high as 8 GHz when. S11 simulations results were successfully compared with measured ones. Applications include different mobile generations, TV and different IoT systems.

1. Introduction

Multi-band antennas capture various segments of the frequency spectrum simultaneously. Yet, a challenging aspect of maintaing the required gain appears [1-6]. Of course, an antenna should have high efficiency to be able to capture many frequencies. Moreover, new technology requires smaller front end devices, and hence antenna size follows. Moreoever, selected frequencies are targetted, to cope with licenced frequency bands. Hence, research is still developing in this area.

Multiple frequency operation has become recently desired to engage IoT applications in same device. It allows performing technologies that require and or scan wide range of frequencies. For example, it is desirable to work with two mobile generations simultaneously [7]. Adapting to wearable technologies, [8], presented a relatively large antenna operating for different applications with up to six frequency bands. Alternatively, different bands related to same generation can be combined in same antenna [9]. Moreover, switches were mounted to achieve multiple resonances, while changing the switching scenario changes operation band being the L or S band [10,11]. Metamaterials were also exploited in [12] for multi-band operation.

The need to minimize space occupied by modern devices such as phones or sensors renders multi-band antennas more attractive since they replace many antennas [10], while still consistently frequencies in MHz range are mainly emitted by large antennas that are quarter wavelength, i.e. more than 12 cm in length and width [4,11,13,14]. This creates a challenge if high frequency ratio is targeted within same antenna. supporting different active parts of wireless spectrum. Low

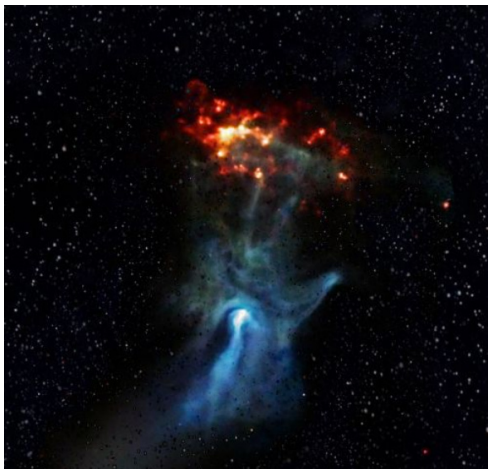


Figure 1: Galaxy named 'Hand of God' by NASA

This paper suggests a novel design that covers multiple low and high communications frequency bands, serves IoT applications, and maintains a compact size. The design is inspired by a photo of a galaxy captured recently by NASA, entitled 'Hand of God' [15]. Hence, this antenna duplicates a cosmic shape using

* Mervat Akram Madi, mm102@aub.edu.lb

electromagnetic radiation modelling, in an articulate way. Ultimately, the antenna presented here, resonates at frequencies near 0.1 [GHz] up to 8 [GHz], knowing that different techniques are used to enhance these aspects.

First, as power input position is varied, operating frequencies shifted. Hence, position of feed is an important parameter that is thoroughly investigated in the paper. Second, slots were easily incorporated in the design and enhanced increase of frequency operation in addition to bandwidth ratio. Third, varactors are added across slots to study their effect on frequency reconfigurability. As a result, frequency is tuned in the MHz range and resonance at 100 MHz is achieved, in spite of the very small physical length of the antenna.

The design of multiple extensions given by a hand shape is elaborated in this paper, since naturally each frequency is linked to a radiating element. Yet, the operation bands obtained outnumber the dipole, finger like extensions.

The antenna has compact volume of 4 x 5 x 0.16 cm^3, a new design that succeeds in combining more than six bands and captures ten bands in one specific scenario, with high frequency ratio of 53 starting at 0.15 [GHz], if no diodes are used. Frequencies above 8 [GHz] were not included in this research. 3G, 4G and IoT frequency applications are still main concern in antenna design. Authors present the 'Hand of God' for the first time, novelty is achieved in design, size and range of application [16]. In following sections, the antenna design significance is shown FEM modeling in HFSS. Third section shows the simulations results for feed through one chosen position. Section Four details all interesting scenarios in feeding positions and displays a comparison in S11 (return loss) results between measurements and simulations for the various fabricated prototypes. In section five, the effect of mounting varactors while DC voltage control is applied is shown. Sections six and seven show the current density and the corresponding radiation patterns for the resonant frequencies relative to the first discussed prototype.

2. Antenna Design Methodology

Due the importance of monopole, L and U shaped antennas in microstrip antenna design, the hand design holds importance in relation of design and shape. As the first draft of the patch design had shown multiple band characteristic, parametric sweeps that modify the shape for better resonance were done in FEM software tool HFSS. In fact, there is a conjecture that physical shapes in nature cope well with a certain frequency, and these shapes manifest certain cosmic frequencies, which is literally a direct link between shape and frequency. In fact, both ends of physics and metaphysics meet at the antenna design platform.

Moreover, the effective relation between having multiple extensions from a basic feed to a multiple frequency operation has immediately inspired me to recreate the hand shape on the patch after seeing the NASA photographs of a far galaxy. Yet, the use of FEM calculation tool for antenna design has facilitated turning the antenna conceptual design into an operating one. Here, experience in antenna design and persistence in using the HFSS software to explore all parametric influences on their own or in relation to other parameters has led to good results and even unexpectedly 9 resonant bands when slots were used.

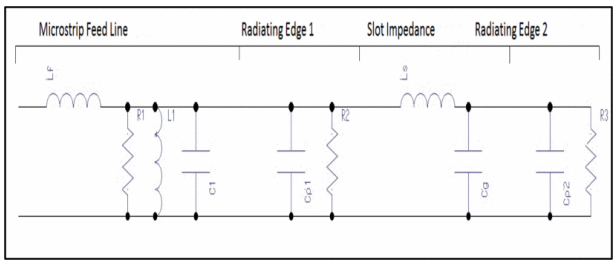


Figure 2: Equivalent circuit of slot loaded antenna

The 'Hand of God' [15], is source of inspiration of this design due to the photo captured by NASA telescope. The finger shapes and hand is a renovated form of dipole antenna array, with main 'palm' space acting as a relatively large capacitance that stabilizes antenna radiation operation at large frequencies in GHz range. This is no the first time a cosmic/natural shape is successfully exploited in electromagnetics, examples include the spiral shape (inspired from a galaxy [17]) and many others showing the beauty of science in relation to nature.

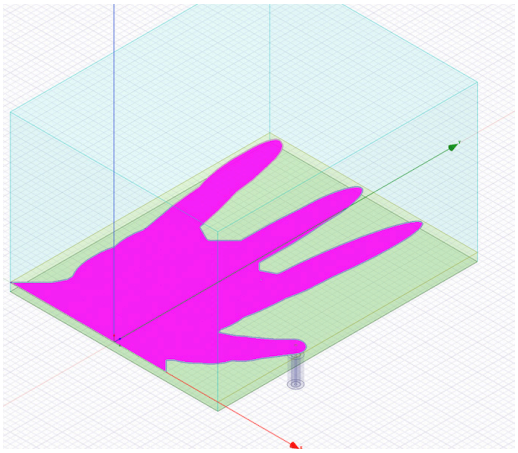


Figure 3: Simulated feed at 1st fingertip

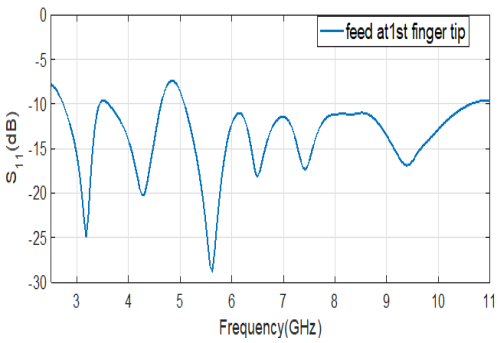


Figure 4: S11 simulated results of feed at 1st fingertip in Figure 3.

The antenna with four fingers combines inductive and capacitive effect of each extension. Yet, it forms a new way of

combining four dipole antenna extensions in a compact design. These were responsible for multiple frequency operation. In particular, each element has different length and shape, and spaces between elements play crucial role in minimizing destructive interference, since targeted frequencies have wavelengths comparable to antenna size. After performing many simulations on common area in the patch and length of the extensions, final design was achieved as shown in Figure 3. Initial setting of the model was even varied in total area and substrate material, but a double-sided FR4 substrate was most successful. Finally, small area of 4 x 5 cm ^{2} accommodated the multiple 'finger like' extensions to achieve the targeted multi-band operation.

Figure 5: First feed scenario prototype at 4th fingertip.

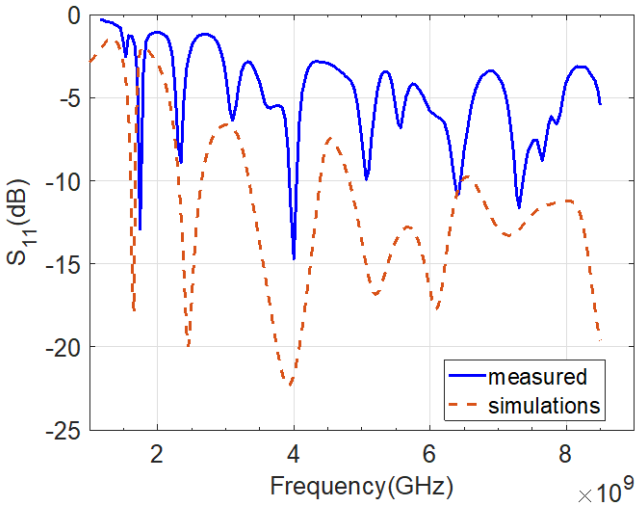


Figure 6: Comparison of S11 obtained in measurements and simulations results for 4th prototype (Figure 5).

Slots are used to increase bandwidth. Loading the radiating patch with slots or stub elements can extend bandwidth of the patch antenna. Additionally, varying the dimension of cutting slots on different locations and adjusting the shape of the slots like rectangular, circular, octagonal, fractal, and modified/partial ground plane, can help to obtain the desired bandwidth at estimated resonant frequency [18]. Slots are also used in size miniaturization of antenna [19] and increasing number of resonant frequencies [20].

Slot impedance is considered as series inductance and parallel capacitance, which adds to the total inductance and capacitance as shown in Figure 2, [21]. In fact, slots helped increase the number of resonant frequency bands over [0.1-8] GHz to nine with tunability over frequency range [0.2-0.8] GHz, and total area = 4 x 5 cm ^{2}. In a similar approach, [22] has used 3 slots to achieve three resonant bands tunable at 3 frequencies over [0.6-2.7] GHz, and an area = 5 x 10 cm^2. Slots with modes controlling method were used in [23]. Starting with a monopole at 0.25 λ then adding an open slot at 0.25 λ then a tuning pad, resulted in 2 bands [0.8-1.02] GHz and [1.69-2.73] GHz, knowing that the area = 10 x 6 cm^2. Multiple slits on a rectangular patch were also used in [20], to achieve 4 bands over [5.8-7.87] GHz and antenna area of 2.6 x 7.75 cm^2. Other methods can be listed here with corresponding characteristics, for example using multiport and theory of characteristic modes (TCM) to isolate frequency bands, resulting in 3 bands over [0.8-2.2] GHz [24]. In comparison, [25] mounted a 3D loop folded antenna (volume 60×5×5 mm^3) on 100×60×0.8 mm^3 FR4 substrate with full ground plane, for a mobile handset, to design a 3 bands tunable antenna over [0.8-2.5] GHz. Consequently, the present method is better improved over other methods due to carving the slots on the antenna patch along the extended finger shapes, which enhanced the resonant characteristic modes by doubling the mode effect being a result of splitting into two arms in a U shape.

3. Simulations basic design and effect of feed position

The antenna is a microstrip patch, with full ground, an area of 4 x 5 cm ^{2} and 1.6 mm thick epoxy FR4 substrate. The dimensions of finger shapes (see Figure 5) are 25.8 mm x 6 mm for 4th finger, 19.6 x 6 mm for 3rd finger , 21.5 mm x 6 mm for 2nd finger , 14 mm x 9.5 mm for 1st finger. The antenna feed is a coaxial cable as shown in Figure 3, where axes x, y, and z are in red, green, and blue respectively for later reference. In fabrication, an SMA connects the cable to the antenna at the variable positions to be discussed, where 50-Ω matching is considered.

Figure 7: Antenna prototype with the feed at the 2nd fingertip.

The design is modeled in FEM software, HFSS. In order to preserve the universal shape, and improve results and easily accommodate the shape in small area for better performance, simulations were the main guide to refine the design, where there

is no imperial formula to apply in this case. The underlying design

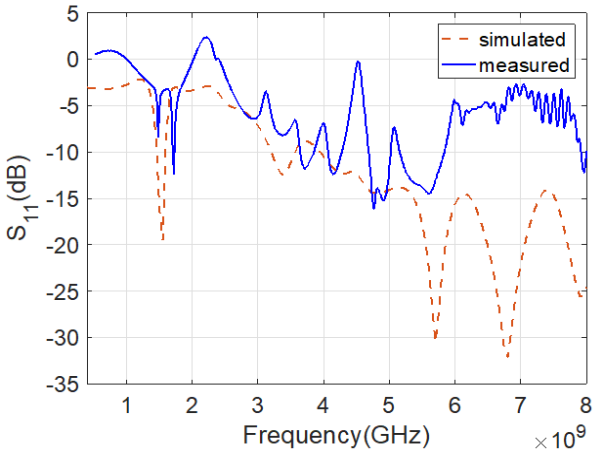


Figure 8: Comparison of S11 obtained in measurements and simulations results for 2nd prototype (Figure 7).

is an overlap of geometrical shapes like polygons, ellipses, circles, and lines in coherence with basics of drawing depicted from human anatomy. The area being restricted to 4x5 cm^2 has less extension of the wrist, which has shown no effect on multi-band aspect.

Feeding has been tested over 10's of potential positions. Yet, resonance was easily achieved when the coaxial feed is placed the fingers tips. For example, an ultra-wide band (121% bandwidth) between 2.7 and 11 [GHz] is obtained when the feed is at the tip of the first finger as shown in Figure 4. The results at the remaining fingertips are shown in comparison with measurements.

Figure 9: Antenna prototype with slots with the feed at the 1st fingertip.

4. Comparison between measurements and simulations results

A. First prototype, no slots and the power input at the fourth fingertip (Figure 5) :

The first prototype of feeding position at 4th fingertip is shown in Figure 5. It is fabricated as first test of the antenna simulations results. The reflection coefficient S11 is measure using a network analyzer that spams up to 8.5 [GHz].

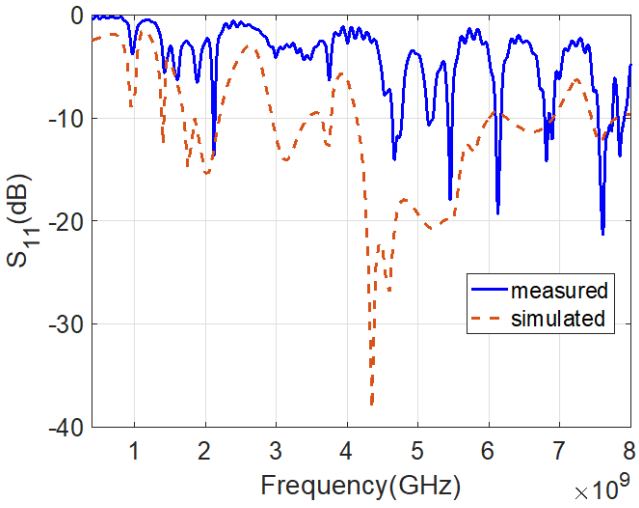


Figure 10: Comparison of S11 obtained in measurements and simulations results for prototype in Figure 9.

S11 results recorded in measurements coincide very well with simulated ones as shown in Figure 6. When comparing the resonance frequencies for the simulated and fabricated cases, small drift was obtained; i.e., (1,66, 1.74), (2.43, 2.34), (3.95, 4), (5.2, 5.06), (6.15, 6.42), and (7.2, 7.4) [GHz]. 0.27 [GHz] is the maximum difference between measured and simulated results due to many reasons, which include:

1) Modeling at a center frequency, while multiple frequencies are targeted and being discovered. As a result, slight differences

Figure 11: Antenna prototype with slots with the feed at the 2nd fingertip.

appear due targeting different frequencies while setting far field in the simulations at fixed wavelength.
2) Fabrication errors, resistive mismatch in the SMA connector, soldering and losses in measurement wires.
3) Since losses increase with frequency, discrepancy increases at

higher frequency to maximum of 0.27 [GHz].

B. Second prototype, no slots with feed at the second fingertip (Figure 7) :

For feeding position at second fingertip, comparing the S11 results as displayed in Figure 8 shows discrepancy at 6.85 GHz being a resonant frequency that did not appear in measurements unlike simulations. However, almost good matching is obtained for measured frequencies at 1.85, 3.85, 4.2, [4.8-5], 5.7 and 7.98 [GHz]. Largest differene between measurements and simulations in listed values is at 3.47 GHz obtained in simulations versus 3.78 GHz in simulations. Such discrepancies results are descirbed in paragarph A of section 4.

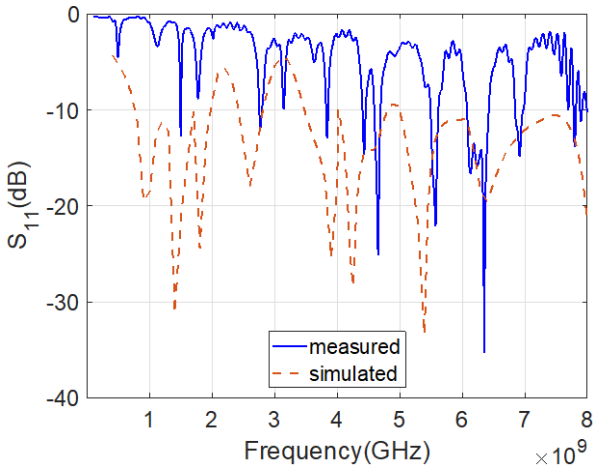


Figure 12: Comparison of S11 obtained in measurements and simulations results for prototype in Figure 11.

C. Third prototype, with slots and feed at the first fingertip (Figure 9) :

Adding slots along all fingers has created multi band operation for feed at all fingertips due to elongating the current path. Here the feed is at the first fingertip as shown simulated results show more like ultra wide band operation below 4 GHz. Common nature of multi-band operation is seen in both results with peaks almost coinciding or falling within matching region i.e for S11 belwo -10 dB. in Figure 9. Comaprision between measured and simulated results is shown in Figure 10. Measured results resonance is obtained at 2.15, 4.5, 5, 5.5, 6.05, 6.8 and 7.7 [GHz], while simulated results show more ultra wide band operation above 4 GHz in addition to several multi-band rsonances belwo 4 GHz. Multiple bands are seen in both results,

Figure 13: Antenna prototype with slots with the feed at the 3rd fingertip.

with peaks almost coinciding or falling within matching region i.e for S11 below -10 dB.

D. Fourth prototype, with slots and feed at the second fingertip (Figure 11) :

Adding slots with the coaxial feed position at the second fingertip is shown in Figure 11. The S11 measurements show good matching with simulation results (Figure 12) at the following resonant frequencies : 1.4, 1.83, 2.72, 3.9, 4.2, 4.78, 5.38, 6.32,6.97,8 [GHz] being the simulated obtained values with maximum difference of 0.25 [GHz]. Resonance at 1 [GHz] did not appear strong enough in measurements, hence it is not considered.

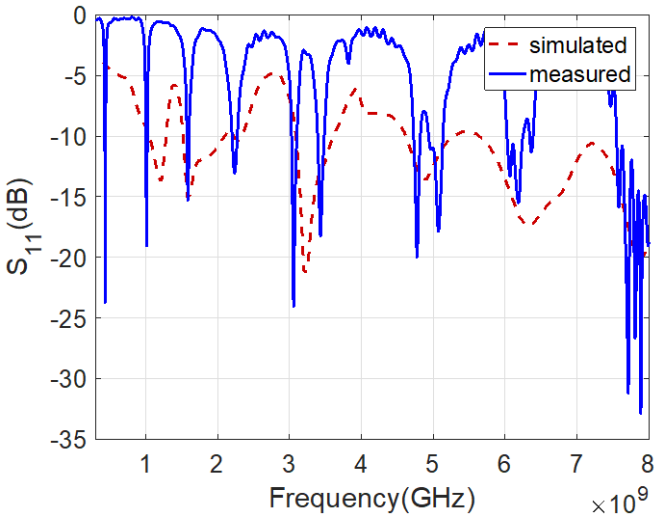


Figure 14: Comparison of S11 obtained in measurements and simulations results for prototype Figure 13.

E. Fourth prototype, with slots and feed at the third fingertip (Figure 13) :

Adding slots with the coaxial feed position at the third fingertip is shown in Figure 13. The S11 measurements show good matching with simulation results in Figure 14, at the

following resonant frequencies : 0.15, 1, 1.65, 2.25, 3, 3.4, 4.8, 5.08, 6.2, and [7.6-8] [GHz]. Resonance in measurements at 0.4 [GHz] could not be obtained in simulated results and hence it is not considered.

5. Dc biasing of mounted varactors

A simple experimental approach of adding varactors across slots was done as an arbitrary trial to check resonances and reconfigurabilty with DC biasing [26]. Simulations were not done to approve results as this experiment falls in further studies to explore the full frequency reconfigurability upon adding varactors. Moreover, experimental tools did not allow DC biasing to the full range of voltage, which is up to 60 V due to summing up the voltage across varactors, which take up to6 V each. As shown in Figure 15, varactors were added at the slots carved through second and third fingers. In order to DC bias the 13 varactors. Inductors were added across slots in 1st and 4th fingers as shown in Figure 15. Hence DC current was drawn through the inductors, which act as AC block at the same time [26]. S11 resonance at low frequencies is remarkable considering the size of the antenna. Frequency reconfigurability over [2.15-2.25] and [0.2-0.8] GHz is obtained upon DC biasing the antennas between [0-10] V as shown in Figure 16. Future studies are to be made by further increasing the DC voltage biasing range. Applications to such frequencies fall in DVB, TV broadcast, LTE, and many other telecommunications standards, which allow -6 dB S11 at maximum return loss [14]. Additional new frequency resonances are obtained at1, 1.6 and 2.15 [GHz] due to adding varactors as shown in Figure 16. The multi-band feature of the antenna at wide range of frequencies enables additional applications such as IOT [6] and energy harvesting [4,27].

Figure 15: Antenna prototype with slots with the feed between 1st and 2nd fingertip and using varactors along 2nd and 3rd fingers slots.

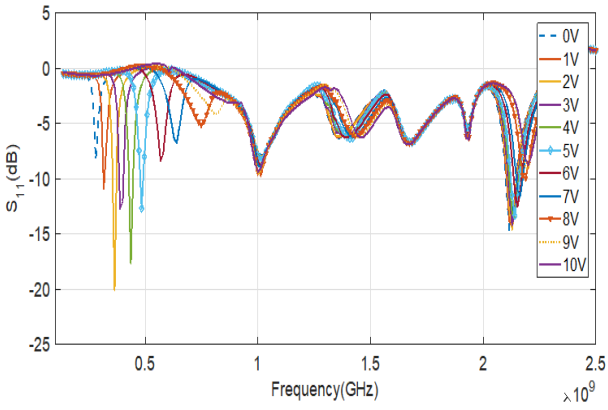


Figure 16: Comparison of S11 obtained in measurements and simulations of prototype in Figure 15 with frequency tunability upon varying the DC bias voltage.

6. Surface current density

Current flow at the antenna surface changes with frequency, consistently with the electric length required for frequency resonance [28]. Surface current density distribution is plotted from simulations for four frequencies as a demonstration. The first antenna prototype in Figure 6 is chosen as an example. Here the feed is at the 4th fingertip. Figure 17 shows the four current graphs corresponding to 1.66, 3.95, 5.2 and 7.2 [GHz] with same color scale with maximum at 6 [A.m^{-1}] and minimum at 0.012 [A.m^{-1}]. However, at 1.66 [GHz] the current density achieves highest value of 60 [A.m^{-1}]. Normally, lowest frequency corresponds to largest electric length and current density values. The current is first

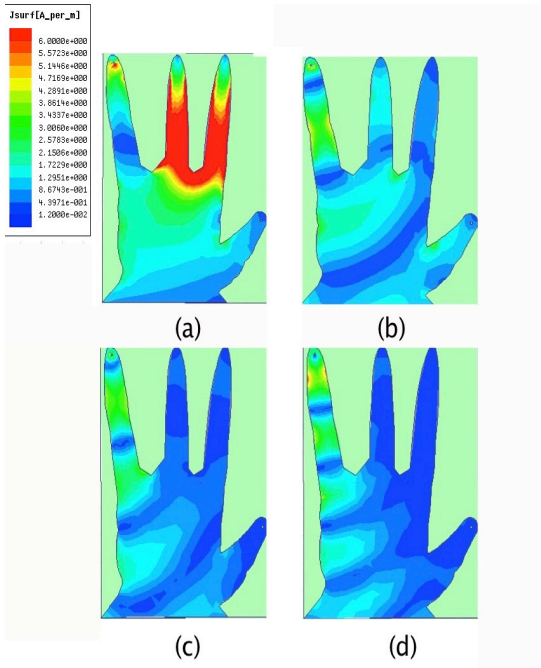


Figure 17: Simulated results for the current distribution for the 4 resonant cases corresponding to 1.66 GHz (Figure a), 3.95 GHz (Figure b), 5.2 GHz (Figure c), and 7.2 GHz (Figure d).

strong especially over the mid part of the hand 1.66 [GHz] in graph(a) of Figure 17; then the current surface area diminishes as

frequency increases as shown in graphs (b,c,d) of Figure 17. Of course, longest current path from input feed to furthest points in the patch assures the lowest frequency resonance. And obviously, the current is drawn to smallest area around the fourth fingertip where the feed is placed, at the highest frequency of 7.2 [GHz] as shown in graph (d) in Figure 17.

7. Radiation pattern

The current flow distribution in Figure 17 radiates the different patterns shown in Figure 18, [28]. The radiation pattern at the 2.43 GHz frequency is shown in Figure 18 (a). The common observed feature at low resonant frequencies is the symmetric current distribution with respect to the right diagonal resulting in a common radiation pattern at 1.66 and 2.43 [GHz]. Here, the axes are same as in Figure 5. Figure (b) has many peaks due to non-symmetric current density at 6.15 [GHz]. Figures 18 (a, b) show the main trend of radiation pattern at low (1.66 and 3.95 [GHz]) and high frequencies (above 5.2 [GHz]) respectively, as a

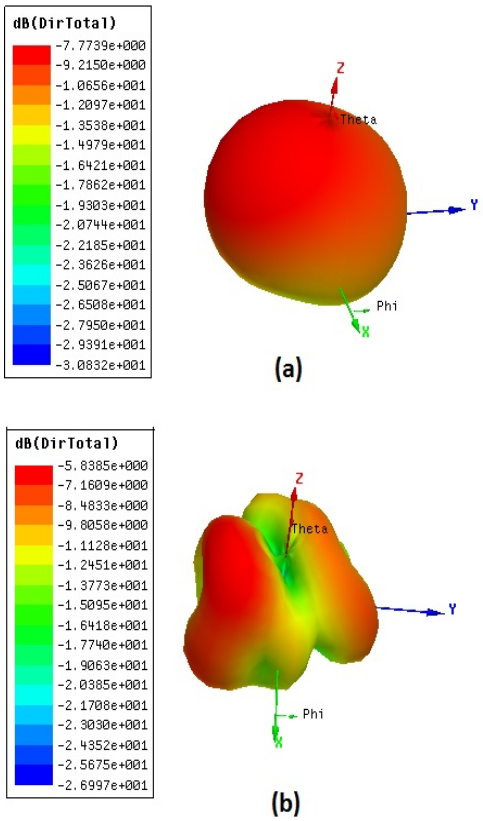


Figure 18: Radiation pattern at 2.43 and 6.15 [GHz] in (a) and (b) respectively.

consequence of change of current density symmetric distribution on the surface. A more description plot of the radiation pattern at 2.43 [GHz] is shown in Figure 19.

8. Conclusion

The 'Hand of God' motif has been successfully translated in antenna design modeling and served in achieving a multi-band operation. Slots are incorporated in the antenna showing an enhanced multi-band operation. In fact, 10 resonant frequencies were obtained over the interval [0.15-8] GHz for one prototype having slots and with the feed at third fingertip. Interesting increase in electrical length of the antenna results when varactors are added across slots. Consequently, resonance at low frequencies of 100 MHz is obtained. Moreover, frequency reconfigurability is observed at frequencies below 1 [GHz]. Finally, the shape, size, and wide range of antenna operation frequencies render it as multi-functional and suitable for different modern applications.

References

[1] Balanis, C. A. 'Antenna Theory'. (Wiley and Sons, INC., 1997).
[2] Ghosal, A., Majumdar, A., Das S. K., & Das A. (2018). 'A multiband microstrip antenna for communication system'. 2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT), Kolkata, 1-4.
[3] Xu, K. D., Li, D., Liu, Y., & Liu, Q. H. (2018). 'Printed Quasi-Yagi Antennas Using Antennas Using Double Dipoles and Stub-Loaded Technique for Multi- Band and Broadband Applications'. IEEE Access, 6, 31695-31702.
[4] Chen, Y., Liu, Q., Tang, X., Mo, Z., Li, C., & Li, F. (2018). 'Compact triple band circularly polarized directional antenna for UHF/ISM RFID mobile readers with GNSS band'. 2018 International Workshop on Antenna Technology (iWAT), Nanjing, 1-3.
[5] Li, Y., Sim, C., Luo Y., & Yang G. (2018). 'Multiband 10-Antenna Array for Sub-6 GHz MIMO Applications in 5-G Smartphones'. IEEE Access, 6, 28041-28053.
[6] Mansour, A. M. et al. (2016). 'Compact reconfigurable multi-size pixel antenna for cognitive radio networks and IoT environments'. 2016 Loughborough Antennas & Propagation Conference (LAPC), Loughborough, 1-5.

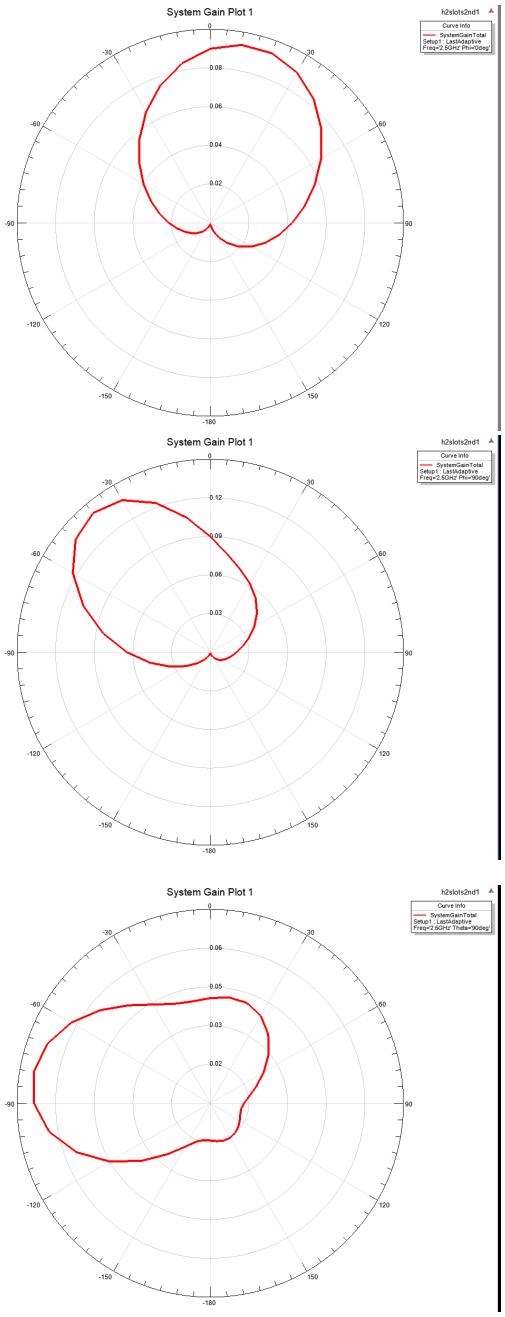


Figure 19: 2D radiation Pattern in XOZ, YOZ, and XOY planes from top to bottom respectively.

[7] Hui, D. (2017). 'A new type of multi-band mobile phone antenna for 3G&4G network', 659-63.

[8] Wang, J. J. H., and Adley, J. C. (2017). '30-2000 MHz multi-band body wearable antenna (MBWA)', 1771-2.

[9] Ozbakis, B., Okuyucu, S., Seemen, M. & Yegin, K., (2017). 'Multi-band frequency tunable LTE antenna for mobile phone applications', 939-43.

[10] Shahzad, M. U., Ahmed, W., Bukhari, S. M. R., Hassan F., & Shahid, H. (2017). 'Multiple band reconfigurable dual H-shaped patch antenna'. 2017 International Symposium on Recent Advances in Electrical Engineering (RAEE), Islamabad, 1-4.

[11] patents.google.com/patent/US8629811B2/en, 2014.

[12] Labidi, M., Salhi, R., and Choubani, F., (2017). 'A design of metamaterial multi-band bowtie antenna based on omega-shaped resonator', Applied Physics A, 123(5), 1-6.

[13] Zhang, Q., & Gao, Y. (2017). 'Compact low-profile uwb antenna with characteristic mode analysis for uhf tv white space devices'. IET Microwaves, Antennas Propagation, 11(11).

[14] Al-Husseini, M., Kabalan, K. Y., El-Hajj, A., Christodoulou, C. G., 'Reconfigurable microstrip antennas for cognitive radio', in Kishk, A., 'Advancement in Microstrip Antennas with Recent Applications'. (InTech, 2013).

[15] en.wikipedia.org/wiki/Hand_of_God(art), 2018.

[16] Mahesh, A., Shushrutha, K. S. & Shukla, R. K. (2017). 'Design of multi-band antennas for wireless communication applications'. 2017 IEEE Applied Electromagnetics Conference (AEMC), Aurangabad. 1-2.

[17] https://i.pinimg.com/564x/91/84/14/91841435effbcaddf955790f09ed4402.jpg

[18] K. Sharma, B. V. R. Reddy and A. Mittal, "Slot Loaded Microstrip Patch Antenna for WLAN and WiMax Applications," 2015 IEEE International Conference on Computational Intelligence & Communication Technology, Ghaziabad, 2015, pp. 597-599.

[19] S. E. Muthumani, Vallikannu R and H. R. Patnam, "Compact slot loaded Koch fractal microstrip patch antenna," 2013 IEEE Applied Electromagnetics Conference (AEMC), Bhubaneswar, 2013, pp. 1-2.

[20] S. Verma, J. A. Ansari and M. K. Verma, "A novel compact multi-band microstrip antenna with multiple narrow slits," Microwave and Optical Technology Letters, vol. 55, (6), pp. 1196-1198, 2013.

[21] S. Ghosal and S. R. B. Chaudhuri, "Analysis of a rectangular slot on a microstrip patch antenna with an equivalent circuit model," 2013 IEEE Applied Electromagnetics Conference (AEMC), Bhubaneswar, 2013, pp. 1-2.

[22] Q. Bai et al, "An Independently Tunable Tri-Band Antenna Design for Concurrent Multiband Single Chain Radio Receivers," IEEE Transactions on Antennas and Propagation, vol. 65, (12), pp. 6290-6297, 2017.

[23] C. Deng et al, "A Novel Low-Profile Hepta-Band Handset Antenna Using Modes Controlling Method," IEEE Transactions on Antennas and Propagation, vol. 63, (2), pp. 799-804, 2015.

[24] K. K. Kishor and S. V. Hum, "Multiport Multiband Chassis-Mode Antenna Design Using Characteristic Modes," IEEE Antennas and Wireless Propagation Letters, vol. 16, pp. 609-612, 2017.

[25] Y. Li et al, "Compact Heptaband Reconfigurable Loop Antenna for Mobile Handset," IEEE Antennas and Wireless Propagation Letters, vol. 10, pp. 1162-1165, 2011.

[26] Madi, M., Al-Husseini, M. & Kabalan, K. Y. (2018). 'Frequency Tunable Cedar-Shaped Antenna for WIFI and WIMAX'. Progress In Electromagnetics Research Letters. 72, 135–143.

[27] Alsharoa, A., Neihart, N. M., Kim, S. W., & Kamal A. E., (2018). 'Multi-Band RF Energy and Spectrum Harvesting in Cognitive Radio'. Networks. 2018 IEEE International Conference Communications (ICC), Kansas City, MO. 1-6.

[28] Madi, M. & Kabalan, K. Y. (2018). 'Multi-Band Hand of God Antenna for Wireless Communications'. ICT conf., St Malo, France, July, IEEE proc.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 235-243 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Energy Consumption Evaluation of Air Cooled Chiller With Cold Storage System Powered by Photovoltaic Modules

Ishaq Sider [1], Maher Al-Maghalseh[*,2], Zaid Alnather[3]

[1]*Mechnical Engineering Department, Palestine Polytechnic University, Palestine*

[2]*Electrical Engineering Department, Palestine Polytechnic University, Palestine*

[3]*Hebron Industrial Secondary School, Palestinian Ministry of Education and Higher Education, Palestine*

ARTICLE INFO

Article history:
Received:21 April, 2019
Accepted:06 June, 2019
Online: 26 June, 2019

Keywords:
Photovoltaic
Thermal Storage
Air-cooled Chiller
TRNSYS
Cold storage system

ABSTRACT

Renewable energy becomes an appealing technology used in many applications in our life. Environmentally, it reduces CO_2 emissions and enhances the sustainability of the system. This paper study the benefits of using a photovoltaic system with a thermal storage tank to power air-cooled chiller, in two different scenarios. The simulation methodology is adopted in this research to study the various scenarios of the combination of the utility, photovoltaic system, thermal storage tank, and air-cooled chiller. The scenarios are based on TRNSYS simulation software. The two scenarios investigated in this study include supplying an air-cooled chiller using a photovoltaic system integrated with the grid. While the second one is to study the photovoltaic system integrated with the grid as well as a thermal storage tank. It was found that the reduction of the consumed energy in the first scenario reduced by 81%. Also, CO_2 emissions reduced by 72%. In addition, the payback period equalled nine years and generated $4,350 in total profit along the project life cycle. The second scenario saves 75.6% of the utility energy consumption and decreases CO_2 emissions by 68%. Moreover, the payback period becomes 12.4 years with $3,202 in total profit generated.

1. Introduction

As innovations advance and the cost of fossil fuel assets develops rapidly, the expanding focus on renewable energy assets is tracked [1]. Recently, the use of solar photovoltaic (PV) systems has increased significantly due to the reduction in the cost of using this technology, in addition to the negative environmental impact of the use of conventional fossil fuels [2,3]. There are several types of PV systems, mostly grid connected, with some also serving as standalone systems. Solar PV systems can be integrated into several devices and applications [4,5]. For example the use of solar photovoltaic energy to power refrigerators has great potential for lowering running costs while also providing high reliability. In addition, this power source has the potential to lengthen the lifespan of kerosene refrigerators, as well as diesel generators, which have been generally used in remote areas. On the other hand, thermal energy storage (TES) systems can play a remarkable role in energy savings by shifting it from on-peak load to off-peak load for cooling by use by the TES system. TES is one of the best methods energy management methods and provides several economic advantages [6-8]. Dincer [9] introduced various examples of cooling thermal energy storage (CITES) and analyzed them through from energy efficiency, environmental and economic advantages.

Many researchers have developed studies and experiments based on vapor compression refrigeration systems driven by photovoltaic cells. Deshmukh1 et al. [10], developed a performance evaluation of a PV system designed for a DC refrigerator. Performance testing of the PV system at no load and full load conditions were carried out to assess the system's technical viability and average PV conversion efficiency. It was found that energy efficiency at both no load and full load conditions were about 8.5% and 11% respectively. Kalbande et al. [11], developed a photovoltaic system for a DC refrigerator were designed and developed to meet the needs of most rural areas which have no access to a national grid or with unstable and erratic electricity supply. The solar photovoltaic operated DC

* Maher Al-Maghalseh, Email: maherm@ppu.edu

vapor compression refrigeration system under test was able to maintain the temperature as specified by the World Health Organization (WHO) for the vaccine preservation (2-8°C). The average photovoltaic conversion efficiency and exergy efficiency of the refrigerator was found to be closer to about 12.05% and 14.20% on full load condition. Fatehmulla et al. [12], studied the performance of the refrigeration system with a PV module is significantly good. Cost comparison between the PV (photovoltaic energy or solar energy) and conventional energy (electrical energy) demonstrates the economic effectiveness of the energy efficient low power PV refrigeration system which, is green, clean and safe, in view of the calculations and the initial cost of our PV system including the initial electrical installation cost to run the low power refrigeration system. The thermal storage tank is a promising new technology that is used to transfer thermal power during periods when power is conserved and stored until it is needed. Rismanchi et al. [13], provides an economic cost-benefit analysis of the system, including the chiller and storage systems. The study was conducted for a range of 100–2000 tonnes of refrigeration (TR) (352–7034 kW), under two storage methods, specifically full storage and load leveling storage strategies. In Rahdar et al. [14], a vapour compression A/C system was analyzed through two strategies of hybrid systems. First, an ice thermal energy storage (ITES) system is used in the a.m. hybrid system; and thereafter a phase change material (PCM) tank is used as a full storage system in order to shift the load from on-peak to off-peak mode. This A/C system is modelled and analyzed from an energy, economic and environmental perspective in both cases.

Oró et al. [15], This paper provides an overview of existing Spanish and European potential energy savings and CO_2 mitigation by incorporating TES systems to cold storage and transportation systems. The total energy demand for cold applications in Spain and in Europe was calculated, and after that, the energy reduction and therefore CO_2 emissions mitigation was determined to assume full implementation of the phase change material (PCM) TES system. The industrial sector shows the highest potential for benefit across all analyzed sectors. With regard to economic savings, Spain would be able to conserve between 2,309 and 11,674 GWh/year, with a potential yearly CO_2 emission reduction of between 1195 to 5,902 [1000 tCO_2/year]. Liu et al. [16] presented an innovative refrigeration system incorporating a phase change material (PCM) proposed to maintain refrigerated trucks at desired thermal conditions. In addition, the system consumes less energy and produces much lower local greenhouse gas (GHG) emissions.

The objective of this research is to merge studies that consider supplying the air-cooled chiller from the PV-system [9-12] with a thermal storage tank used the air-cooled chiller [13-16]. The combination of the two approaches is evaluated for application in the city of Hebron from three standpoints, namely energy consumption, economic feasibility and environmental impact.

2. Methodology

The total heat required for removal from the refrigerated space in order to bring it to the desired temperature and maintain it with refrigeration equipment is known as cooling load. The purpose of load estimation is to determine the size of refrigeration equipment required to maintain the internal design conditions during periods of maximum exterior temperatures. The cooling load is seldom effected by any one source of heat. Rather, it is an accumulation of heat that usually evolves from several different sources, some of which are more common sources of heat that impact the load on the refrigeration equipment, making equipment walls heat gain, the product heat gain, infiltration heat gain, packing heat gain, defrosting heater heat gain and fan motor heat gain [17]. In this study, only the wall heat gain, product heat gain, and infiltration heat gain are estimated. Other cooling load sources such as the fan motor, door lamp, and other components are considered to be very small loads, and are covered by using a factor of safety of 15%, which is added to the total cooling load. The total annual cooling loads are modelled and simulated using TRNSYS simulation software.

2.1. Thermal and Electrical Load Simulation

TRANSYS is a transient system simulation program. The software includes a large library of built-in components, often validated by experimental data [18-20]. TRNSYS consists of suitable programs. In this study, only two of these programs have been deployed which are TRNSYS simulation studio and Multi-zone building (TRNBuild) [21]. TYPE56 (Multi-zone building model) in TRNSYS is selected to simulate the heat conduction, convection and infiltration through surfaces of the one-meter cubic refrigeration chamber cavity, in order to use this type of refridgeration system in two separate processing program. The first process, the TRNBuild program reads in and processes a file containing the chamber description and generates two files. The second process occurrs in the TRNStudio program, the two generated files will be used by the Type 56 component during a TRNSYS simulation. The TRNSYS mathematical model calculations are affected by outdoor climatic conditions, indoor design conditions and the refrigeration chamber envelop structure. The heat balance method is used by TRNSYS as a base for all calculations. For conductive heat gain at the surface on each wall, TRNSYS use Transfer Function Method (TFM) as a simplification of the arduous heat balance method [22, 23]:

$$q_{s,i}^{\cdot} = \sum_{k=0}^{nb_s} b_s^k T_{s,o}^k - \sum_{k=0}^{n_{cs}} C_s^k T_{s,i}^k - \sum_{k=1}^{nd_s} d_s^k q_{s,i}^k \quad (1)$$

$$q_{s,o}^{\cdot} = \sum_{k=0}^{na_s} a_s^k T_{s,o}^k - \sum_{k=0}^{n_{bs}} b_s^k T_{s,i}^k - \sum_{k=1}^{nd_s} d_s^k q_{s,o}^k \quad (2)$$

where:

q_s: Conduction heat flux throw the wall [kJ/h]

a_s, b_s, c_s, and d_s: z-transforms of the surface temperature and heat flux determined by the z-transfer function routines of literature [24].

k: refers to the term in the time series, and it specified by the user within the TRNBUILD description.

The heat gain by convection is calculated the following equation [24]:

$$q_c = h_{convc} \times \Delta T \quad (3)$$

where:
q_c: conviction heat flux [kJ/h]
h_{convc}: conviction heat transfer coefficient [W/m^2^{o}C]
ΔT: surface temperature deference [^{o}C]

The Latent heat gain by the ventilation or infiltration is calculated by using [24]:

$$q_{\inf} = \rho \times V_f^* \times C_p \times (T_o - T_i) \quad (4)$$

where:
ρ: Air density [1.25 kg / m^3].
C_p: Specific heat of the air [1000 J / kg.^{o}C].
V^*_f: Volumetric flow rate of infiltrated air [m^3/s].
T_o: Outside temperature [^{o}C].
T_i: Inside temperature [^{o}C].

TRNStudio program is used to model the product load inside the refrigeration chamber by using equation 5.

$$q_{\Pr od.} = \frac{m \ .C_p \ .\Delta T}{C.T} \quad (5)$$

where:
q_{Prod}: Cooling product load in [kW].
m: Mass of the product in [kg]. (100 kg of water used in this study)
C_p: Specific heat of the product in [kJ/kg. ^{o}C]. (C_P for water equal to 4.18 kJ/kgK)
ΔT: Temperature deference for the product [^{o}C].
C.T: Cooling time [sec].

The total annual load of the refrigeration chamber is the summation of the load from the wall from TRNBuild program using type 56 and the product load using the equation type in the TRNStudio program. Figure 1 shows the load from the wall obtained from the TRNbuild program, the product load obtained from TRNStudio program and the total annual cooling load for the refrigeration chamber.
In this study, a half refrigeration ton (1.75 kW) air-cooled chiller selected for covering the cooling load demand. This chiller produced by ChillX company with the model (CXF050DRS) [25]. The chiller used ethylene glycol-water mixture in order to decrease the mixture freezing point under -10 ^{o}C, when the glycol percentage equal to 40% (by the volume) in the water mixture the freezing point temperature equal to -23.8 and the specific heat for the mixture at this point equal to 3.4 kJ/kg K [26]. Type 655 used to model a vapour compression air cooled chiller. This type requires input parameters, specifically the inlet fluid (water ethylene glycol) temperature, inlet fluid flow rate, setpoint temperature, ambient temperature, and fluid specific heat in order to calculate the annual thermal cooling load demand. Inlet fluid temperature is taken zero ^{o}C, inlet fluid flow rate calculated using equation 6, setpoint temperature is taken -5 ^{o}C, ambient temperature taken from the weather data file for Hebron city and the specific heat of the fluid equal to 3.4 kJ/kgK.

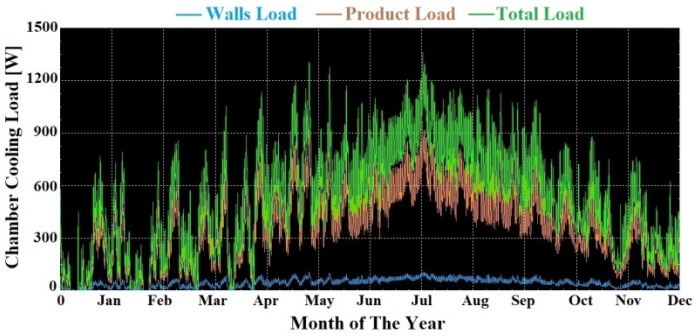


Figure.1. Walls load, product load and total annual cooling load for the refrigeration chamber.

$$m^* = \frac{Q_{total}}{C_p \times \Delta T} \quad (6)$$

where:
Q_{total}: Thermal cooling load for the chamber [W].
m^*: Mass flow rate [kg/sec].
C_p: Specific heat of the water ethylene glycol in [kJ/kg. ^{o}C].
ΔT: Temperature deference for inlet and outlet fluid [^{o}C].

In order to calculate the total electrical power required to cover the total thermal cooling load, type 655 requires sample data for the COP of the utilized chiller. This data is entered into the simulation model (type655) using the text file from the selected chiller data sheet [27]. Type655 then calculates the electrical power using equation 7.

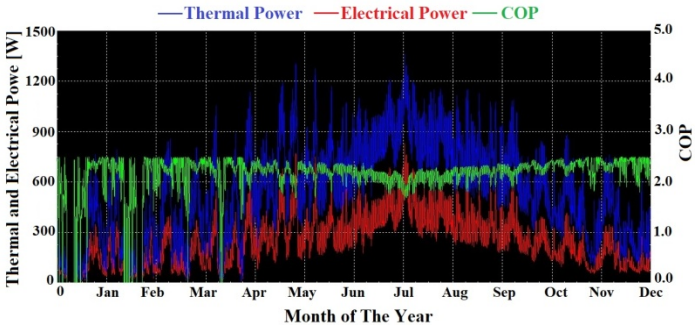
$$P_{elec} = \frac{Q_{total}}{COP} \quad (7)$$

where:
Q_{total}: Thermal Cooling load for the chamber [W].
P_{elec}: Electrical Power [W].
COP: Coefficient of performance for the chiller.

Figure 2 illustrates the total thermal power for the refrigeration chamber, the total electrical power needed for the chiller in order to cover the thermal power and the COP values. As shown in figure 2 the maximum electrical power value equal to 835 W in July, where the thermal power reaches its maximum value. The total annual electrical energy is calculated by integrating the electrical power profile (red profile) this value equal to 2,107 kWh/year.

2.2. System Load Coverage Scenarios

In order to operate the air-cooled chiller that used to cover the total annual cooling load of the refrigeration chamber, two scenarios are used to achieve this goal. In the first scenario, only PV system used to operate the air-cooled chiller during on-peak periods, in this case, the system contains PV modules, on grid inverter and air-cooled chiller as shown in figure 3.



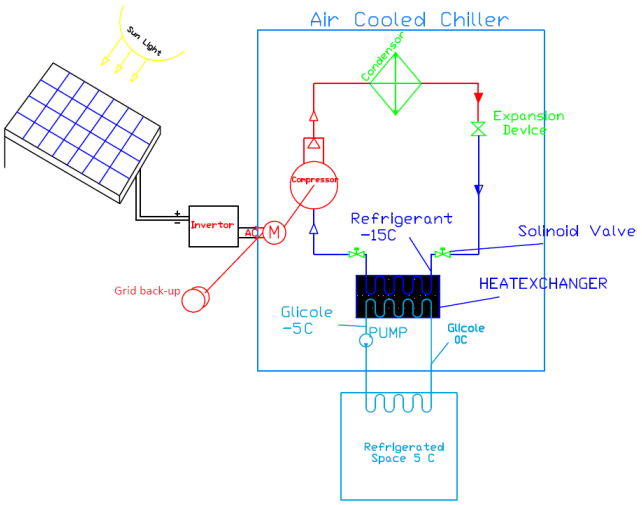
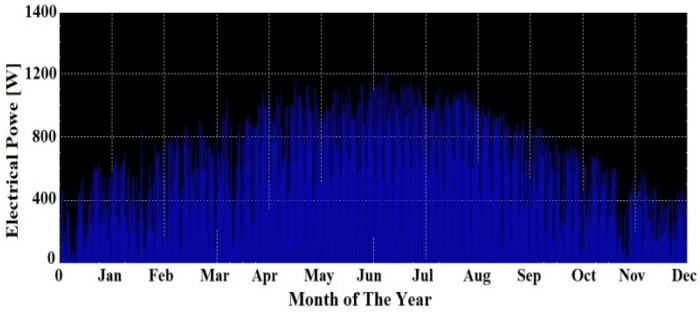
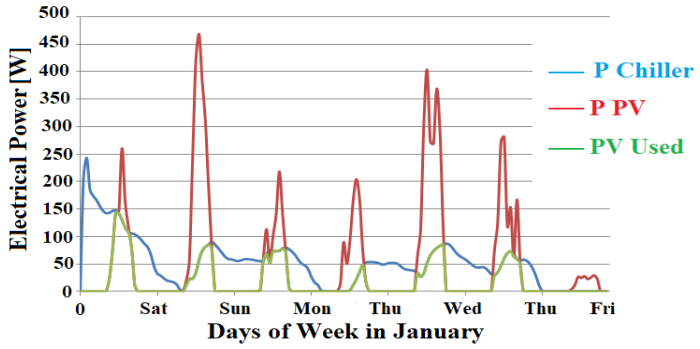


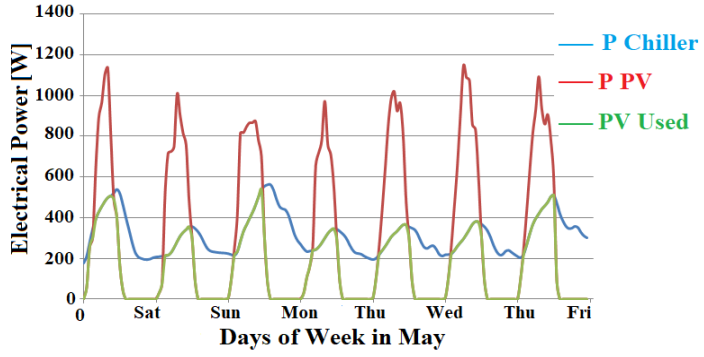
Figure.3. System components in the first scenario.

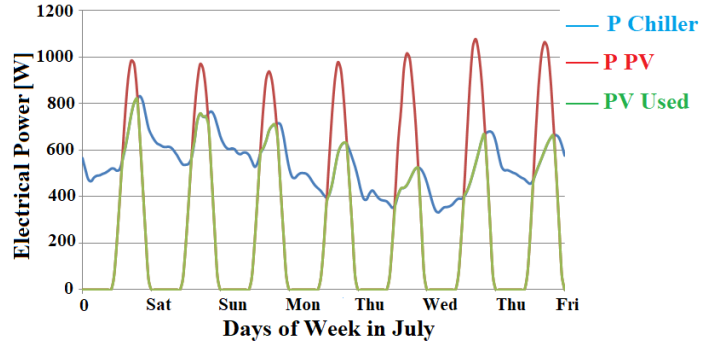
Four Q CELLS (325 W_p) solar PV modules are connected in series which converts solar radiation into electric power as direct current (DC). The 2kW EVERSOL inverter with 97% efficiency converts DC into alternating current (AC) which is needed to drive the chiller compressor. The chiller converts the AC power to the required thermal cooling power. The air-cooled chiller is supplied as a back-up with an electric AC power from the grid, when there is not enough DC power from the PV-array, especially at night, in the evening and morning times of the day when there is not enough solar radiation to drive the chiller. Type 194b (PV-inverter) is used to model the PV array and the inverter from the two scenarios in this study, the model is based on the calculation method presented by De Soto et al. [28]. Type 194b requires many input parameters from Hebron, specifically data files such as the total radiation on tilt angle, beam radiation, sky diffuse radiation, ground reflected diffuse radiation, the slope of the surface, wind speed and ambient temperature. The model of the electrical characteristics from the PV module data sheet, such as short circuit current and open circuit voltage at Standard Test Condition (STC), module voltage and current at the maximum power point (MPPT), temperature coefficient at the short circuit current, as well as the open circuit voltage, the number of cells in series, the number of module in series, the number of module in parallel, the normal operating conditions test (NOCT) and the module area [29]. Also, the inverter parameters efficiency and power are used in type 194b. By running the simulation program for the first scenario, the output electrical power obtained from the PV array along the year shown in figure 4.

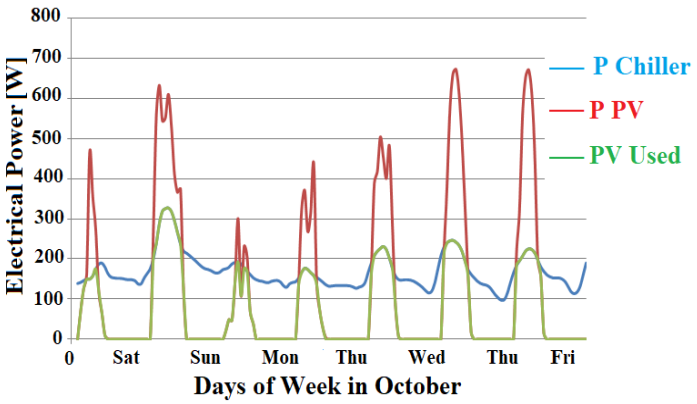


The obtained electrical power values moved to excel sheet file in order to compare between the chiller power and the power produced by the array. Figure 5 to figure 8 illustrate the weekly distribution of electrical powers of the system in winter, spring, summer and in autumn for the chiller and the PV array.

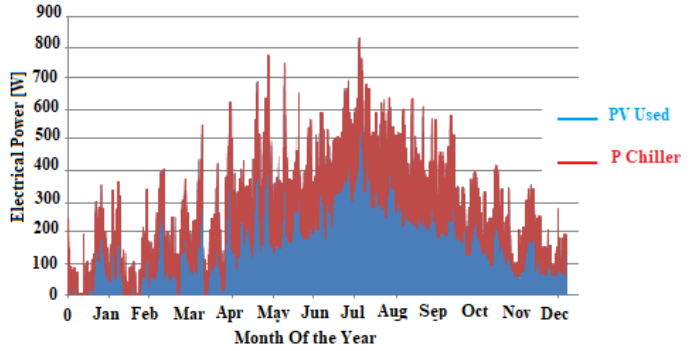






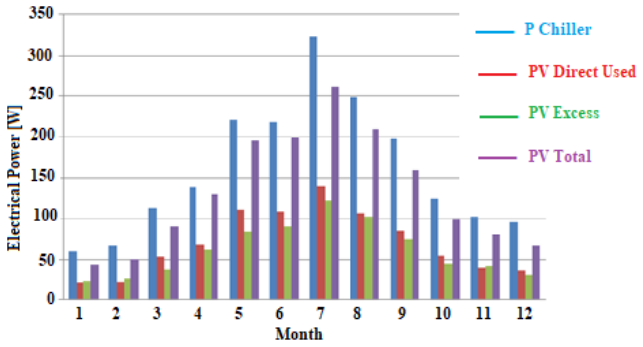


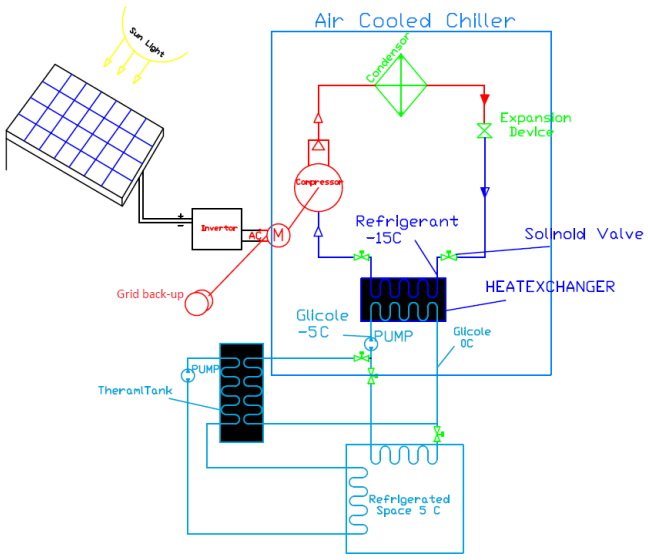
The power that direct used from PV array in order to run the chiller in the on-peak hours is represented in figure 9 in the blue profile, this profile integrated using excel file to calculate the total annual power that direct used from the PV array, which equals to 950.5 kWh/year, this value represents 45% of the total annual power (2107 kWh/year) that needed to run the chiller along the year as illustrated in figure 9.



In this scenario the excess power from PV array is supplied to the grid, getting an equal value for it during night hours - in case of electricity consumption for the chiller is greater than the PV system production in one month - based on Power Authority laws in Palestine, this value equals to 752 kWh/year which represents 36% of the total annual power needed. Based on the first scenario results, 81% of the total annual electrical power required to run the chiller was generated. Figure 10 depicts the electrical power consumed by the chiller and the power produced by the PV-system, the blue profile represents the total electrical power needed to run the chiller, the purple profile represents the total electrical power produced by the PV-system, the red profile represents the electrical power direct used from the PV array at on-peak periods and the green profile represent electrical power that transferred to the grid and used at the off-peak periods.

In the second scenario PV system used in order to run the air-cooled chiller during on-peak periods, also the thermal storage tank used in order to shift the excess power from on-peak periods to off-peak periods, this case contains the same components of the system in the previous scenario, in addition to the thermal storage tank as shown in figure 11.





The working fluid used in the storage system is the same fluid that used for the chiller, specifically a glycol-water mixture (40% glycol) with 3.4 kJ/kgK specific heat and 1110 kg/m^3 density [24]. The volume of the thermal storage tank is designed according to the maximum excess of thermal power produced by the PV in a day, this is equal to 8000 Wh/day in June. By using equation 8, the tank volume equals 1.5 m^3 (1500 L).

$$V = \frac{m}{\rho} \tag{8}$$

where:
V: Volume of the tank [m^3].
m: Fluid mass [kg].
ρ: Fluid density [kg/m^3].

The fluid mass equals to 1745.45kg calculated using maximum excess the thermal power by using equation 9.

$$m = \frac{Q_{excess}}{Cp \times \Delta T} \tag{9}$$

where:
m: Fluid mass [kg].

Cp: Fluid specific heat[kJ/kgK].
ΔT: Fluid temperature difference [°C].

The selected tank is (HF1500) produced by Reflex, a Germany company [30]. The heat loss of this tank is equal to 5.1 kWh/24h (212 Wh) in the worst case of heat loss.

In this scenario the excess electrical power produced by the PV array in the first scenario used to run the chiller in TRANStudio, then chiller produced thermal power that stores in the thermal storage tank. The storage tank supplies this power to the refrigeration chamber during the night hours. The output thermal load from the thermal tank depending on its heat losses which equal 212 Wh, the output thermal load calculated using excel sheet file. Figure 12 shows the thermal tank input and output thermal power. The thermal losses in storage tank equal to 310.6 kWh/year.

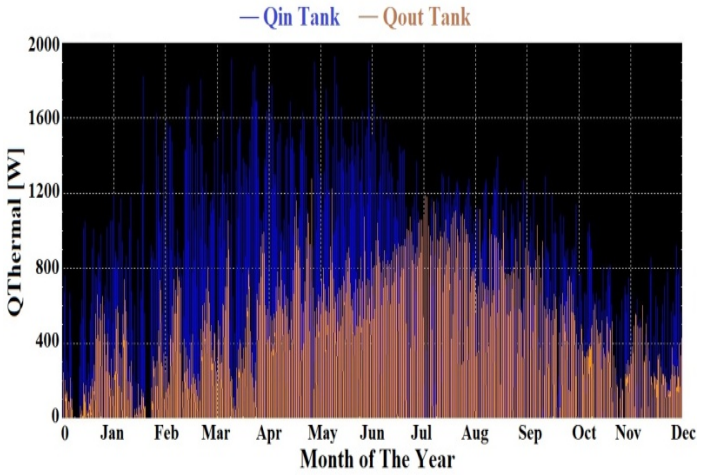
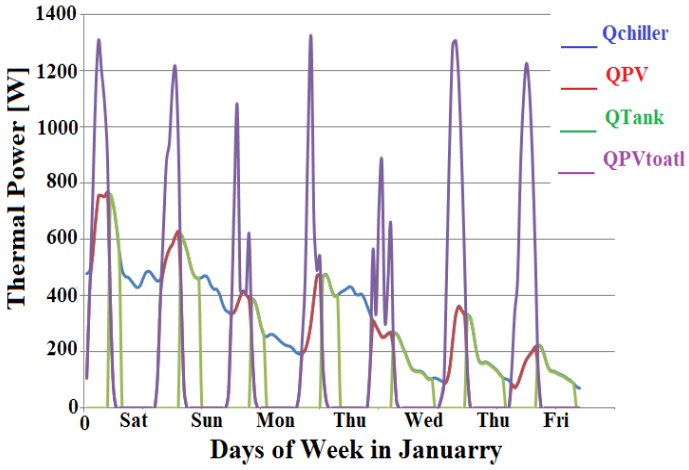
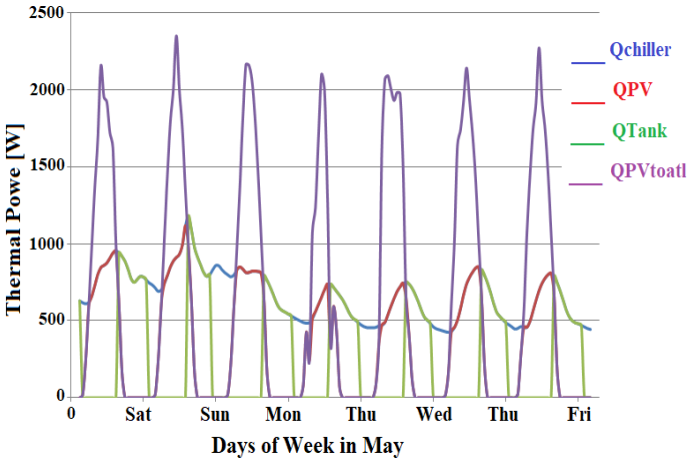
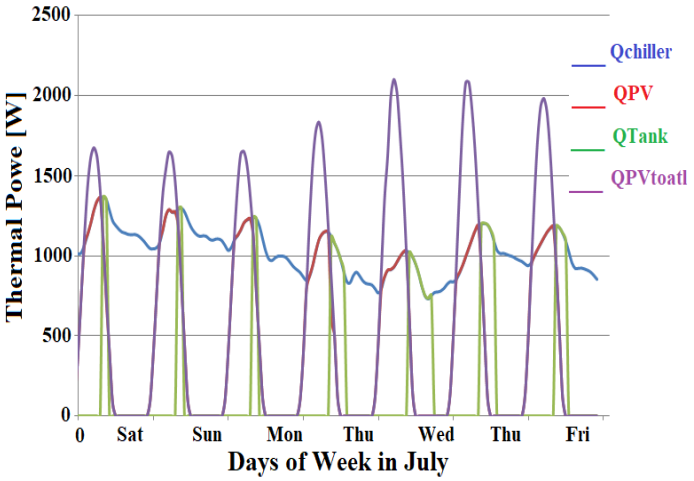


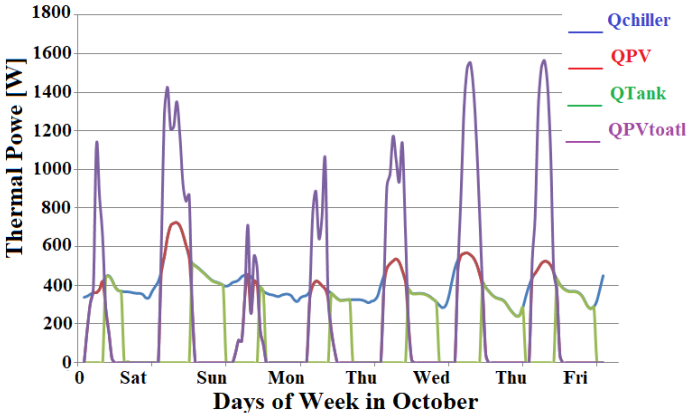
Figure.12. Thermal tank input and output power.

Figure 13 to Figure 16 illustrates the weekly energy distribution during the winter, spring, summer and autumn for the total thermal power produced by the PV array, thermal power direct used from the PV array and thermal power used from the tank in order to cover the chiller thermal load.



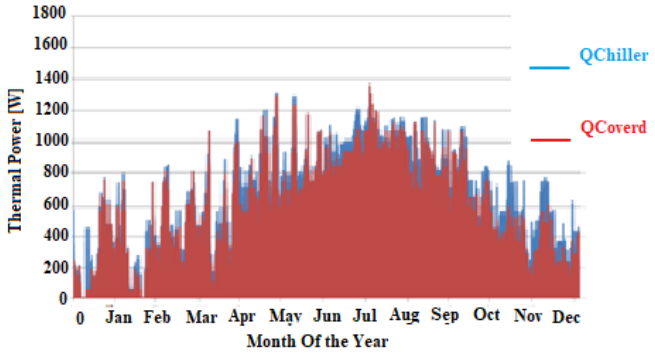






As shown in figures 6.13 to 6.16, the purple profile represents the total thermal power produced by the PV, the red profile represents the thermal power direct used from the PV array at on-peak periods, the green profile represents thermal power obtained from storage tank at off-peak periods and the blue profile represents the chiller thermal power. In winter, 70% of the chiller load was covered, 38% direct from the PV array and 32% from the storage

tank. In spring, 89% of the chiller load was covered, 49% direct from PV array and 40% from the storage tank. In summer 59% of the chiller load was covered, 42% direct from PV array and 17% from the storage tank. Finally, during the autumn, 74% of the chiller load was covered, 40% direct from PV array and 34% from the storage tank. In this scenario, the excess power from PV array is moved to the thermal storage tank in order to use it at night hours, as shown in figure17 the total needed load expressed using blue profile, and the total thermal power that covered using PV and storage tank expressed in a red profile. The red profile integrated using excel file to calculate the total annual thermal power that direct used from the PV array and storage tank, which equal to 3,474 kWh/year, this value represent 75.6% (45% PV and 30.6% storage tank) of the total annual thermal power (4,595.4 kWh/year) that needed to run the chiller along the year as illustrated in figure 17.



According to the second scenario results', the value of the electrical power that reduced in this scenario equal to 1,593 kWh/year, this power obtained by using PV array and storage tank which represent 75.6% of the total electrical power needed to run the chiller (2,107 kWh/year).

2.3. Economic Evaluation

The economic feasibility of investments in PV systems with thermal storage tanks have been conducted in this study for the two scenarios. The selected indicators for this kind of assessment are the payback period (PBP) and total profit generated. On the other hand, the environmental advantage in comparison to traditional sources of energy is evaluated through the reduction of carbon dioxide emissions. The project life period is taken as 25 years according to the PV module performance guarantee. Payback periods describe the length of time required to recover the initial cost of system implementation, while the total profit is the number of dollars that decreased after the payback period, it can be calculated by knowing the annual revenue in the years after the payback period. This values can be estimated by calculating the total capital cost and the total annual cost for the two scenarios using the following equations [31-34].

$$CRF\left(i\%, n\ year\right) = \frac{i\left(i+1\right)^{n}}{\left(i+1\right)^{n} - 1} \tag{10}$$

$$AC = IC \times CRF \tag{11}$$

$$TAC = AC + OC + MC \tag{12}$$

$$TCC = \frac{TAC}{CRF} \tag{13}$$

$$Total\ Annual\ Saving = E_{Annual} \times T_{r} \tag{14}$$

$$Payback\ Time\ \ (PT) = \frac{Total\ Capital\ Cost}{Annual\ Saving} \tag{15}$$

$$Total\ \ \text{Pr}ofit = Annual\ \ \text{Re}venue \times \left(n - PT\right) \tag{16}$$

where :
AC: Annual Cost
IC: Initial Cost
MC: Maintinance Cost
OC: Operation Cost
TAC: Total Amount Cost
TCC: Total Capital Cost
CRF: The Constant Rate Factor is the default quantity *CRF* = 0.071 CRF/year.
n: Project life Period (25 years).
i: Loan interest took (5%), in this study represents the annual depreciation of the system during the life of the project.
T_r: Tariff price in Hebron City (0.58 NIS [35]).
E_{Annual}: Total electrical energy obtained using each scenario.

E_{Annual} equal 1,702.5 kWh/year for the first scenario (81% of the chiller electrical energy) and 1,593 kWh/year for the second scenario (75.6% of the chiller electrical energy). The initial cost contains the total cost that needed to operate the system at the beginning of the project life period such as PV modules cost (205 $/module), inverter cost (US$ 752), storage tank cost (US$ 700) and system installation cost. Installation cost found the range from US$ 0.064-US$ 0.1/Wp [31] (in this study taken 0.1), so it is equal US$ 130 for the first and second scenario (4 modules x 325 Wp) Table 1 shows the required initial cost of both scenarios.

Table 1: Scenario initial costs

Initial Cost	First Scenario	Second Scenario
Modules Cost($)	820	820
Inverter Cost ($)	752	752
Storage Tank Cost ($)	0	700
Installation Cost ($)	130	130
Total initial Cost ($)	1,702	2,402

Annual cost contains the operation and maintenance cost which represent the annual money that needed to operate and maintain the suggested system (including PVs, inverter and storage tank), according to the operating cost 1.5% [36] added to the annual cost that calculated using equation 11, and US$ 0.04/Wp [37] for system maintenance (US$ 52 for the first and second scenarios). Then the total annual cost calculated using equation 12. the payback period and total profit for the two scenarios as shown in table 2.

Table 2: Economic calculations for the study scenarios

Economical Calculations	First Scenario	Second Scenario
E annual (kWh/year)	1,702.5	1,593
Initial Cost ($)	1,702	2,402
CRF / year	0.071	0.071
Annual Cost ($)	120.8	170.5
Operating Cost ($)	1.8	2.6
Maintenance Cost ($)	52	52
Total Annual Cost ($)	175	225
Total Capital Cost ($)	2,460	3,170
Total Annual Saving ($)	272	255
Payback Period / year	9	12.4
Total Profit ($)	4,350	3,202

As illustrated in table 2 the payback periods equal to 9 and 12.4 for the first and second scenarios respectively. Furthermore, the total profit values for the two scenarios equal to US$ 4,350 and US$ 3,202 respectively. According to the economic assessment for the two scenarios, the most economical scenario in this study is the first one, an on-grid PV system without a thermal storage tank.

2.4. Environmental Evaluation

Using renewable energy sources allows for the reduction of environmental pollution. In order to evaluate the environmental advantages in this study, a comparison between CO_2 emissions released by the study scenarios and the ones released by using grid electricity produced by fossil fuels. The quantified emissions released by a PV system equal to 81 gCO_2/kWh, of which 93.7% are caused by PV panel manufacture [38]. On the other hand, fossil sources emissions were quantified in 771 gCO_2/kWh [39]. In this research, running the chiller using grid electricity only produced 1,624.5 $kgCO_2$/year. The first scenario used 1702.5 kWh/year from the PV-system and 404.5 kWh/year from grid electricity, so the CO_2 emissions equal to 450 $kgCO_2$/year. The second scenario used 1,593 kWh/year from the PV system and 514 kWh/year from the grid, so the CO_2 emissions equal to 525.3 kg CO_2/year.

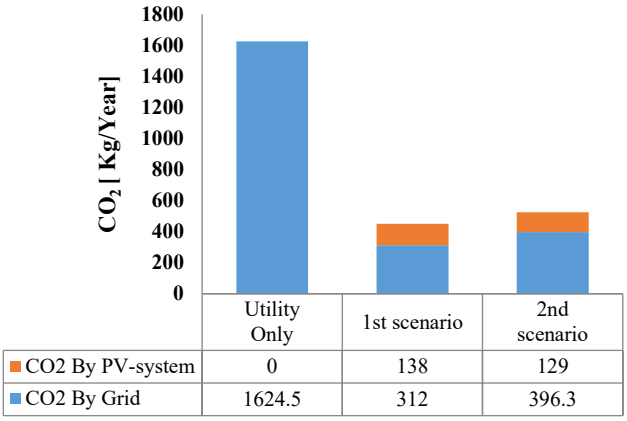


Figure.18. The amount of CO_2 emissions for the system scenarios.

Based on the CO_2 emissions for the system scenarios shown in figure 18. The CO_2 emissions reduced by 72% for the first scenario compared with supplying the chiller from the grid only. In the same manner, the amount of CO_2 reduction for the second scenario equals 68%. According to these results, the first scenario has the lowest impact on the environment.

3. Conclusions

This investigation of the utilization of PV-systems with thermal storage tanks to power air-cooled chillers has yielded very effective and efficient results in reducing electrical consumption from this utility, as well as an alternative energy source which provides a reduction in CO_2 emissions and enhances the payback period and the system profit. Powering the chiller with the grid and PV-system saves 81% of the electrical power, compared with supplying it from the grid only. This result is a payback period of 9 years and a profit of US$ 4,350 generated by the system, in addition to reducing CO_2 emissions by 72%. By adopting thermal storage tank of 1.5 m^3 with the utility and PV-system the electrical consumption decreased by 75.6% and CO_2 emissions by 68%, also the payback period becomes 12.4 years, and the total profit equal to US$ 3,202. According to the above listed economic and environmental examinations, the on-grid system is the most economical scenario. On the other hand, the first scenario has the lowest impact on the environment. The main findings of this study may be applied to larger systems, as well as systems in remote areas, to reduce electricity cost and reduce the amount of CO_2 emissions. Furthermore, the usage of thermal storage units enable for a reduction in scheduled maintenance required by PV-systems with electrical batteries.

References

[1] Vasarevicius, D., & Martavicius, R. (2011). Solar irradiance model for solar electric panels and solar thermal collectors in Lithuania/I lietuvoje naudojamas sales baterijas ir kolektorius patenkancio sales energies auto models. Elektronika IR Electrotechnical, (2 (108)), 3-7.

[2] Al-Maghalseh, M. M. (2018). The Impacts of Distribution Generators Size and Location on Power Efficiency and Voltage Profile in Radial LV Networks. Advances in Science, Technology and Engineering Systems Journal, 3, 276-283.

[3] Al-Maghalseh, M. M. (2018). Evaluating the Reliability worth Indices of Electrical Medium Voltage Network: Case Study. Procedia Computer Science, 130, 744-752.

[4] Al-Maghalseh, M., Odeh, S., & Saleh, A. (2017, October). Optimal sizing and allocation of DGs for real power loss reduction and voltage profile improvement in radial LV networks. In 2017 14th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT)(pp. 21-25). IEEE.

[5] Al-Maghalseh, M. M., Abutemeha, A. K., & Iyadiyyeh, M. M. (2017). Modelling a hybrid system for electrical generation and wastewater treatment using photovoltaic and fuel cells. Desalination and Water Treatment, 100, 250-257.

[6] Al-Maghalseh, M., & Mahkamov, K. (2018). Methods of heat transfer intensification in PCM thermal storage systems. Renewable and Sustainable Energy Reviews, 92, 62-94.

[7] Al-Maghalseh, M. M. (2017). Investigate the Natural Convection Heat Transfer in A PCM Thermal Storage System Using ANSYS/FLUENT. Jordan Journal of Mechanical & Industrial Engineering, 11(4).

[8] Al-Maghalseh, M. (2014). Compact solar thermal energy storage systems using phase change materials (Doctoral dissertation, Northumbria University).

[9] Dincer, I. (2002). On thermal energy storage systems and applications in buildings. Energy and Buildings, 34(4), 377-388.

[10] Deshmukh, S. D., & Kalbande, S. R. (2015). Performance evaluation of photovoltaic system designed for DC refrigerator. Int J Sci Res, 4(2), 18-23.

[11] Kalbande, S. R., & Deshmukh, S. (2015). Photovoltaic Based Vapour Compression Refrigeration System for Vaccine Preservation.

[12] Fatehmulla, A., Al-Shammari, A. S., Al-Dhafiri, A. M., & Al-Bassam, A. A. (2011, April). Design of energy efficient low power PV refrigeration system. In 2011 Saudi International Electronics, Communications and Photonics Conference (SIECPC) (pp. 1-5). IEEE.

[13] Rismanchi, B., Saidur, R., Masjuki, H. H., & Mahlia, T. M. I. (2012). Energetic, economic and environmental benefits of utilizing the ice thermal storage systems for office building applications. Energy and Buildings, 50, 347-354.

[14] Rahdar, M. H., Emamzadeh, A., & Ataei, A. (2016). A comparative study on PCM and ice thermal energy storage tank for air-conditioning systems in office buildings. Applied Thermal Engineering, 96, 391-399.

[15] Oró, E., Miró, L., Farid, M. M., Martin, V., & Cabeza, L. F. (2014). Energy management and CO2 mitigation using phase change materials (PCM) for thermal energy storage (TES) in cold storage and transport. International Journal of Refrigeration, 42, 26-35.

[16] Liu, M., Saman, W., & Bruno, F. (2012). Development of a novel refrigeration system for refrigerated trucks incorporating phase change material. Applied Energy, 92, 336-342.

[17] Dossat Roy, J. (1961). Principles of refrigeration. John Wiley And Sons Inc; New York.

[18] Calise, F. (2012). High-temperature solar heating and cooling systems for different Mediterranean climates: Dynamic simulation and economic assessment. Applied Thermal Engineering, 32, 108-124.

[19] S.A. Klein, et al. (2006). TRNSYS. A Transient System Simulation Program, Solar Energy Laboratory, University of Wisconsin, Madison.

[20] Energy modelling and building physics resource base, Software, TRNSYS, University of Cambridge.

[21] TRNSYS 16.(2009). A TRaNsient System Simulation program – Volume 1 Getting Started. Solar Energy Laboratory, University of Wisconsin-Madison.

[22] Bhaskoro, P. T., & Gilani, S. I. U. H. (2011). Transient Cooling Load Characteristic of an Academic Building, using TRNSYS. Applied Sci, 11, 1777-1783.

[23] TRNSYS Group. (2003). TRNSYS16 manual, < http://www.aiguasol.coop/files/file46-3.pdf>.

[24] Mitalas, G. P., & Arseneault, J. G. (1970). Fortran IV program to calculate z-transfer functions for the calculation of transient heat transfer through walls and roofs. National Research Council Canada, Division of Building Research.

[25] Jannot, Y., & Degiovanni, A. (2018). Thermal properties measurement of materials. John Wiley & Sons.

[26] ChillX company, home page 1/4-1 Ton Compact Low Temp. Chiller, last visited website page 12/2018, <http://chillxchillers.com /compact-chillers/chillx-025-1-ton-compact-low-temp-chillers?sku= CXF025DRS1>.

[27] The Engineering ToolBox, Freezing Point of Ethylene Glycol based Water Solutions, last visited website page 12/2018, <https://www.engineeringtoolbox.com/ethylene-glycol-d_146.html>.

[28] De Soto, W., Klein, S. A., & Beckman, W. A. (2006). Improvement and validation of a model for photovoltaic array performance. Solar Energy, 80(1), 78-88.

[29] Al-Maghalseh, M. (2018). Experimental Study to Investigate the Effect of Dust, Wind Speed and Temperature on the PV Module Performance. Jordan Journal of Mechanical & Industrial Engineering, 12(2).

[30] The reflex company, homepage, HF Storage Tank, last visited website page 12/2018, <www.reflex.de>.

[31] Al-Maghalseh, M. M., & Maharmeh, E. M. (2016). Economic and Technical Analysis of Distributed Generation Connection: A Wind Farm Case Study. Procedia Computer Science, 83, 790-798.

[32] Al-Maghalseh, M., & Saleh, W. (2017, October). Design and cost analysis of biogas based power plant: Jenin perspective. In 2017 14th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT) (pp. 31-35). IEEE.

[33] Al-Maghalseh, M . (2018). Techno-economic assessment of biogas energy from animal wastes in central areas of Palestine: Bethlehem perspective. International Journal of Energy Applications and Technologies, 5 (3), 119-126. DOI: 10.31593/ijeat.444089.

[34] Al-Maghalseh, M . (2019). Generation unit sizing, economic analysis of grid-connected and standalone PV power plant. International Journal of Energy Applications and Technologies, 6 (1), 1-7. DOI: 10.31593/ijeat.444091.

[35] Hebron Electric Power Company, website page, Tariff cost, last visited website page 12/2018, <http://hepco-pal.com/ar/services/news2>.

[36] ERSS, X., PTP, C. E., C&R, C. E., & Engineer's Building, D. (2017). Central e

Electricity Regulatory Commission New Delhi.

[37] Bolinger, M., & Weaver, S. (2014). Utility-scale solar 2013: An empirical analysis of project cost, performance, and pricing trends in the United States.

[38] Cucchiella, F., & D'Adamo, I. (2012). Estimation of the energetic and environmental impacts of roof-mounted building-integrated photovoltaic systems. Renewable and Sustainable Energy Reviews, 16(7), 5245-5259.

[39] Cucchiella, F., D'Adamo, I., & Gastaldi, M. (2015). Financial analysis for investment and policy decisions in the renewable energy sector. Clean Technologies and Environmental Policy, 17(4), 887-904.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 244-250 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Sentiment Analysis on Twitter for Predicting Stock Exchange Movement

Yuliana Tanulia, Abba Suganda Girsang*

Computer Science Department, BINUS Graduate Program-Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480

ARTICLE INFO

Article history:
Received:14 May, 2019
Accepted:06 June, 2019
Online: 26 June, 2019

Keywords:
Stock Market
Twitter
Support Vector Machine
Nonnegative Matrix Factorization

ABSTRACT

This paper is proposed to build a model by applying two methods, namely support vector machine and nonnegative matrix factorization in the process of predicting stock market movement using twitter and historical data. The stock exchange is based on the LQ 45 stock with period from August 2018 - January 2019. The features consist of closing price, volume, percentage of topics and sentiment. The price and volume are taken from yahoo finance data, while topics and sentiment are taken from comments of each stock market in LQ45. NMF method is used to get the topic percentage feature in the tweet data, while the analysis sentiment value is obtained using SVM. The evaluation results obtained by getting the level of accuracy using confusion matrix. The accuracy value of stock movement predictions in this study is 60.16%.

1. Introduction

Many companies use twitter to analyze products and improve relationships also trust between consumers and companies [1] One of the using twitter in the economic field is predicting economic movements and sales indicators for a product [2]. One prediction in the field of economics is predicting stock prices. Stock price prediction is an interesting topic for business people and researchers. Stock is one of the capital market products which is one of the investments for the long term [3]. Stock price predictions involve a complex process because of the noise and the very irregular conditions. Stock market prediction is an important issue for transaction participants in assisting in decision making.

Sentiment analysis is related to the increase and decrease in stock prices [4]. Several studies compared several methods used in sentiment analysis. From previous studies, methods that are commonly used in the process of sentiment analysis are naïve bayes classifier and support vector machine. The comparison results of that techniques, the SVM method is obtained with the highest level of accuracy. In the case study of twitter sentiment analysis in the FIFA 2013, the SVM method provides a higher level of accuracy compared to the naïve bayer classifier [5].

Many previous studies focused on classification methods such as support vector machine (SVM), naïve bayes, and maximum entry to classify tweets. Coletta, Hruschka, Felix, & Hruschka [6], combine SVM classifier and C3E-SL clustering technique and produce a better classification than just using the SVM method. The process of tweeting clustering is generally done by k-means and also non-negative matrix factorization (NMF). When both results are compared they have similar results, but NMF can run faster and the results are easier to interpret [6].

Research that discusses about stock market predictions on Indonesia Stock Exchange (IDX) with sentiment analysis on Twitter has not been done much. In fact, market sentiment analysis is one of the benchmarks of stock price movements. Furthermore, there are not many sentiment analyzes using the support vector machine method as a predictive method and non-negative matrix factorization in the feature extraction process. Therefore, in this study, twitter sentiment analysis done for predicting stock price movements by combining support vector machine methods and non-negative matrix factorization methods.

The main objective of this research is to build a model by applying two methods, namely support vector machine and nonnegative matric factorization. For increasing the accuracy and performance of the sentiment analysis method on twitter for predicting stock prices on Indonesia Stock Exchange.

2. Related Work

Sentiment analysis is one of social media mining applications that can be used to find out sentiments. Niu, Yin, & Kong [8] research on the analysis of commitment to microblogging. In this study, they increase the ability of the Bayesian text classifier method to select word features, determine

* Abba Suganda Girsang, Email: agirsang@binus.edu

the weight and classification of commitment. One application of sentiment analysis is a prediction of stock price movements. Sentiment analysis is related to the increase and decrease in stock prices [4]. Rao & Srivastava [9], implemented the Naïve Bayes classifier in his research.

The sentiment analysis process can be done with tools such as OptionFinder that produce positive and negative, Google Profile of Mood States (GPOMS) that produce calm, alert, sure, vital, kind and happy. The results of these tools were analyzed using granger casuality and self-organizing fuzzy neural networks resulting in 86.7% accuracy in predicting increases and decreases of DJIA and MAPE reduction of more than 6% [2].

Jadhav & Wakode [1] make comparisons of sentiment classification techniques including SVM, NB, NBSVM, MNB, SentiStrengh + Twitter Sentiment, SEntiStrength and Twitter Sentiment. From the comparison result, SVM method is obtained the highest level of accuracy. The method that is widely used in sentiment analysis process is naïve bayes classifier and support vector machine. Sentiment analysis process in case study of FIFA 2013, the SVM method provides a higher level of accuracy compared to the naïve bayer classifier [5].

Many previous studies focused on classification methods such as support vector machine (SVM), naïve bayes, and maximum entopy to classify tweets. The combination of text clustering and text classifier increase the accuracy of tweets classifier. The combination of SVM classifier and C3E-SL clustering technique can correct tweets classification, compared by only using SVM method [7]. Zhu, Jing, & Yu [10] in their study stated that NMF has a good performance used in real-world data texts. Shahnaz, Berry, Pauca, & Plemmons [11], done reasearch for cluster documents using non-negative matrix factorization. Tweet clustering can be done by k-means and also non-negative matrix factorization (NMF), but NMF can run faster and the results are easier to interpret [6]. The research conducted by Nguyen, T. H., Shirai, K., & Julien Velcin, 2015 which analyze stock market movements using close price, volume, sentiment analysis and topics as features.

3. Method

Figure 1 is an overview of the research methods in this study.

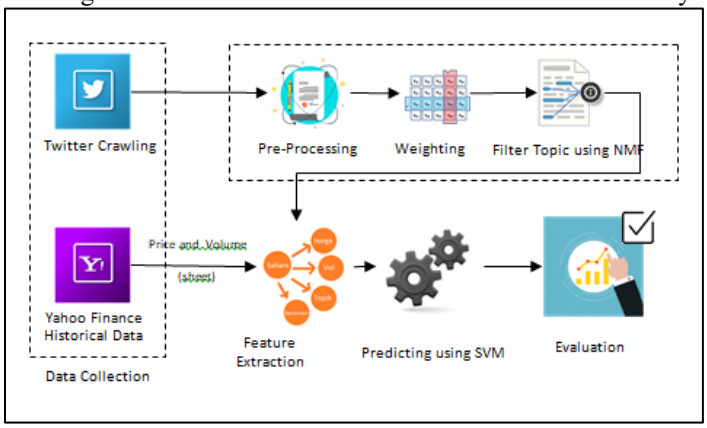


Figure 1: Research Method

3.1. Data Collection

The tweet used comes from verified account on Twitter. The list of companies in this study are companies that included in the LQ 45 period from August 2018 - January 2019, including Adhi Karya (Persero) Tbk., Adaro Energy Tbk., AKR Corporindo Tbk., Aneka Tambang (Persero) Tbk., Astra International Tbk. ,Bank Central Asia Tbk., Bank Negara Indonesia (Persero) Tbk., Bank Rakyat Indonesia (Persero) Tbk., Bank Tabungan Negara (Persero) Tbk., BPD Jawa Barat dan Banten Tbk., Sentul City Tbk, Bank Mandiri (Persero) Tbk., Barito Pasific Tbk., Bumi Serpong Damai Tbk., Elnusa Tbk., XL Axiata Tbk., Gudang Garam Tbk., H.M. Sampoerna Tbk., Indofood Sukses Makmur Tbk., Indika Energy Tbk, Indah Kiat Pulp & Paper Tbk., Indocement Tunggal Prakarsa Tbk., Indo Tambangraya Mehag Tbk., Jasa Marga (Persero) Tbk., Kalbe Farma Tbk., Lippo Karawaci Tbk., Matahari Department Store Tbk., Medco Energi Internasional Tbk., Media Nusantara Citra Tbk., Perusahaan Gas Negara (Persero) Tbk., Tambang Batu Bara Bukit Asam (Persero) Tbk., PP Properti Tbk., Surya Citra Media Tbk., Semen Indonesia (Persero) Tbk., Sri Rejeki Isman Tbk., Sawit Sumbermas Sarana Tbk., Telekomunkasi Indonesia (Persero) Tbk., Chandra Asri Petrochemical Tbk., United Tractors Tbk., Unilever Indonesia Tbk., Wijaya Karya (Persero) Tbk., Waskita Beton Precast Tbk., dan Waskita Karya (Persero) Tbk [8]. From 45 companies in the LQ45 list, only 15 companies have verified accounts, with 24 accounts. Table 1 is the list of stock code and twitter users.

Table 1: List of Stock Code and Twitter Users

Stock Code	Twitter Username
BBCA	bankbca
	HaloBCA
BBNI	bni
	BNICustomerCare
BBRI	BANKBRI_ID
	kontakBRI
BBTN	BankBTNcoid
BMRI	bankmandiri
	mandiricare
EXCL	XLAxiata_Tbk
	myXLCare
	myXL
INDF	indofood
JSMR	PTJASAMARGA
	OFFICIAL_JSMR
KLBF	KALBEfamily
LPPF	GayaMatahari
PGAS	Gas_Negara
SCMA	IndosiarID
	SCTV
SMGR	semenku
TLKM	TelkomCare
	TelkomIndonesia
UNVR	UnileverIDN

3.2. Pre-processing

The pre-processing process is performed to eliminate noise, and data normalization. Data must be processed to improve the performance of text mining process.

Example: *@TelkomIndonesia Berkebun itu kegiatan yang mengasyikan lho Sobat! #TelkomEdu*

Figure 2: Pre-processing

1. Case Folding

The first step is case folding. The initial stage is the cleansing process, then convert all character into a lower case.
Example :*telkomIndonesia berkebun itu kegiatan yang mengasyikan lho sobat telkomedu*

2. Tokenizing

At this stage, the sequence of words in the tweet is cut, and form to pieces of words according to system requirements.
Example: *telkomIndonesia* | *berkebun* | *itu* | *kegiatan* | *yang* | *mengasyikan* | *lho* | *sobat* | *telkomedu*

3. Filtering

Filtering / Stopwords is a process of removing characters, punctuation, and general words that have no meaning.
Example: *berkebun* | *kegiatan* | *mengasyikan*

4. Stemming

Stemming process of reducing a word to its word stem that affixes to suffixes and prefixes or to the roots of words known as a lemma.
Example: *kebun* | *giat* | *asik*

3.3. Filter Topic Using NMF

Non-Negative Matrix Factorization (NMF) gives a non-negative matrix V to produce a non-matrix factor *W* and *H* as Eq. (1).

$$V \approx W\,H \tag{1}$$

At first the values of the *W* and *H* matrices are random, then the matrix values are updated to get the results that are getting closer to the *V* matrix.

The *H* matrix is updated using Eq. (2)

$$H = H * \frac{(W'V)}{(W'WH)+\epsilon} \tag{2}$$

The W matrix is updated using Eq. (3)

$$W = W * \frac{(VH')}{(WHH'+\epsilon)} \tag{3}$$

3.4. Feature Extraction

Feature extraction is important to get stock daily features. Data that used in this step are from Yahoo Finance and the results of the Twitter API. Nguyen, Shirai, & Julien Velcin [13], have compared the features used in the stock prediction process, namely Price, Topics and Sentiment analysis scores. In this study, researchers used volume as an additional feature in the stock prediction process. The features used in this study are in Table 2.

Table 2: List of Features

Features	Desciption
Pc_{t-2}	Close Price T-2
Pc_{t-1}	Close Price T-1
Vol_{t-2}	**Volume T-2**
Vol_{t-1}	**Volume T-1**
PT_{t-2}	Topic percentage T-2
PT_{t-1}	Topic percentage T-1
Sen_{t-2}	Sentiment score T-2
Sen_{t-1}	Sentiment score T-1

3.5. Predicting using SVM

The learning method for predicting stock movement is SVM. SVM tries to find the best line dividing the two classes, and then classifies the test documents based on which side of the line they appear. The features used in predicting stock movements are illustrated in table 3.

Table 3: Illustration of SVM Application Method

	Date	Pc_{t-2}	Pc_{t-1}	Vol_{t-2}	Vol_{t-1}	PT_{t-2}	PT_{t-1}	Sen_{t-2}	Sen_{t-1}	Y
Data Train	01072016	2000	2010	3000	4000	30	40	1	-1	1
	…									
	31032017									
Data Test	01042017									
	…									

In table 3 it is illustrated that, the prediction results on stock movements are *Y*. Where the value of *Y* is between [0, 1]. Every day, it only has 1 data in predicting stock market movements. Where the data is divided into two, 25% data test and 75% data train. The results of stock movements are divided into two classes, such as up and down.

3.6. Evaluation

The evaluation method that used in this study was confusion matrix. Confusion matrix is a method used to calculate accuracy in the concept of data mining [9]. The evaluation results with confusion matrix are the values of accuracy, precision and recall. Accuracy in classification is the percentage of data record provisions that are classified correctly after testing the classification results [10]. Precision or confidence is the proportion of cases predicted to rise which are also positively correct in the actual data. Recall or sensitivity is the proportion of actual positive cases that are correctly predicted positively [11] .

Table 4: Confusion Matrix Model

Correct Classification	Classified as	
	Up	Down
Up	TP	FP
Down	FN	TN

Eq. (4), Eq. (5), Eq. (6) and Eq.(7) are the formula for calculate Accuracy, Precision, Recall and Error Rate, respectively.

$$Accuracy = \frac{\text{TP+TN}}{All\ Value} \tag{4}$$

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \tag{5}$$

$$Recall = \frac{\text{TP}}{\text{TP+FN}} \tag{6}$$

$$Eror\ Rate = 1 - Accuracy \tag{7}$$

Table 5: Twitter Data Results

Stock Code	Twitter Username	Number of tweets
BBCA	bankbca	6224

	HaloBCA	21718
BBNI	bni	8848
	BNICustomerCare	8675
BBRI	BANKBRI_ID	6090
	kontakBRI	10981
BBTN	BankBTNcoid	435
BMRI	bankmandiri	14073
	mandiricare	13424
EXCL	XLAxiata_Tbk	1376
	myXLCare	12587
	myXL	9510
INDF	indofood	398
JSMR	PTJASAMARGA	25191
	OFFICIAL_JSMR	1241
KLBF	KALBEfamily	1733
LPPF	GayaMatahari	2416
PGAS	Gas_Negara	2982
SCMA	IndosiarID	19791
	SCTV_	11546
SMGR	semenku	1122
TLKM	TelkomCare	29971
	TelkomIndonesia	9327
UNVR	UnileverIDN	862
	Total Tweets	220521

Table 6: Historical Data Result

Stock Code	Number of days
BBCA	174
BBNI	174
BBRI	174
BBTN	174
BMRI	174
EXCL	174
INDF	174
JSMR	174
KLBF	174
LPPF	174
PGAS	174
SCMA	174
SMGR	174
TLKM	174
UNVR	174
Total	2610

4. Data Analysis and Results

There are two data used in this study, such as stock price historical data from Yahoo Finance, and data tweets taken using NodeXL Pro. This study only processes data from 1 August 2018 to 1 April 2019. Table 5 is twitter data results.

Tabel 6 is historical data taken using Yahoo Finance from 1 August 2018 to 1 April 2019.

All tweets will go through preprocessing data tweet stage. This process is done to get data that has been clean / normalized. The steps taken in this process are divided into several steps including cleansing, case folding, tokenizing, stopwords and stemming. This step is done using the Python nltk library and Sastrawi.

4.1. Topic Extraction

At this stage, each tweet will be classified by using the sklearn library in python with Non-Negative Matrix Factorization (NMF) algorithm to get a daily topic. Before classification, each tweet through a weighting process using TfidfVectorizer. Tweets that are processed in this stage are tweets that have through preprocessing stage. Non-Negative Matrix Factorization (NMF) is used to find two non-negative matrices (W, H) whose results are close to the non-negative matrix V [12]. This factorization process used for dimension reduction, source separation or topic extraction. The first process in NMF is created a weight matrix (V) to produce a matrix W and H. Table 7 is preprocessed tweets that used for create matrix V.

Table 7: Preprocessed tweets

Num.	Tweet
T1	terima kasih
T2	mohon tunggu konfirmasi laku telepon ya
T3	transaksi finansial periode libur lebaran
T4	khawatir cabang bca daerah buka layan operasional batas libur idul fitri
T5	lapor proses kait mohon tunggu proses jalan ya
T6	hai langsung terima transfer rekening bca rekening bank ya h 1 ya
T7	selamat malam aplikasi saku mudah download aplikasi google play app store ya informasi lengkap klik link
T8	sih hati suci mari maaf fitri selamat raya idul fitri 1439 h
T9	hai top saldo flazz saku transaksi saku saku plus kena biaya informasi lengkap putar saku klik link ya terima kasih

The first process in NMF is to create a weight matrix (V) to produce a matrix W and H. Table 8 is matrix V is the result of a weighting process with TfidfVectorizer.

Table 8: Matrix V

Term \ Tweet	T1	T2	T3	T4	T5	T6	T7	T8
1439	0	0	0	0	0	0	0.2	0.
aplikasi	0	0	0	0	0	0.5	0	0.
app	0	0	0	0	0	0.2	0	0.
bank	0	0	0	0	0.31	0	0	0.
batas	0	0	0.3	0	0	0	0	0.
bca	0	0	0.2	0	0.26	0	0	0.
biaya	0	0	0	0	0	0	0	0.20
buka	0	0	0.3	0	0	0	0	0.
cabang	0	0	0.3	0	0	0	0	0.
daerah	0	0	0.3	0	0	0	0	0.
download	0	0	0	0	0	0.2	0	0.
fitri	0	0	0.2	0	0	0	0.5	0.
flazz	0	0	0	0	0	0	0	0.20
google	0	0	0	0	0	0.2	0	0.
hai	0	0	0	0	0.26	0	0	0.17
hati	0	0	0	0	0	0	0.2	0.
idul	0	0	0.2	0	0	0	0.2	0.
informasi	0	0	0	0	0	0.2	0	0.17
jalan	0	0	0	0.33	0	0	0	0.
kait	0	0	0	0.33	0	0	0	0.
kasih	0.74	0	0	0	0	0	0	0.15
kena	0	0	0	0	0	0	0	0.20
khawatir	0	0	0.3	0	0	0	0	0.
klik	0	0	0	0	0	0.2	0	0.17
konfirmas	0	0.4	0	0	0	0	0	0.
laku	0	0.4	0	0	0	0	0	0.
langsung	0	0	0	0	0.31	0	0	0.
lapor	0	0	0	0.33	0	0	0	0.
layan	0	0	0.3	0	0	0	0	0.
lengkap	0	0	0	0	0	0.2	0	0.17
libur	0	0	0.2	0	0	0	0	0.
link	0	0	0	0	0	0.2	0	0.17
maaf	0	0	0	0	0	0	0.2	0.
malam	0	0	0	0	0	0.2	0	0.

mari	0	0	0	0	0	0	0.2	0.
mohon	0	0.3	0	0.28	0	0	0	0.
mudah	0	0	0	0	0	0.2	0	0.
operasion	0	0	0.3	0	0	0	0	0.
play	0	0	0	0	0	0.2	0	0.
plus	0	0	0	0	0	0	0	0.20
proses	0	0	0	0.67	0	0	0	0.
putar	0	0	0	0	0	0	0	0.20
raya	0	0	0	0	0	0	0.2	0.
rekening	0	0	0	0	0.62	0	0	0.
saku	0	0	0	0	0	0.2	0	0.69
saldo	0	0	0	0	0	0	0	0.20
selamat	0	0	0	0	0	0.2	0.2	0.
sih	0	0	0	0	0	0	0.2	0.
store	0	0	0	0	0	0.2	0	0.
suci	0	0	0	0	0	0	0.2	0.
telepon	0	0.4	0	0	0	0	0	0.
terima	0.66	0	0	0	0.20	0	0	0.13
top	0	0	0	0	0	0	0	0.20
transaksi	0	0	0	0	0	0	0	0.17
transfer	0	0	0	0	0.31	0	0	0.
tunggu	0	0.3	0	0.28	0	0	0	0.
ya	0	0.2	0	0.2	0.37	0.1	0	0.12

Then, after the V matrix initialize the NMF function. Researchers determined 3 daily topics for every stock code. Table 9 matrix H that formed in python using the function fit_transform.

Table 9: Matrix H

Term \ Topic	K1	K2	K3
1439	0.	0.	0.13729
aplikasi	0.	0.	0.
app	0.	0.	0.
bank	0.	0.	0.
batas	0.	0.	0.13438
bca	0.	0.	0.10971
biaya	0.	0.	0.
buka	0.	0.	0.13438
cabang	0.	0.	0.13438
daerah	0.	0.	0.13438
download	0.	0.	0.
fitri	0.	0.	0.39452
flazz	0.	0.	0.
google	0.	0.	0.
hai	0.	0.	0.
hati	0.	0.	0.13729
idul	0.	0.	0.25211
informasi	0.	0.	0.
jalan	0.	0.15815	0.
kait	0.	0.15815	0.
kasih	0.871	0.	0.
kena	0.	0.	0.
khawatir	0.	0.	0.13438
klik	0.	0.	0.
konfirmasi	0.	0.22982	0.
laku	0.	0.22982	0.
langsung	0.	0.	0.
lapor	0.	0.15815	0.
layan	0.	0.	0.13438
lengkap	0.	0.	0.
libur	0.	0.	0.10971
link	0.	0.	0.
maaf	0.	0.	0.13729
malam	0.	0.	0.
mari	0.	0.	0.13729
mohon	0.	0.35382	0.
mudah	0.	0.	0.
operasional	0.	0.	0.13438
play	0.	0.	0.
plus	0.	0.	0.
proses	0.	0.35058	0.
putar	0.	0.	0.
raya	0.	0.	0.13729
rekening	0.	0.	0.
saku	0.041	0.	0.
saldo	0.	0.	0.
selamat	0.	0.	0.11218
sih	0.	0.	0.13729
store	0.	0.	0.
suci	0.	0.	0.13729
telepon	0.	0.22982	0.
terima	0.783	0.	0.
top	0.	0.	0.
transaksi	0.	0.	0.
transfer	0.	0.	0.
tunggu	0.	0.35382	0.
ya	0.	0.25208	0.
1439	0.	0.	0.13729
aplikasi	0.	0.	0.
app	0.	0.	0.
bank	0.	0.	0.
batas	0.	0.	0.13438
bca	0.	0.	0.10971

Table 10 is matrix W that formed by components_ function.

Table 10: Matrix W

	T1	T2	T3	T4	T5	T6	T7	T8
K1	0.79	0	0	0	0.08	0	0	0.15
K2	0	0.84	0	0.83	0.06	0	0	0
K3	0	0	0.86	0	0	0	0.93	0

The topic extraction result is used as a daily topic percentage in stock features. Furthermore, the topic value is also used to obtain tweet sentiment values that are included in the highest daily topics. Eq. (8) is the formula to get the topic percentage value.

$$Topic\ Percentage = \frac{Total\ Topic}{Total\ Tweets} \quad (8)$$

Table 11 is an illustration of the calculation results for the percentage value of the topic. From the calculation of the percentage of topics, it was found that the highest topic on that day was topic 1 (K1), with the number of tweets as many as 5. Tweets that are incorporated in topic 1 (K1) are T1, T5, T6, and T8 which will then be processed to obtain the value of his sentiment.

Table 11: Topic Percentage Calculation

	Topic Percentage
K1	4/8 = 0.5
K2	2/8 = 0.25
K3	2/8 = 0.25

4.2. Sentiment Analysis

Sentiment analysis was done after preprocessing, where the tweets that be analyzed only tweeted that have the highest topic value on that day. This is done to reduce the number of tweets processed. At this stage the researcher uses LinearSVM in library sklearn SVM Python. The list of negative and positive words used refers to the list of words at https://github.com/masdevid/ID-OpinionWords (Masdevid, 2017). At this stage each tweet in the highest topic on that day will be processed to get the value of sentiment, and its probability. Table 12 is the result of sentiment analysis using Linear SVM.

The results of the probability will be taken as a feature in the formation of sentiment features in the next stage. The Eq. (9) is a formula to get sentiment values.

$$Sentiment = \frac{Total\ Value\ of\ Positif - Total\ Value\ of\ Negatif}{Total\ Topic} \quad (9)$$

where total value of positive = total percentage value of positive sentiment, total value of negative = total percentage value of negative sentiment obtained from the sentiment analysis process

Table 12: Sentiment Analysis Result

Topics	Tweets	Sentiment	Probability
T1	terima kasih	0	0.7508999502661317
T5	hai langsung terima transfer rekening bca rekening bank ya h 1 ya	0	0.7137279798102268
T6	selamat malam aplikasi saku mudah download aplikasi google play app store ya informasi lengkap klik link	0	0.7278331821231244
T8	hai top saldo flazz saku transaksi saku saku plus kena biaya informasi lengkap putar saku klik link ya terima kasih	0	0.7946268383337876

The Eq. (10) is the result of calculation of sentiment values according to the formula (9).

$$Sentiment = \frac{(0.7509 + 0.7137 + 0.7278 + 0.7946) - (0)}{4} = \frac{2.987}{4} = 0.74675 \quad (10)$$

4.3. Formation of Stock Features

After obtaining the topic value and tweet sentiment, then we can proceed by forming a stock feature consisting of closing price at T-2, closing price at T-1, closing volume at T-2, closing volume at T-1, total percentage of the highest topic on T-2, total percentage of the highest topic on T-1, sentiment score on T-1, and sentiment score on T-2.

Table 13 is a form of daily stock data obtained from merging price data from Yahoo Finance with the results of sentiment extraction with LinearSVM and topics with NMF.

Table 13: Stock Daily Data of BBRI

Date	Close	Volume	Sentiment	Topic
8/1/2018	3190	147098300	0.380237	0.5
8/2/2018	3250	158406800	0.13912	0.492647
8/3/2018	3330	100589700	0.009038	0.568627
8/6/2018	3410	143310900	0.261233	0.52809
8/7/2018	3350	89404000	0.449981	0.777778
8/8/2018	3330	90710200	0.485755	0.396825
8/9/2018	3330	69654300	0.312497	0.571429
8/10/2018	3390	115961500	0.142377	0.534247
8/13/2018	3140	179328600	0.443025	0.46
8/14/2018	3130	154831700	0.119308	0.565217
8/1/2018	3190	147098300	0.380237	0.5
8/2/2018	3250	158406800	0.13912	0.492647
8/3/2018	3330	100589700	0.009038	0.568627

4.4. Support Vector Machine Process

The features that have been formed from Price, Volume, Sentiment and Percentage of topics per day are processed using Python Support Vector Machine library, which is the library from sklearn library in Python. The data collection is divided into 75% training data and 25% testing data for predicting stock price movements with predictive results down and up. Table 14 is the prediction of stock price movements result.

Table 14 : *Stock Movement Prediction Results*

Stock Code	RBF	POLY	LINEAR
BBCA	0.5813953488	0.4186046512	0.4418604651
BBNI	0.6046511628	0.5581395349	0.6511627907
BBRI	0.5348837209	0.4651162791	0.4186046512
BBTN	0.5813953488	0.5813953488	0.6046511628
BMRI	0.6046511628	0.3953488372	0.4418604651
EXCL	0.5348837209	0.4651162791	0.5813953488
INDF	0.5348837209	0.4651162791	0.5813953488
JSMR	0.7209302326	0.2790697674	0.6744186047
KLBF	0.6511627907	0.3488372093	0.488372093
LPPF	0.7441860465	0.2558139535	0.5581395349
PGAS	0.5813953488	0.4186046512	0.511627907
SCMA	0.6976744186	0.3023255814	0.4418604651
SMGR	0.5348837209	0.4418604651	0.5581395349
TLKM	0.6744186047	0.3255813953	0.5813953488
UNVR	0.4418604651	0.5348837209	0.6279069767
Average	0.6015503876	0.4170542636	0.5441860465

4.5. Comparison with The Previous Research

In a previous study conducted by Nguyen, Shirai, & Julien Velcin [13], the features were closing prices, sentiments and topic percentages. The addition of volume features can increase the accuracy value to 60.16%, while the previous feature is 57.21%. Table 15 is a comparison of stock movements with previous research.

Table 15 : Comparison Accuracy Results

Features	RBF	POLY	LINEAR
Close Price, Sentiment, Topic [13]	0.572	0.479	0.491
Close Price, Sentiment, Topic and **Volume**	0.602	0.417	0.544

The predictions results of stock market movements carried out using volume feature have a higher accuracy value compared to previous research without volume feature. LI & Zhu [19] comparing the performance of MA and VMA it was found that the application of technical indicators that used volume information is more effective than the pure price index on the stock market. Furthermore, Investopedia [20] said volume is an important indicator in technical analysis because it is used to measure the value of market movements. If the market has made a strong price move up or down, it can depend on the volume for that period. The higher volume during the price movement, the more significant movement. The analysis of the authors that underlies the addition of volume usage in this study is that if there is a high increase in volume on the previous day, then it can be interpreted that there are many investor's interests in that stock so it can affect the next day price. Other, result of Pearson correlation analysis is 0,1857, for value under 0.5 said that it has weak relation between price and volume.

5. Conclusion

The features proposed in this study are closing price at T-2, closing price at T-1, closing volume at T-2, closing volume at T-1, total percentage of the highest topic on T-2, total percentage of the highest topic on T-1, sentiment score on T-1, and sentiment score on T-2. The process of predicting stock market movements using SVM with the RBF kernel has the highest level of accuracy. It is proven that the addition of volume features can increase the accuracy value to 60.16%.

References

[1] R. Jadhav and M. S. Wakode, "Survey : Sentiment Analysis of Twitter Data for Stock Market Prediction," *IJARCCE Vol. 6, Issue 3,* pp. 558-562, 2017.

[2] J. Bollen, H. Mao and X.-J. Zeng, "Twitter mood predicts the stock market," *IEEE Computer,* pp. 91-94, 2010.

[3] Bursa Efek Indonesia, "Sekilas Saham," 2017. [Online]. Available: http://yuknabungsaham.idx.co.id/sekilas-saham-detail.php. [Accessed 23 Juni 2017].

[4] D. Sonachalam, "Using Twitter to predict Stock Market Returns (August 2015)," *International Journal of Scientific & Engineering Research, Volume 6, Issue 10, October-2015,* pp. 76-77, 2015.

[5] A. L. Alves, C. d. S. Baptista, A. A. Firmino, M. G. d. Oliveira and A. C. d. Paiva, "A Comparison of SVM Versus Naive-Bayes Techniques for Sentiment Analysis in Tweets: A Case Study with the 2013 FIFA Confederations Cup," *WebMedia '14 Proceedings of the 20th Brazilian Symposium on Multimedia and the Web,* pp. 123-130, 2014.

[6] L. F. Coletta, E. R. Hruschka, N. Felix and E. R. Hruschka, "Combining Classification and Clustering for Tweet Sentiment Analysis," *Brazilian Conference on Intelligent Systems (BRACIS-2014),* pp. 210-215, 2014.

[7] D. Godfrey, C. Johns, C. Sadek, C. Meyer and S. Race, "A Case Study in Text Mining: Interpreting Twitter Data From World Cup Tweets," *arXiv preprint arXiv:1408.5427,* 2014.

[8] Z. Niu, Z. Yin and X. Kong, "Sentiment Classification for Microblog by Machine Learning," *Fourth International Conference on Computational and Information Sciences,* pp. 286-289, 2012.

[9] T. Rao and S. Srivastava, "Analyzing Stock Market Movements Using Twitter Sentiment Analysis," *ASONAM '12 Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012),* pp. 119-123, 2012.

[10] Y. Zhu, L. Jing and J. Yu, "Text Clustering via Constrained Nonnegative Matrix Factorization," *11th IEEE International Conference on Data Mining,* pp. 1278-1283, 2011.

[11] F. Shahnaz, M. W. Berry, V. P. Pauca and R. J. Plemmons, "Document clustering using nonnegative matrix factorization," *Information Processing and Management 42,* pp. 373-386, 2006.

[12] Sinarmas Sekuritas, "Unduh Daftar Saham LQ45," Agustus 2018. [Online]. Available: http://www.sinarmassekuritas.co.id/uploads/filelist/2018/daftar-saham-lq-45_1.pdf. [Accessed 1 Januari 2018].

[13] T. H. Nguyen, K. Shirai and Julien Velcin, "Sentiment Analysis on Social Media for Stock Movement Prediction," *Expert Systems With Applications Volume 42, Issue 24,* pp. 9603-9611, 2015.

[14] P. Mayadewi and E. Rosely, "Prediksi Nilai Proyek Akhir Mahasiswa menggunakan Algoritma Klasifikasi Data Mining," *Seminar Nasional Sistem Informasi Informasi 2-3 November 2015,* pp. 329-334, 2015.

[15] J. Han and M. Kamber, Data Mining Concepts & Techniques 2nd Edition, San Fransisco: Elsevier, 2006.

[16] D. Powers, "Evaluation: From Precision, Recall, and F-Measure to ROC, Infomedness, Markedness & Correlation," *Journal of Machine Learning Technologies,* pp. 37-63, 2011.

[17] Scikit-learn, "Scikit-learn Documentation: sklearn.decomposition.NMF," 2011. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.NMF.html. [Accessed 3 3 2019].

[18] Masdevid, "List of Opinion Words (positive/negative) in Bahasa Indonesia for Sentiment Analysis," 24 March 2017. [Online]. Available: https://github.com/masdevid/ID-OpinionWords. [Accessed 20 3 2019].

[19] G. LI and J. Zhu, "Research on the Effectiveness of Technical Indicators with the Volume," *International Conference on Education, Management and Computing Technology (ICEMCT 2014),* pp. 436-439, 2014.

[20] Investorpedia, "Volume Definition," 10 May 2019. [Online]. Available: https://www.investopedia.com/terms/v/volume.asp. [Accessed 10 May 2019].

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 251-259 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

Proposal of Ledger Technology to Apply to a Public Organization in Ecuador

Segundo Moisés Toapanta Toapanta[*,1], Adrian Alberto Chávez Monteverde [1], Javier Gonzalo Ortiz Rojas[1], Luis Enrique Mafla Gallegos [2]

[1]*Departament of Engineering Systems, Universidad Politécnica Salesiana (UPS), Guayaquil, Ecuador*

[2]*Faculty of Engineering Systems, Escuela Politécnica Nacional del Ecuador (EPN), Quito, Ecuador*

ARTICLE INFO

Article history:
Received:02 April, 2019
Accepted:19 June, 2019
Online: 30 June, 2019

Keywords:
Public Organization
Ecuador
Blockchain
Ledger Technology
Algorithm
Database
Distributed System

ABSTRACT

Ledger technology has now changed financial systems around the world, applying this technology to the improvement of the public sector, mainly to the debt collection system, we improve the speed of payment and its immediate registration, we make use of the advantages of Ledger technology with respect to reliability, security and speed. The objective is to propose how we apply Ledger technology within a public organization, considering the improvement of this, without diminishing its efficiency, security and current productivity. We use a deductive method to analyze the information obtained from the scientific articles reviewed. The results of our research are reflected in an algorithm that shows the operation and application of Ledger technology. It was concluded that the strengths of the algorithm, together with the good distribution and application of Ledger technologies solve problems regarding efficiency, safety and savings.

1. Introduction

Today, the public sector is seen as an ineffective, ineffective service, and citizens need better performance from the public sector. The public service in Ecuador many years ago needs a very profound change that is effective and demonstrable in the services offered to citizens, because of this reason we can find solutions in various technological fields, we focus on a solution offered by Ledger technology applied to a public organization.

Technology advances day by day to immeasurable levels from which we can obtain benefits, some countries have begun to implement Ledger technology in their systems of collection of debts, goods, etc. In the countries where Ledger technology is implemented, the results have become positive in every field in which it was applied, showing results that can be used as an example to achieve objectives in Ledger technology applications.

Today the public organizations that manage the debts with the citizenry are collected through banks and other private financial institutions that provide their service and that has a high degree of security and reliability that offers to the citizens, but for the transaction to take place an intermediary in the collection of the debt, which is the private sector, because of this it takes away the speed of the payment process and the reflection thereof between the citizen and the private financial entity and this in turn with the public entity, which generates an additional cost that is paid by the citizen [1].

If the current system is robust, safe and effective, why should Ledger technology be deployed to a public organization? The current way to pay a debt to a public entity is through private sector entities which has a cost that is charged by the private sector intermediary and paid by the citizen this generates in an increase in the cost of transaction in total and should also be added the protection to our information, the private sector have possession of this, can make use of it or at worst have leaks or thefts of our private information.

Ledger technology helps us reduce costs by paying for transactions, which is a saving for the average citizen, and save budget for hardware needed by the public sector to have the current collection system run between a public organization and a private organization, which gives us robustness [2], in the case of the application of the Ledger technology gives us an advantage that the payment is made directly and anonymously with the participation of the citizen and the public organization, without offering our personal and bank information to a private entity, which gives us security and efficiency [3].

[*] Segundo Moisés Toapanta Toapanta, Email: stoapanta@ups.edu.ec

An important advantage is the saving of money, both by the citizen who saved the cost by performing the transaction through the private sector, also the public institution will save on hardware, because the processing cost was shared by all participating nodes within the network, this way less hardware was needed for the network to work also decreases costs in different areas: hardware cost, electrical power, personnel capable of handling specialized hardware [3].

If the benefits are higher, then why is the study and implementation not applied? Because the change that is taking place could generate conflicts for the citizen, by explaining all the functions and advantages that a Ledger technology would have in applying it, the beneficiaries can resist the change, so it would be a problem because people do not want to participate in the application of the Ledger technology, another problem is the change of certain functions in the public organization, leaving aside some traditional actors but, In the same way, new participants enter the collection system. For this reason, it is necessary to create a new definition of the responsibilities of the infrastructure and of the participants in the provision of collection services.

The articles reviewed for this research are:

Centralized solution to securely transfer payment information electronically to banks from multiple enterprise resource planning (ERP) systems [1]: The model is basically composed of ERP business resources planning applications, this model offers security, effectiveness and cost savings of up to 75% without geographical restrictions. The main users are large business units that frequently carry out banking or financial transactions such as payment to local or foreign suppliers, payroll, etc., thus having a relationship with our system due to the collection and payment of debts, which instead of being a bank, it will be a public organization, from which we obtain an implementation already applied to financial systems, in this way we will be able to obtain results according to the exposed thing guaranteeing in our systems equal or superior results.

Lightweight and Manageable Digital Evidence Preservation System on Bitcoin [2]: We found a structure for the preservation of light digital evidence that has the characteristics of privacy-anonymity, audit-transparency, function-scalability and light-operation, such characteristics that we will apply in our system, obtaining information from studies conducted using bitcoin systems (cryptocurrency) which helps our system to use virtual money and its characteristics by applying them in the correct way guaranteeing results.

Towards dependable, scalable, and pervasive distributed ledgers with blockchains [3]: We find the distributed general ledger technology (DLT), its structure, classification and applications in three generations: 1.0 (cryptocurrency), 2.0 (Apps) and 3.0 (omnipresent applications). It presents all aspects of the blokchain, showing how the block chain systems are balanced, from which we have learned and used the study done to the blockchain, which allows us to apply it in the best way to our system, having the knowledge offered by blockchain.

Redecentralizing the Web with Distributed Ledgers [4]: Presents that the contracts or accounting books distributed represent a service of reliability, responsibility and security in transactions without the need for centralized validation authorities, projecting the web as a true decentralized autonomous system, which helps us to understand the benefits offered by the application of daily books and at the same time be able to offer the advantages of the same to our implementation.

CoC: Secure Supply Chain Management System Based on Public Ledger [5]: CoC (supply chain on blockchain) is a supply chain management system that provides a security mechanism to circumvent any access that is not registered as authorized to the general ledger database, since the general ledger generally lacks security, we will use this method of building blocks within our system to be able to make use of the advantages of it, This way applying a more level of security to our locks that handles the default Ledger technologies.

PQChain: Strategic design decisions for distributed ledger technologies against future threats [6]: The importance of a well-structured strategy for an appropriate chain of blocks is highlighted, when instances are created means that a large number of participants must be dealt with and it is the trust in the centralized authorities that determines the security guarantees provided by the cryptography, in this way we have created and structured good chains of blocks to be able to manage within the network, in this way we ensure the security of each block received and sent by the network for the benefit of the application within the public organization.

Distributed Ledger [7]: It demonstrates a systematic description of the most remarkable principles of the DL field, because there is currently no structural approach and definition of DL, from which we can learn clear concepts that help us to apply Ledger technologies to our system, which we take advantage of so that knowing the fundamentals we can succeed in implementation.

Security of Distributed Ledger Solutions Based on Blockchain Technologies [8]: Provides information on the security aspects of blockchain technology, identifies the most relevant security threats and challenges for technology development, with these studies we will be able to apply them to provide the necessary security to be able to apply the blockchain, through this study we will be able to understand and apply the security that we will have to offer to our entire system for a successful application.

Distributed Ledger Technology: Blockchain Compared to Directed Acyclic Graph [9]: Develops a new model in the field of distributed accounting technology, it is compared with tools like Ethereum, Bitcoin and Nano, with the blockchain being the two most important paradigms for distributed systems. In order to achieve the reduction in the size of the ledger and the effectiveness of the transactions that are generated, we obtain the comparison with the blockchain, In this way we can know advantages and disadvantages compared to the blockchain and by this comparison we can obtain better results in the application of Ledger technologies.

Blockchain-based Proof of Delivery of Physical Assets with Single and Multiple Transporters [10]:. It presents a solution for the Proof of Delivery (PoD) of the physical assets negotiated based on the blockchain. They use the aspects of the chain of blocks Ethereum so that payments are automated and generate records that can identify fakes, all this in order to gain reliability and

present transparency, we can find a relationship in the handling of money to be able to carry out transactions within ledger technologies, using electronic money (cryptocurrencies) and in turn understand the benefits that it guarantees with the use of cryptocurrencies which we take advantage of in our application.

A Low Storage Room Requirement Framework for Distributed Ledger in Blockchain [11]: It proposes to reduce the storage space required by the blockchain to store all transactions, using a coded network (NC) and distributed storage (DS): NC-DS, which proposes to use the NC-DS encodingDS achieves in a reduction of our data storage to use for the backing of all monetary transactions within the analyzed system in this way reducing cost compared to the current systems.

Distributed ledger technology for decentralization of manufacturing processes [12]: The use of distributed general ledger (DLT) is studied to innovate industrial cyber physical systems, establish limits of the DLT comfort zone, measurable performance and high impact indicators (KPI), allowing us to apply these innovations to our public systems and thus apply them in the best way to obtain the best possible results according to the needs.

The method used was deductive research to proceed to analyse the information of the articles under study and to obtain the steps for the development of the algorithm.

The result obtained in this phase is a prototype of an algorithm in which flow diagramming techniques were used.

The objective is to analyze the application of Ledger technology in a public organization, to determine the positive impacts it has on the improvement of its operation.

2. Materials and methods

Within the Ledger technology we find the blockchain (block chain), which is a digital journal, in which are recorded all the transactions that have been made grouped in blocks that are linked linearly between them, that is, the first block is linked to the second block, the second block to the third block and thus successively. Ledger technology ensures that every transaction made is valid and unalterable. In this way we will see each block as if it were a page of a virtually infinite accounting book, but in this case the transaction can neither be erased nor repeated because each transaction has a unique fingerprint, with this we guarantee that all transactions are immutable.

It uses a distributed system that will integrate the nodes and the main system which includes all the information necessary for the recording of transactions and equally security rules and encryptions in this way will be controlled by the public organization to ensure that transactions are more secure and robust [4].

The blockchain can be private and public, this is responsible for defining how is the management of the distributed network, if in the case it is a public blockchain, all the participants within the distributed network have the same level of authority and level of access to the information allowing all the nodes to have the same hierarchical level within the network and at the same time all these nodes have full access to the information that exists within the network, this is a disadvantage for the reason that the nodes considered citizens are not can have information from other nodes as it would leak information within the network, instead it is a private blockchain, the main node that is the public organization has higher authority than the other participants or nodes and this is responsible for establishing rules and conditions to participate in the network and in addition to controlling and verifying the information already processed by all the nodes, a final information and already processed, since only this main node has access to the final information, it guaranteed more security and anonymity to the system [4].

For example, rules and conditions were established for entering or performing transactions within the network, the citizen who enters as a node to the system must be over 18 years old to be able to enter the network, the citizen's digital portfolio must have sufficient funds to be able to carry out a transaction, must have active debts to be able to make a payment, must have a PC with internet connection and offer it to be a node within the system in order to be able to enter the network, etc., any node that does not comply with one of the rules must not be accepted or removed from the distributed system and any transaction that does not comply with the conditions must be rejected, both restrictions serve to maintain the safety and efficiency of the network.

The security offered by the blockchain along with a CoC (supply chain on blockchain) encryption, which was in charge of encrypting the transactions within the distributed system, has a greater degree of security than the blockchain offers us, which worked as follows: information is encrypted by encryption algorithms offered by the blockchain technology along with CoC encryption and an additional measure of security is that the transaction is replicated to each node within the network and each node verified this encryption, if all the nodes accepted the validity of the encryption. the encryption is proceeded to complete the transactions that were sent to each node, without this acceptance the transaction is not made, this additional security measure increased the security when making the transactions, with these benefits offered by the Ledger technology, it became a very safe and robust system, for which it is useful and very considerable when taking into account for the implementation [5].

When applying CoC encryption to our network, we must take into account the advantages and disadvantages generated by the use of this type of encryption, the advantages that we consider greater, the encryption is greater, offering more security to each transaction, the latency of the network will decrease considerably unlike the basic use of the blockchain, a storage scheme that relieves the data overload that the blockchain will have, each participating node keeps a copy of the chain which allows having a less load within the network when validating the encryptions , the participating entities do not need to trust each other due to the behavior of the encryption, allows only registered users to the network to write in the transaction blocks, allows to establish entry and operation protocols for the participants in this way, allows adjusting the network in a better way to the operating conditions.

Now we must consider the disadvantages, in this way we will know the disadvantages and improvements for our network, being a centralized network and the CoC works better with decentralized networks should be implemented with some adjustments to maintain the expected functionality, since a copy of the chain in all

the nodes and being a centralized network this causes a double expense and the network suffers delays at the time of encryption, it can be difficult to achieve a consensus on the nodes committed to our network at the moment of building the chains; it needs that the set of nodes within the network is complete, all the participants must maintain a history of transactions which in the long run generates histories of very large sizes that can affect the functioning of the participants.

Figure 1 contains a prototype general outline of a distributed system for a public organization and its members (nodes), which allows us to show the structure of the distributed system that allows us to obtain an optimization of the resources of the network and thus obtain an appropriate performance by applying Ledger technologies [6].

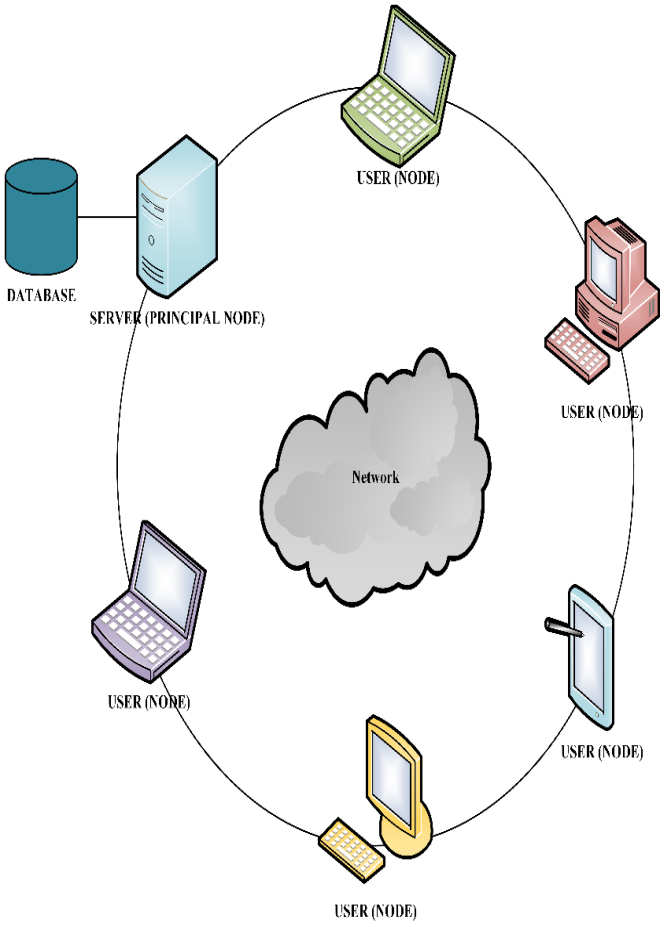


Figure 1: General outline prototype of the distributed system.

A private blockchain was established for this case, where the highest authority within the network will be the public organization, which will determine the rules and conditions of use for all its nodes (citizens) and also verifies the hash of each transaction, is the only one in charge of verifying this information, it does it to each of the transactions, in this way it can detect intruders with false validations [7], instead the nodes can only verify the encryption hash of the previous node, it is not necessary to verify the encryption of other nodes, thus saving node resources.

Such technologies need to maintain the highest level of security both inside and outside the network, in order to ensure a moral integrity that all transactions made have not been modified by administrators of the main node belonging to the public organization, Private Ledgers can have many owners. When a successful transaction is completed and the journal is registered, the ledger is checked to maintain its integrity through a consensus process. This is carried out by a different public organization, which will perform a role of trusted agent, this process helps to ensure the integrity of all transactions that are performed, thus avoiding alterations made within the public organization for malicious purposes or for common errors committed by main node administrators.

In Figure 2 we show how a block is treated by a node, is encrypted, each block has a previous hash of the last block which contains the last hash, a time stamp that tells us how long it takes to find the correct encryption for each block, if this time is exceeded the transaction is cancelled, a nonce that is a random number that allows us to verify that the old hash cannot be used again by for example repeated attacks carried out by intruders and finally the way in which the hash of each block of a transaction.

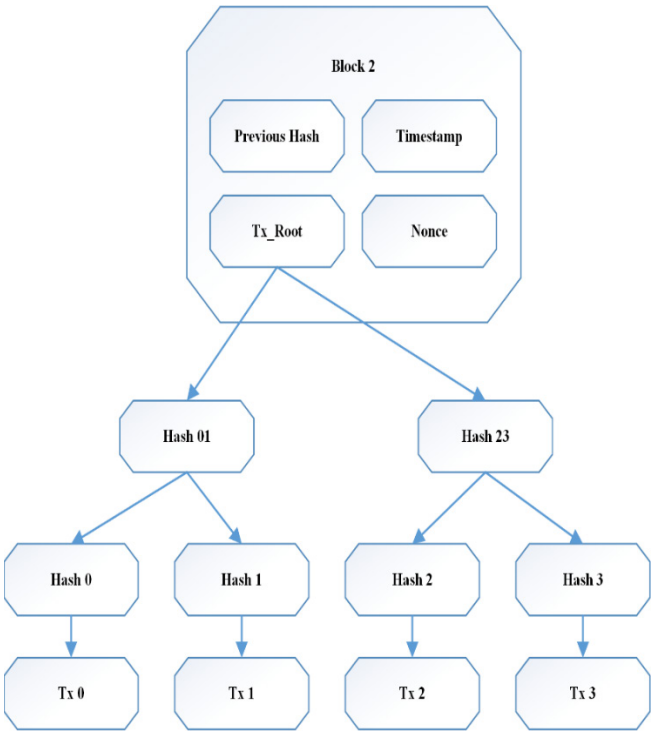


Figure 2: Encryption of each hash transaction in each block

As shown in figure 2 we will use a type of algorithm to generate the hash, the one that offers greater advantage for the blockchain is SHA-256 due to its balance between security and complexity at the moment of generating the keys, for the moment of creating the SHA-256 type key, it will be possible to identify if the transaction has been modified in the trip to its destination, for which the transaction would be rejected when detecting that the hash has changed. By the time the hash has been generated with its algorithm, it will be distributed through the network using the

dispersion method, which is commissioned by a search function to find the position of the hash needed, due to this implementation access to keys is almost direct and in a short time, because you would only need one or two memory attempts to achieve your goal by avoiding collisions within the network when distributing hash keys, the search time for each key is independent of how many keys exist within the network.

Applying this set of cryptography to the network we achieve security and robustness, but we must consider the performance of the network due to these applications, we can reduce collisions by not applying this set of cryptographic functions. When the described cryptography is not applied to the network, we will find two main problems that will affect the network seriously, generating two keys pointing to the same index will not be able to store the information creating collision problems that will cause the network to have delays when sending information from node to node, another problem that we found is the agglomeration that occurs at the time the structure causes that commonly used keys tend to fall very close to each other and can fall consecutively, this degrades the network because it accumulated the keys making the cost of processing and time to solve it significantly large in the search. We solve this by applying detailed methods, due to the method of dispersion the search accelerates the desired hash thus avoiding the collision with little use of memory and process achieving a network more robust in security and more efficient when making secure transactions and the crowding will not affect the network due to the search in a few attempts and times.

To develop the system within a distributed scope, it was necessary that the money that is used within the digital portfolios within the virtual net (bitcoin), in this way was guaranteed security and anonymity for each citizen, since no other node will be able to know where the payment comes from or where the payment goes only the main node being the control within the network, it knew at the end of the transaction, who made the payment and to which debt it was addressed.

In order to demonstrate the functioning of the network and all its participants and the functions they performed, the order of operation of the network shall be followed:

The network must have a minimum of 5 active nodes, that is to say that it is not carrying out any activity and that it is waiting for a request, it must be stated with this minimum amount due to the required transactions, that all transactions sent by the nodes will always be divided into 5 tasks that will be solved by the active nodes in this way an order is maintained within the network and an overload of the network or the disutilization of resources is avoided, if there are more than 5 active nodes and they are waiting for requests, these nodes will work as a backup in the event that a node may fail or may no longer be available at the time of the transaction, in this way we ensure that the network is always complete in order to be able to resolve the request. And if there are no nodes available, the transaction will be put on hold until we have the minimum acceptable to resolve the transaction.

2.1. The transaction was generated

A virtual journal was responsible for recording all transactions that have been made within the network, this method was used to record all monetary transactions that were made within the network. The transaction was generated by a node (citizen) and is encrypted under the rules and CoC encryption, which was sent to the administrator node (server of the public entity) of the distributed network, which was in charge of verifying all the rules and conditions are fulfilled.

The transaction was sent to other node that worked as blocks within the chain to solve the complete transaction, different small jobs are assigned to the participating nodes, to be able to resolve the transaction as: resolve that the transaction was made with satisfaction, another node took care that the citizen's monetary balance was reduced from the value paid and another node of the public entity's monetary balance was increased from the value paid by the citizen, another node gathered the information that will be shown once it was confirmed or the transaction was denied, another node was responsible for verifying the encryption of the virtual currencies (bitcoin) is correct, due to this the processing load of the whole transaction was divided into the different nodes belonging to the network, which expedited the workloads and helped the centralized systems of public entities have less transaction burdens, so only the public server was in charge of registering the transactions once they were completed and informing the citizen through their information systems (web page) [8].

2.2. The movements of the network were verified

A user of the blockchain network made a transaction, the transaction was divided into several tasks: verifying the legality of the transaction, decreasing the monetary balance, increasing the monetary balance, recording in the digital journal, these tasks were performed by a different node within of the network, through Ledger technology, the tasks were performed anonymously and distributed throughout the network, in this way each transaction is isolated from the knowledge that the owner of the node [9] can have, without knowing the information that contains the transaction, it was possible to include anonymity to the transactions and the user (citizen), in comparison with a transaction made with a private institution, which earns a commission for carrying out the transaction and which has access to private information generated by the public institution, due to the transaction will run the risk of a leak of information that exists on behalf of the public institution.

The movements that exist in the network are verified both node to node that information is sent by encryption that contains the hash of the past node for a verification that the sent information has not been manipulated or changed, just as the main node is in charge of verifying that the entire block transaction has been successfully completed and at the same time controlling the encryption of each of the blocks, in this way the transaction has not been modified by any intruder external to the network, since its hash will not be recorded within the network, it is assumed that the transaction has been manipulated [10].

2.3. Effectiveness of the transaction

When all the transactions have been resolved by the nodes of the network, and the encryption keys of all the nodes have been verified by the main node, it can be assumed that the transaction has been completed and the main node which is the public institution was in charge of registering the complete transaction, in

a centralized database in which no node has access to be able to protect the information of all the participating nodes since there will only be private connection between the main node and the database, with the help of an NC-DRDS framework is responsible for coding all the incoming information of the transactions resolved by the nodes, this decreases the load both in the network and in the database, by sending the compressed data in small blocks, in which the framework is responsible for compressing them in their minimum allowed equivalence and at the same time when it is necessary to decompress it for the use of information [11].

The transaction information once completed, and stored in the database, the web server used this information to replicate it on the public institution's website for the user's use, as for example to be able to see reflected the debt but already once the payment transaction was carried out, in your digital portfolio you can see the decrease in your monetary balance by performing the payment transaction or proof of payment of the debt by means of a certificate issued by the public organization in order to be able to support the physical payment [12].

3. Results

The research demonstrated a model that allowed the public organization to implement Ledger technology more effectively, because of the benefits of Ledger technology can offer to an infrastructure of a public organization.

Figure 3 will show the algorithm that served as a guide for the total resolution process of a user's transaction with a public institution:

Start: The citizen required to make the request for a transaction to make the payment to a public organization, becoming a node of the network.

Data entry: All data necessary to perform the transaction was requested, data will belong to the citizen and public organization, such as digital portfolio, node data, debt list, which are treated only by the main node to ensure data integrity.

Encrypting the transaction: The participating node as the public organization is in charge of encrypting the information they are going to send to the network with CoC encryption, both nodes when encrypting the information ensure that the message cannot be manipulated either by other nodes of the same network or external agents of the same, guaranteeing security to the information of both parties.

Verification of the encryption: The public organization that is the main node receives the information necessary to carry out the transaction and be sent to the network and is in charge of verifying that the encryption of the information that has passed through the network is correct and not has been manipulated and changed, this way we add a more level of security.

Validation of rules and encryption: The main node that is the public organization was in charge of validating that the node requesting the transaction complies with the rules that the server has established, which is greater than 18 years, who has money in his account, who has outstanding debt and to verify the encryption of the node, otherwise if he does not comply with the rule or the encryption is not correct the request is rejected waiting for another request from another node.

Transaction accepted and divided: The transaction was accepted fulfilling all the requirements of the server, a node takes care that the transaction is divided into small processes that were resolved by the other nodes within the network.

Transactions sent to nodes: When the transaction is divided, it is sent to each node of the network, which each will have a different encryption, where each node verifies its own encryption and the transmitting node, avoiding manipulation of information in the course of the network.

Decrease balance: One node took care to resolve this process of diminishing the balance of the digital portfolio of the node that the payment transaction requires.

Increase balance sheet: One node was responsible for resolving this process of increasing the balance of the digital portfolio of the public organization by collecting the debt of the transaction.

Collect information: A node was in charge of gathering the information of the citizen and the debt to be able to show this information at the time of completion.

Verify encryption: A node was in charge of verifying the encryption that the transaction needs once completed, in order to send it to the main node.

Distribute information: A node is in charge of gathering the necessary and public information for the server to publish in its media (web page).

Validation of transactions and their encryptions: validates that all processes ordered to each node have been resolved and that the encryption of all nodes are correct otherwise the transaction is cancelled completely.

Sending to main node: Each node sent to the server that does the main node function, all information already processed.

Gathering information: The main node is responsible for attaching all the information I receive and use to give you.

Database: All information collected by the server was recorded and sent to be stored in the database of the public organization.

Display public information: We proceeded to display information that may be public for the node that requested the transaction by means of media such as web pages of the public organization.

End: Completes the transaction and is expected by another node that requires the resolution of another.

Now, in Figure 3 we describe the algorithm that helped us solve the established needs.

Figure 3 shows us by means of a flow diagram how the system works, but it is not exempt from problems or faults, which are solved by establishing operating conditions within the network, One of the most important problems is when a node has its designated transaction and is no longer available for the network,

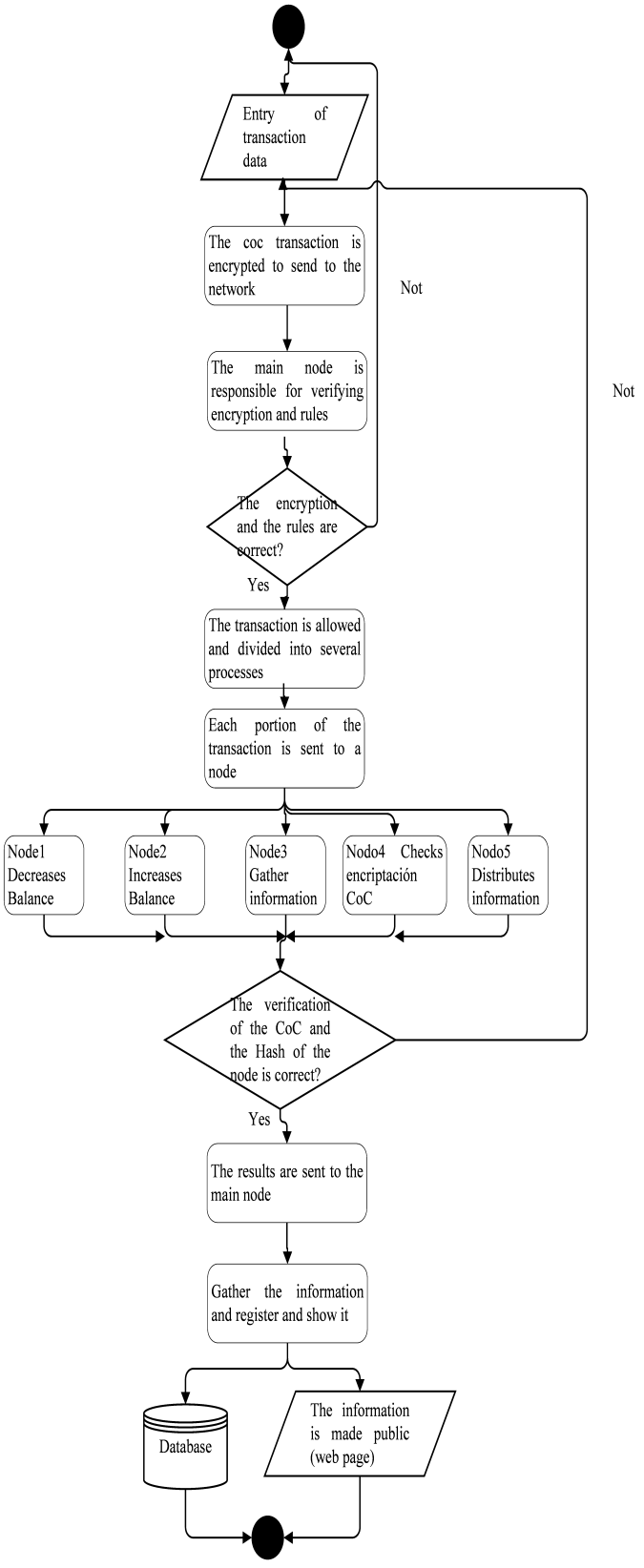


Figure 3: Prototype of the system operating algorithm.

for this a node that is available within the network and is waiting for a transaction, will replace the node that is no longer available, in this way we ensure that the network is always complete to resolve a transaction, although due to this substitution the network will suffer a small delay in operation due to the replacement of nodes, but in this way we ensure its operation and compliance.

In Figure 4 we show a simplified form of the main steps of resolving a distributed transaction.

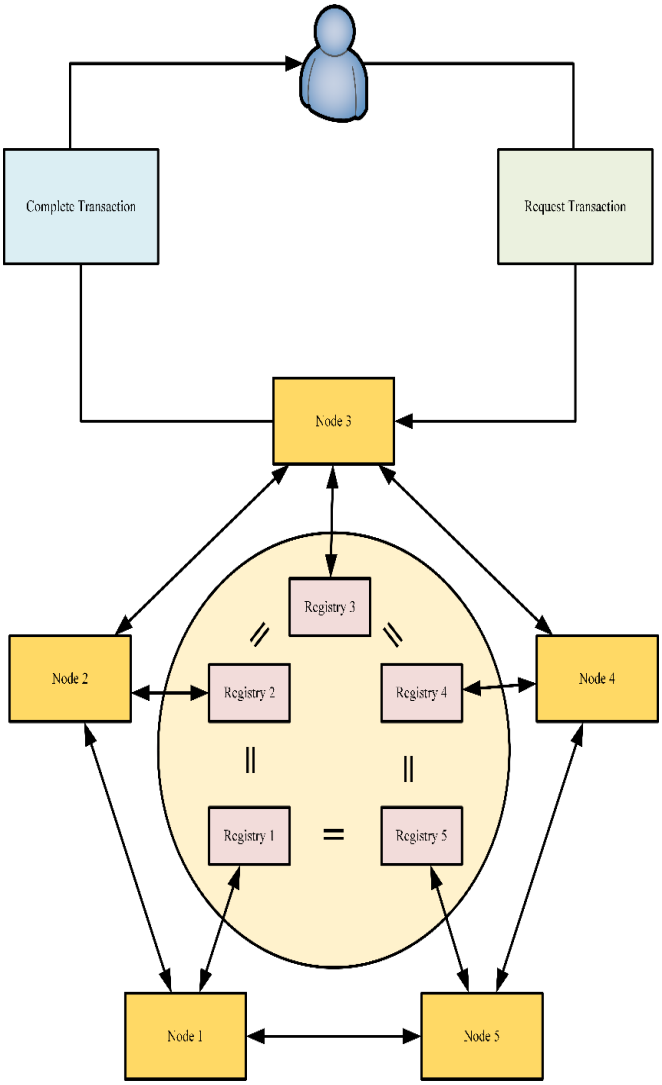


Figure 4: Resolution scheme within the distributed

A user requires the resolution of a payment transaction, the node belonging to the user sends this request to a main node, requesting its processing. Once the application is received, it is sent to the network and other nodes verify the identity of the applicant and certify the transaction in a way as the consensus mechanism declares, confirming that the applicant has the necessary credentials to perform the transaction. Each node oversees resolving each part of the transaction, sending the information resolved by each node to the main server of the public organization, which handles it and shows it to the citizen who made the request.

This scheme ensures a distributed system according to the Ledger technologies, in this way guaranteeing security, efficiency to the whole network and its operation, thus achieving objectives necessary for the success of the implementation.

For the encryption of the messages that are sent in each node, basic encryption offered by the Ledger technology is used, which we will proceed to explain in equation 1 along with its mathematical sustenance:

$$A = g^a \bmod p \quad (1)$$

A random prime number is established which will be p, and a generator g of prime values smaller than p, where the node chooses a number at random smaller than p and is recorded in a, with this proceeds to the calculation of A and sent to the node with which it communicates.

$$B = g^b \bmod p \quad (2)$$

In equation 2 the node does the same work as the node with which it communicates, establishes a prime number, uses a G generator and is calculated in the same way as B and is sent to the node with which it communicates.

$$K = A^b \bmod p = B^a \bmod p \quad (3)$$

In equation 3, it states that the value of K must be equal in both nodes to know that the encryption of both nodes is correct, for which both nodes use the result sent by the other one, This way each node can get its value of K.

At the time a request is sent, the issuer is in charge of searching for the public key that has the receiver's encryption, at this time it encrypts its request with the receiver's key, and when the message reaches the receiver, is in charge of deciphering it using its own hidden key.

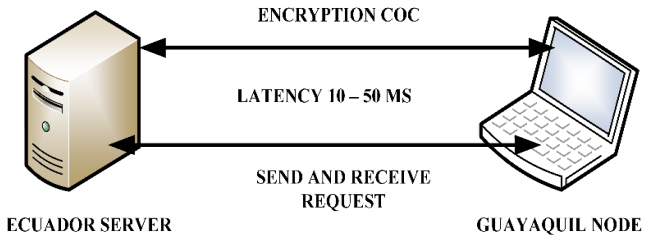


Figure 5: Operation of a node and impact on the network.

As we can see in Figure 5 the performance of a node with the server affecting the network, the main server sends the first message with CoC encryption, which will have a latency of 20 to 30 milliseconds, in order to meet these objectives it must be have a minimum standard for the network, the main one that the network has a bandwidth of 1 megabyte and that the node and server comply with distance limits, in this case by the guidelines generated by the main node, they are within the same country that governs the Public Organization. Once the server and the node have sent and received the encryption, the server sends the request and the node resolves it, for this transaction the latency is affected by the increased information traffic, handling the minimum standard of 1 megabyte of bandwidth, latency can vary between 20 to 50 milliseconds, the difference between the two interactions that have the server and the node, for encryption when using the search method ensures that the network sends less information through the network affecting the latency in the network, decreasing it, on the other hand for the interaction of resolving the request, the latency is increased in the network, because the complete transaction is processed, affecting the latency.

4. Discussion

According to the results of this research of our model we will obtain that the implementation of the blockchain together with a distributed network and centralized systems manage to guarantee benefits such as safety, speed, efficiency, effectiveness, reduction of costs and anonymity between the citizen and his environment except for the public organization that needs such knowledge for the operation of the collection model.

Our algorithm along with the encryption that the network has proposes a more efficient way than the current one in order to carry out the transaction following the standards of the blockchain achieving the proposed objectives.

5. Future Work or Conclusion

The blockchain model has benefits that provide viable features to this model applied to the environment under study, but there are improvements in information encryption and block chain management, we have chosen the methods proposed by the obtained results, but there are improvements with different methods of encryption and system design.

The algorithm generated along with the developed scheme, offer security advantages and are efficient, but for better security and avoid failures within the running model, the public entity (higher authority within the system) must apply very strict rules and maintain control, which will set a higher level of security that will help make the model more robust and safer.

Acknowledgment

The authors thank the Salesian Polytechnic University of Ecuador, the research group of the Guayaquil Headquarters "Information Technology, Security and Information for a Globalized World" (CSITGW) created in accordance with resolution 142-06-2017-07-19 and the Secretariat of Higher Education Science, Technology and Innovation (Senescyt).

References

[1] M. Kohli and E. Suarez, "Centralized solution to securely transfer payment information electronically to banks from multiple enterprise resource planning (ERP) systems," Proc. - 2016 15th Int. Conf. Inf. Technol. ICIT 2016, pp. 275–282, 2017.

[2] M. Wang, Q. Wu, B. Qin, Q. Wang, J. Liu, and Z. Guan, "Lightweight and Manageable Digital Evidence Preservation System on Bitcoin," J. Comput. Sci. Technol., vol. 33, no. 3, pp. 568–586, 2018.

[3] K. Zhang and H. A. Jacobsen, "Towards dependable, scalable, and pervasive distributed ledgers with blockchains," Proc. - Int. Conf. Distrib. Comput. Syst., vol. 2018–July, pp. 1337–1346, 2018.

[4] L. D. Ibáñez, E. Simperl, F. Gandon, and H. Story, "Redecentralizing the web with distributed ledgers," IEEE Intell. Syst., vol. 32, no. 1, pp. 92–95, 2017.

[5] L. Xu, L. Chen, Z. Gao, Y. Lu, and W. Shi, "CoC: Secure Supply Chain Management System Based on Public Ledger," 2017 26th Int. Conf. Comput. Commun. Networks, pp. 1–6, 2017.

[6] R. El Bansarkhani, M. Geihs, and J. Buchmann, "PQChain: Strategic design decisions for distributed ledger technologies against future threats," IEEE Secur. Priv., vol. 16, no. 4, pp. 57–65, 2018.

[7] D. Burkhardt, M. Werling, and H. Lasi, "Distributed Ledger," 2018 IEEE Int. Conf. Eng. Technol. Innov., pp. 1–9, 2018.

[8] M. R. Ogiela and M. Majcher, "Security of Distributed Ledger Solutions Based on Blockchain Technologies," no. c, pp. 1089–1095, 2018.

[9] F. M. Benčić and I. P. Žarko, "Distributed Ledger Technology: Blockchain Compared to Directed Acyclic Graph," Proc. - Int. Conf. Distrib. Comput. Syst., vol. 2018–July, pp. 1569–1570, 2018.

[10] H. R. Hasan and K. Salah, "Blockchain-based Proof of Delivery of Physical Assets with Single and Multiple Transporters," IEEE Access, vol. PP, no. 8, pp. 1–1, 2018.
[11] M. Dai, S. Zhang, H. Wang, and S. Jin, "A Low Storage Room Requirement Framework for Distributed Ledger in Blockchain," IEEE Access, vol. 6, pp. 22970–22975, 2018.
[12] M. Isaja and J. Soldatos, "Distributed ledger technology for decentralization of manufacturing processes," Proc. - 2018 IEEE Ind. Cyber-Physical Syst. ICPS 2018, pp. 696–701, 2018.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 260-277 (2019)

www.astesj.com

ASTESJ

ISSN: 2415-6698

A Support Vector Machine Cost Function in Simulated Annealing for Network Intrusion Detection

Md Nasimuzzaman Chowdhury, Ken Ferens*

Department of Electrical & Computer Engineering, University of Manitoba, Winnipeg, MB, R3T5V6, Canada

ARTICLE INFO

Article history:
Received:02 May, 2019
Accepted:22 May, 2019
Online: 30 June, 2019

Keywords:
feature extraction
simulated annealing
support vector machine

ABSTRACT

This paper proposes a computationally intelligent algorithm for extracting relevant features from a training set. An optimal subset of features is extracted from training examples of network intrusion datasets. The Support Vector Machine (SVM) algorithm is used as the cost function within the thermal equilibrium loop of the Simulated Annealing (SA) algorithm. The proposed fusion algorithm uses a combinatorial optimization algorithm (SA) to determine an optimal feature subset for a classifier (SVM) for the classification of normal and abnormal packets (possible intrusion threats) in a network. The proposed methodology is analyzed and validated using two different network intrusion datasets and the performance measures used are; detection accuracy, false positive and false negative rate, Receiver Operation Characteristics (ROC) curve, area under curve value and F1-score. A comparative analysis through empirically determined measures show that the proposed SA-SVM based model outperforms the general SVM and decision tree-based detection schemes based on performance measures such as detection accuracy, false positive and false negative rates, area under curve value and F1-score.

1. Introduction

Big data refers to an extremely large volume of information, whose analysis cannot be done in real-time using standard techniques. Analyzing big data includes, but is not limited to, extracting useful information for a particular application and determining possible correlations among various samples of data. Major challenges of big data are enormous sample sizes, high dimensionality problems and scalability limitations of technologies to process the growing amount of data [1]. Knowing these challenges, researchers are seeking various methods to analyze Big Data through several approaches like different machine learning and computationally intelligent algorithms.

Machine learning (ML) algorithms and computationally intelligent (CI) approaches plays a significant role to analyze big data. Machine learning has the ability to learn from the big data and perform statistical analysis to provide data-driven insights, discover hidden patterns, make decisions and predictions [2]. On the other hand, the computational intelligence approach enables the analytic agent/machine to computationally process and evaluate the big data in an intelligent way [3] so, big data can be utilized efficiently. Particularly, one of the most crucial challenges of analyzing big data using computational intelligence is searching through a vast volume of data, which is not only heterogeneous in structure but also carries complex inter-data relationships. Machine learning and computational intelligence approaches help in big data analysis by providing a meaningful solution for cost reduction, forecasting business trends and helps in feasible decision making considering reasonable time and resources.

One of the major challenges of machine learning and computational intelligence is an effective feature extraction approach, which is a difficult combinatorial optimization problem [4]. A feature is a measurable property, which helps to determine a particular object. The classification accuracy of a machine learning method is influenced by the quality of the features extracted for learning from the dataset. Correlation between features [5] carries great influence on the classification accuracy and other performance measures. In a large dataset, there may be a large number of features which do not have any effect or may carry a high level of interdependence that may require advanced information theoretic models for meaningful analysis. Selecting proper and reasonable features from big data for a particular application domain (such as cyber security, health and marketing) is a difficult challenge and if done correctly, could play a significant role in reducing the complexity of data.

* Ken Ferens, Rm E1-544, Department of Electrical & Computer Engineering, University of Manitoba, 1 204 474 8517, Ken.Ferens@umanitoba.ca

In the domain of combinatorial optimization, selecting a good feature set is at the core of machine learning challenges. Searching is one of the fundamental concepts [6] and is directly related to the famous computation complexity problems such as Big-O notations and cyclomatic complexity. Primarily, any problem that is considered a searching problem looks for finding the "right solution," which is translated in the domain of machine learning as finding a better local optimum in the search space of the problem. Exhaustive search [7] is one of the methods for finding an optimal subset of the solution, however, performing an exhaustive search is impractical in real life and will take a huge amount of time and computational resources for finding an optimal subset of the feature set to provide a solution.

In a combinatorial optimization problem, there is a finite or limited number of solutions available in the solution space. Most of the combinatorial optimization problems are considered as a complicated problem [8]. Simulated Annealing (SA) is one of the computational intelligence approaches for providing meaningful and reasonable solutions for combinatorial optimization problems [9] [10] and can be utilized for feature extraction (example; for cybersecurity threat detection). As per our literature survey, it is found that simulated annealing is usually not utilized as a classifier [11]. However, the SA method is explored a lot for searching optimal solutions to problems such as the travelling salesperson problem [12], color mapping problem [8], traffic routing management problem [13], and clustering of large sets of time series [14].

State of the art research in merging ML and CI algorithms has demonstrated promise for different applications such as electricity load forecasting [15], pattern classification [16], stereovision matching [17] and most recently for feature selection [18]. In a practical application, it is required to find a reasonably better feature set that can be utilized for cyber intrusion detection with relatively better reliability and performance. This paper addresses this challenge empirically using various datasets and proposes a methodological approach.

In this paper, we have introduced an intelligent computational approach merging Simulated Annealing (SA) and Support Vector Machine (SVM) with an aim to provide a reasonable solution for extracting optimum (minimum) features from a finite number of features. The classifier is designed with the goal of maximizing the detection performance measures, and the combinatorial optimizer is designed to determine an optimal feature subset, which is input to the classifier. We have applied this general methodology on two different Network intrusion datasets; UNSW dataset (Australian Centre for Cyber Security) [19] [20] and UNB dataset (Canadian Institute of Cyber Security) [21] in order to analyze the performance of the proposed method and evaluate whether the outcome can provide an optimum feature subset and can detect the presence of intrusion in the network system. Furthermore, the empirically validated outcomes of the proposed method are evaluated in contrast with other machine learning methods like general SVM (without annealing) and decision tree to analyze which methodology provides a better reasonable solution.

2. Background Research

Various research works have been conducted to find an effective and efficient solution for combinatorial optimization problems (optimum feature subset selection) for network intrusion detection to ensure network security and for various other applications. An ideal intrusion detection system should provide good detection accuracy and precision, low false positive and negative, and better F1-score. However, nowadays for the increasing number of intrusions, software vulnerabilities raise several concerns to the security of the network system. Intrusions are easy to launch in a computing system, but it is challenging to distinguish them from the usual network behavior. A classifier (that classifies normal and anomalous behavior) is designed with the goal of maximizing the detection accuracy and the feature subset utilized by the classifier is selected as the optimal feature subset. Researchers have been trying to develop different solutions for different types of scenarios. Finding an optimum feature subset for reliable detection system is a significant combinatorial optimization problem in network intrusion detection. Some of the related works are described below based on the approaches in different sectors (cybersecurity, electricity bill forecasting, tuning SVM kernel parameters) and advantages and disadvantages.

In [22], the authors proposed a combined SVM based feature selection model for combinatorial optimization problem in which they applied convex function programming additionally to the general SVM programming to find an optimal subset of features. This approach consumes more computational resources, and the process is mathematically complex.

In [23-25], the authors provided a signature-based detection method which is capable of detecting DoS and routing attack over the network. In [24] the authors mentioned a signature-based model such that the total network system is divided into different regions, and to build a backbone of the monitoring nodes per region they established a hybrid placement philosophy [26]. However, this method was limited to the known signature models. If the signature is not updated and unknown to the nodes at the different region, it does not find a match, and the intrusion walks inside the system. In this proposed system, there were no approaches to finding an optimal feature set to determine any unknown type attacks.

In [27], the authors also proposed a signature-based model in which each of the nodes will verify packet payload and the algorithm will also skip a large number of unnecessary matching operation resulting low computational costing and comparison differentiate between standard payloads and attacks [26]. This is a fast process of identifying malicious activity but when the complexity of the signatures increases it may be unable to detect the malicious packet.

In [28], the authors proposed an OSVM (Optimized Support Vector Machine) based detection approach in which the outliers are rejected and make it easier for the algorithm to classify attacks with precision. For more massive datasets which have some feature dimension then this algorithm does not perform well as it does not know which features to use as the feature workspace is very high, resulting the algorithm performing an exhaustive search on the whole workspace. The proposed method does not provide any reasonable solution for the optimum feature extraction method.

In [29], the authors proposed a random forest-based intrusion detection mechanism that was applied to both anomaly and

signature-based data samples. The random forest-based approach works fine on the signature-based approach, but the algorithm was unable to detect malicious characteristic with an excellent detection accuracy. Also, when applied on large dataset the complexity of detection was very high for this algorithm to perform.

In [30], the authors proposed decision tree-based wrapper intrusion detection approach in which the algorithm can detect a subset of the feature among all the features available on the KDD dataset. So it reduced the computational complexity of the classifier and provided high accuracy of detection intrusion. However, it also performs an exhaustive search trying all possible feature subsets to provide an output. If the feature numbers are high, also data set is large then doesn't provide good accuracy regarding detection accuracy and consumes much time. In real time scenario, this method may fail to detect an anomaly within the secured time limit.

In [31], the authors introduced a lightweight intrusion detection methodology in which energy consumption is considered as a detection feature to detect abnormal behaviors on the network flow. When the energy consumption diverges from an anticipated value, the proposed method calculates the differences of the values, and the algorithm classifies the anomaly from the normal behavior. They minimized the computational resources by focusing only on the energy consumption, so algorithm works faster and provides an acceptable solution for the intrusion detection. In anomaly-based detection scheme as the characteristic behaviors of the data packets are analyzed, what if the node does not consume more energy it consumes more than the specified time to transfer data over to the network? It may be compromised by modification attack which creates a time delay in the route from source to destination. This algorithm becomes vulnerable if the characteristic of the anomaly is different rather than energy consumption. A single feature is not sufficient enough to detect a particular attack precisely.

In [31], the authors proposed in their research on intrusion detection that network nodes must be capable of detecting even small variations in their neighborhood and the data has to be sent to the centralized system. They proposed three algorithms on the data sent by the node to find such type of anomaly namely wormhole. They claimed that their proposed system is suitable for IoT as it consumes low energy and memory to operate [23]. However, analyzing the data samples using three types of algorithm consumes a massive amount of time and limits the countermeasure effectiveness of the algorithm as its huge taking time to detect attacks in real time scenario. Also when the network facing huge traffic flow the algorithm may not be detecting intrusion in the secured time limit.

In [32], the authors proposed a group-based intrusion detection system, which uses a statistical method and designed a hierarchical architecture-based system. The results were very highlighting as their detection accuracy was very high, low false alarm rate, low transmission power consumption. However, the method does not seem feasible if multiple features are considered and don't provide any information about the process of selecting multiple features. Thus, the combinatorial optimization exists in such a scenario.

In [33], the authors used artificial intelligence artificial neural network scheme where ANN is used to every sensor node. The algorithm provides self-learning ability to the intrusion detection system. ANN is an excellent approach in intrusion detection, but node energy consumption becomes high as its continuously learning from the data packet flow.

In [34], the authors proposed an efficient impostor alert system against sinkhole attacks. In this system, a record of the suspected network nodes is generated by continuously analyzing the data. After that, the data flow information's are used to identify the intrusion in the system. When traffic volume is high, and a lot of data packets are flowing, there may be a scenario that many nodes are in the suspect list and comparing all of them may limit the network performances. The algorithm in this research is performing an exhaustive search for finding an optimal feature subset for sinkhole attack detection.

In [29], the authors proposed a random forest-based intrusion detection mechanism that was applied to both anomaly and signature-based data samples. The random forest-based approach works fine on the signature-based approach, but the algorithm was unable to detect malicious characteristic with an excellent detection accuracy. Also, when applied on large dataset the complexity of detection was very high for this algorithm to perform.

In [30], the authors proposed decision tree-based wrapper intrusion detection approach in which the algorithm can detect a subset of the feature among all the features available on the KDD dataset. So, it reduced the computational complexity of the classifier and provided high accuracy of detection intrusion. However, it also performs an exhaustive search trying all possible feature subsets to provide an output. If the feature numbers are high, also data set is large then doesn't provide good accuracy regarding detection accuracy and consumes much time. In real time scenario, this method may fail to detect an anomaly within the secured time limit.

3. Background of Machine Learning Algorithm

3.1. Support Vector Machine

The Support vector machine (SVM) [35] is a discriminative and normally supervised machine learning methodology that analyzes training samples to process a wide variety of classification problems. The algorithm generates an optimal hyperplane which classifies training examples and new data samples. This supervised learning method can analyze the data samples to perform handwritten character recognition [36], face detection [37], pedestrian detection [38], text categorization.

Consider a training dataset $T = \{(x^i, y^i)\}, i = 1, \dots, p$, where $x^i \in R^n$ represents the training examples with n dimensional input features, p is the number of training examples, and $y^i \in \{-1, +1\}$ represents the desired or labelled output of the i^{th} training data sample. $y^i = 1$ denotes output of the i^{th} positive training samples and $y^i = -1$ denotes the output of the i^{th} negative training samples.

Based on the above consideration, the decision hyperplane is given by (1).

$$\sum_{i=1}^{p} \boldsymbol{w}^T x^{(i)} + b = 0 \quad (1)$$

The $\boldsymbol{w}$ and b represent the weight vector and bias term, respectively. During the training process, the weight and the bias terms are learned. With these learned parameters, the decision hyperplane places itself at an optimum location between the positive and negative training example clouds. SVM places the decision boundary in such a way that it maximizes the geometric margin of all the training data samples. In other words, all training examples have the greatest possible geometric distance from the decision boundary. The optimization problem is given by (2).

$$\min_{w,b} \frac{1}{2} \| W \|^2$$

$$\text{s.t} \quad \sum_{i=1}^{p} y^{(i)}(W^T x^{(i)} + b) - 1 \geq 0, \quad (2)$$

$$i = 1 \ldots p$$

The Lagrange multiplier is used to solve this constrained optimization problem. The Lagrangian for this problem is given by (3).

$$L(w, b, \alpha) = \frac{1}{2} \| W \|^2 - \sum_{i=1}^{p} \alpha(y^{(i)}(W^T x^{(i)} + b) - 1) \quad (3)$$

The Lagrangian multiplier is represented by $\alpha \geq 0$. The dual form of the Lagrangian may be written as (4).

$$\max_{\alpha} L(\alpha) = \max_{\alpha} \left\{ \sum_{i=1}^{p} \alpha_i - \frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} \alpha_i \alpha_j y^{(i)} y^{(j)} x^{(i)} x^{(j)}) \right\}$$

$$\text{s.t} \quad \sum_{i=1}^{p} \alpha_i y^{(i)} = 0,$$

$$\alpha_i \geq 0, i = 1, \ldots, p \quad (4)$$

The solution above drives the optimum decision surface that can distinguish linearly separable positive and negative training data example clouds. However, for non-linearly separable training data, a suitable kernel and regularization may be applied. The Gaussian kernel is widely used in such types of problems. Applying a kernel and regularization to (4) gives (5) (C is the regularization parameter).

$$\max_{\alpha} W(\alpha) = \sum_{i=1}^{p} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{p} \alpha_i \alpha_j y^{(i)} y^{(j)} K(x^{(i)}, x^{(j)})$$

$$\text{s.t} \quad \sum_{i=1}^{p} \alpha_i y^{(i)} = 0,$$

$$0 \leq \alpha_i \leq C, i = 1, \ldots, p \quad (5)$$

3.2. Simulated Annealing

Simulated Annealing can be described as an iterative procedure that is composed of two loops. The outer loop is known as a cooling loop and the inner loop as a thermal equilibrium loop. The algorithm is initialized by several parameters like the number of cooling loops, number of equilibrium loops, and probability function. The purpose of the inner loop is to find the best solution for the given temperature to attain thermal equilibrium at the given temperature state. In each equilibrium loop, the algorithm takes a small random perturbation to create a new candidate solution. Initially, as the algorithm does not know which direction to search, it picks a random direction to search, and an initial solution is created. A cost function determines the goodness of the solution. A small random perturbation is made to the current solution because it is assumed that good solutions are generally close to each other, but it is not guaranteed as the best optimal solution. Sometimes the newly generated solution results in a better solution, then the algorithm keeps the new candidate solution. If the newly generated solution is worse than the current solution, then the algorithm decides whether to keep or discard the worse solution, which depends on the evaluation of Boltzmann's probability function, which is given by (6).

$$P = e^{-(\frac{\Delta E}{k_b T})} \quad (6)$$

The change in energy ΔE can be estimated by the change in the cost function ΔC, corresponding to the difference between the previously found best solution at its temperature state and the cost of new candidate solution at the current state. Boltzmann's constant may be estimated by using the average change in cost function (ΔC). Thus, Boltzmann's function may be estimated using (7).

$$P = e^{-(\frac{\Delta C}{\Delta C_{ave} T})} \quad (7)$$

After running the inner loop for the prescribed number of times, wherein each loop it takes a new better solution or keeps a worse solution, the algorithm may be viewed as taking a random walk in the solution space to find a sub-optimal solution for the given temperature.

The current best solution will be recorded as the optimal solution. The temperature is decreased according to a basic linear decrease schedule. The initial temperature is set to a very high value initially, because it allows the algorithm to explore a wide range of solutions, initially. The final temperature should be set to a low value that prevents the algorithm to accept a worse solution at the last stages of the process. If the number of the outer loops

has not reached zero, then the inner loop is called again otherwise the algorithm terminates.

3.3. Decision Trees

Decision tree [39] is a machine learning mechanism which is mostly used for classification and regression in many application domains. Decision trees are based on conceptual tree analytical model that considers dependency perception of an object in such a way that the branches of the tree represent the dependency, and the leaf of the tree represents the object itself regarding the classification labels (such as logical 0 or logical 1) [40]. Further, research literature shows that decision trees are used to represent the extraction of dependent features from a data set where the branches represent the feature or attribute while the leaf represents the decision using class labels. Decision trees are mostly used in data mining and machine-learning research works [41].

There are several decision tree algorithms such as ID3 [42], C4.5 (improved from ID3) [43] and CART (Classification and Regression Tree) [44]. CART based decision tree algorithm is used mainly for machine classification purposes [45].

CART [46] can be used for classification of categorical data or regression of continuous data. CART algorithm is designed to develop trees based on the sequence of rules. If the object passes a specific rule, it goes into one structure otherwise it is sent to other structure. Further, the rules or questions defines the next step to follow. For example, there are two random variables X_1 and X_2. Let's say there are decision thresholds or rules are t_1 and t_2. If $X_1 < t_1$, go and check if $X_2 < t_2$ otherwise, go and check if $X_1 < t_3$ and so on.

In the CART algorithm, the splitting process (or decision-making process at each step) is the most significant step of the training phase for machine learning. There are several criterions for the task. For example, Gini criterion (for CART) and Information entropy criterion (for C4.5) are widely used. Gini; a statistical measure which can be calculated by summing the random variable's probability q_i (where i is the index for a random variable).

In order to calculate the Gini index for a set of features/attributes with K classes [47], let's assume that $i \in \{1,2,3 \dots \dots \dots K\}$, and let q_i be the fraction of the items labelled with class i in the set. Accordingly, the set of equation is given by (8).

$$I_G(q) = \sum_{i=1}^{K} q_i \sum_{k \neq 1} q_k$$

$$I_G(q) = \sum_{i=1}^{K} q_i(1 - q_i)$$

$$I_G(q) = 1 - \sum_{i=1}^{K} q_i^2 \qquad (8)$$

Therefore, it can be seen that the Gini index $I_G(q)$ for a particular labelled item is a function of the sum of all probabilities in the tree. Research literature and various researchers discussion on blogs [48] [49] indicate that CART and C4.5 algorithms provide robust classification in application domains such as health care, marketing, financial forecasting and cyber security systems. The main advantage of the CART algorithm is that it does not have logarithm calculation in Gini index that makes the algorithm faster and efficient than the C4.5 algorithm.

3.4. Dataset Preprocessing & Attack Types

The first dataset used in this research was obtained from the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) [19]. In this dataset, a hybrid of real modern normal activities and attack behaviors were generated. This dataset contains a total of forty-seven features and contains around 2.5 million sample data [19] [20]. It consists of such type of attacks like Fuzzers, Analysis, Backdoors, DDoS, Exploits, Generic, Reconnaissance, Shellcode, Worms & normal data samples with labels.

In the UNSW dataset, 47 columns represent attributes/features. Each recorded sample consists of attributes of different data forms like binary, float, integer and nominal. The attack data samples are labelled as '1,' and normal data samples are labelled as '0'. Some of the feature data sample values are categorical values. For example, source IP address, destination IP address and source port number. Also, some other feature data sample values are continuous variable. For example, source bits per second, source jitter, and source TCP windows advertisement value. For preprocessing purpose, the features which values are categorical values were assigned a key value and stored in a dictionary. In the dictionary object, any values can be stored in an array, and each recorded item is associated with the form of key-value pairs. Furthermore, all the data samples were normalized using the following normal feature scaling process (9):

$$X' = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (9)$$

The X' is the normalized value and x is the original value. The file was saved into a text file for SVM input. In this way, all the data samples were preprocessed in the same pattern.

The total number of data sample in this dataset is 2,537,715 in which around 2.2 million is normal data samples the rest of the sample is attack data samples. In this dataset, the ratio of the normal and abnormal behavior is 87:13. Total normal data samples are 22,18,764, and total attack samples are 321,283.

The second dataset used in this research paper was collected from Canadian Institute of Cyber Security Excellence at the University of New Brunswick [21] upon request. The dataset is named as CIC IDS 2017, which is an Intrusion detection and evaluation dataset specially designed by collecting real-time traffic data flows over seven days that contains malicious and normal behaviors. In this dataset, there are over 2.3 million data samples, and among them, only 10% represents attack data samples. There are 80 network flow features in this data set.

The traffic data samples contain eight types of attacks namely Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web attack, Infiltration & DDoS [21]. This dataset is one of the richest datasets

used for Network intrusion detection research purposes around the world [50]. The goal of using this data set is to evaluate how the proposed method works on different datasets.

4. Proposed Algorithm

The steps of the proposed algorithm are listed below:

1. Define the number of features, N from the feature space.
2. Select the SVM parameter (Gamma, coef (), nu).
3. Define Number of Cooling (nCL) and Equilibrium loop (nEL).
4. Cooling loop Starts {i=1 to nCL}.
   a) Select n sub-features from the set N.
   b) Define an array D_R_array of size 10.
5. Equilibrium-loop starts {j=1 to nEL}.
   a) Train the SVM with the only n selected feature.
   b) Test the learning performance of SVM.
   c) Store the solution in D_L
   d) A small random perturbation of the features
   e) Repeat steps 5a & 5b
   f) Save the Solution in D_R
   g) DR_array [j %10] = D_R
6. if $j \geq 10$ and Standard Deviation (D_R_array)$\leq \sqrt{2}$
   a) Break from Equilibrium and Continue to Cooling loop
7. If $D_R > D_L$, then $D_L \leftarrow D_R$
8. Else
   a) Find the Probability of Acceptance P
   b) Generate Random Number R
   c) If P > R, $D_L \leftarrow D_R$
   d) If P < R, Check if number of Equilibrium loop (nEL)==0
   e) If (nEL)! =0, repeat 5(d), 5(e), 5(f)
   f) Else If (nEL)==0, Then $D_L \leftarrow D_R$
9. Else If (nEL)==0, Reduce Temperature.
10. Equilibrium Loop Ends
11. Check Number of cooling loop (nCL)==0?
12. If (nCL)==0, Done
    Else repeat procedure 4.

The proposed scheme defines the number of features in the dataset. Furthermore, the SVM parameters (Gamma, coef (), nu), Number of cooling loops (nCL), and number of equilibrium loops (nEL) are defined. At first, in the cooling loop, n number of features among N features are randomly selected (as initially, it does not know where to start with) where $n < N$. Then it moves inside the equilibrium loop, and trains the SVM with the selected n number of features and tests the learning performance of SVM. Then the algorithm saves the solution in DL. This solution is considered as an initial solution, and the goal of this loop is to find a best solution for the given temperature. A small random perturbation of the currently held features is made (seems like a random walk in the feature space) to create a new possible solution, because it is believed that good solution is generally close to each other but it is not guaranteed. The algorithm stores the solution in DR. If the cost of the new candidate solution is lower than the cost of the previous solution, then the new solution is kept and replaces the previous solution. Also, if the solution remains within $\pm 2\%$ and factored by ten times in a row, the algorithm will terminate the current equilibrium loop and will check the cooling loop (if nCL$\neq$ 0) and continue another equilibrium loop to save time, as, it is assumed the algorithm is trapped in a local minimum solution. Sometimes the random perturbation results in worse solution; then, the algorithm decides to keep or discard the solution, which depends on an evaluation of the probability function. In case that the new solution is worse than the previous one then, the algorithm generates a random number R and compares with the probability function. If $P > R$, the algorithm keeps the worse solution and if $P < R$, then the algorithm checks whether it has reached the defined number of equilibrium loops or not. If (nEL)==0 then it moves out from the equilibrium loop to cooling loop and restarts the above described procedure again from the cooling loop. If (nEL)! = 0, then the algorithm starts from random perturbation inside the equilibrium loop and performs the procedure again. When the number of cooling loop reaches zero, then algorithm terminates and provide a meaningful solution [51]. This procedure is consuming less time and does not need and specific hardware configuration.

Table 1 2-feature subset combinations.

#	**Features in this Combination** [19] [20]
1	• The IP address of Source. • Service used (HTTP, FTP, SMTP, ssh, DNS, FTP-data, IRC and (-) if not much-used service)
2	• Number of connections of the same source IP and the destination IP address in 100 connections according to the uncompressed content size of data transferred from the HTTP service last time. • Destination interpacket arrival time (mSec)
3	• Source TCP based sequence Number. • Some flows that have methods such as Get and Post in HTTP service.
4	• Destination TCP based sequence Number. • Mean value of the flow packet size transmitted by the Destination
5	• Actual uncompressed content size of data transferred from the HTTP service • Number of connections is the same destination address & the source port in 100 connections according to the uncompressed content size of data transferred from the HTTP service.
6	• Source Jitter (mSec). • Source retransmitted packet.
7	• Mean of the flow packet size transmitted by the destination • Total packet count from Source to destination.
8	• Destination bits per the second. • Source interpacket arrival time (mSec).
9	• TCP connection setup time, the time between the SYN_ACK and the ACK packets. • Actual uncompressed content size of the data transferred from the server's HTTP service.
10	• Source interpacket arrival time (mSec) • Mean of the flow packet size transmitted by the destination.

This algorithm does not search the whole work/feature-space for reaching the global optimum solution. As a result, a small amount of time is required to provide a reasonable solution. It performed relatively well in large datasets. The efficacy of the proposed solution depends on the selection of essential features that help the intrusion detection process to detect an anomaly accurately. To assure that the algorithm provides a better solution compared to other machine learning methods, it has been tested on two different types of datasets, which was explained in the previous chapter.

5. Experiments and Results

This section is divided into several sections. Both datasets contain different numbers of features. The results are discussed in details using multiple feature subsets (Ex. 2, 3, 4 and 5).

5.1. Simulation Setup and outcomes of the Proposed Algorithm (2 features UNSW Dataset)

In this simulation setup, the proposed algorithm has been applied taking two feature subsets as an initial approach. Afterwards, we will try three, four and five feature subset combinations to inspect how the algorithm performs if the number of features increases in a subset. Table 1 shows 2-feature subset combinations:

The detection accuracy of the proposed method versus the feature combination is shown in Fig. 1.

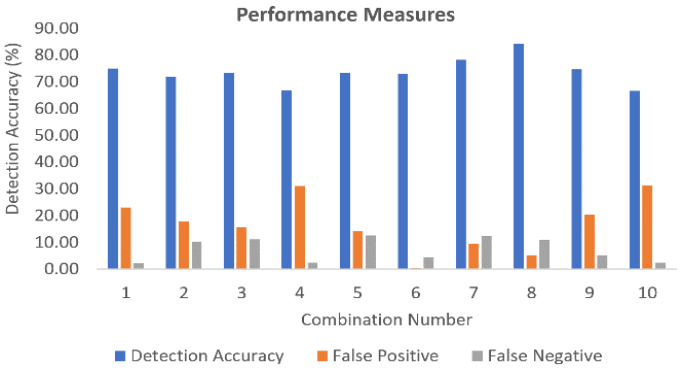


Fig. 1 Detection Accuracy, False Positive and False Negative for 2 feature subset combination.

Table 1 shows the combination number denotes which two features were selected for that combination. The proposed algorithm achieved a detection accuracy recorded as 84.14% when combination number 8 was selected. The lowest detection accuracy among the given results was recorded as 63.56% for combination number 10.

Considering only two feature subsets provided a high false positive and false negative rate of 5.09% and 10.77% respectively. Upon checking the ROC curve (Fig. 2), these combinations provided poor AUC value. It may be caused by taking an imbalanced number of features. The feature subset is selected by the SA process, which may not be linearly separable on the classification space. Using a non-linear typical decision boundary will require more computational efforts than fitting linear decision boundary. Thus, increasing the dimension (or features) may provide better results compared to the 2-feature subsets, as it may allow the hyperplane to separate the data.

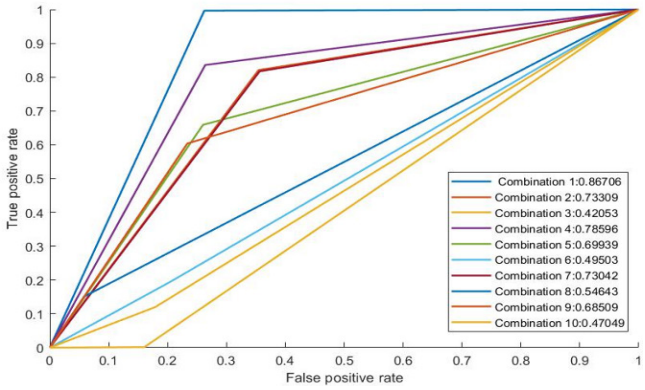


Fig. 2 RoC analysis 2-feature subset.

The F1-score is a statistical analysis of binary classification and a measure of precision. The F1-score can be described as a harmonic mean of the precision and recall, where an F1-score reaches its best value at one and worse at 0. Precision is the ratio of correctly projected positive annotations to the total projected positive annotations, whereas recall is the ratio of correctly predicted positive annotations to all annotations in the actual class. The reason why harmonic mean is considered as the average of ratios or percentages is considered. In this case, harmonic mean is more appropriate then the arithmetic mean. As shown in fig 2**Error! Reference source not found.**, we plotted the F1-score for each combination of 2-feature sets (Table 1). As shown, the F1-score is quite low, except for combination 8, where it achieved almost 0.92; otherwise it was an average of about 0.81. Based on our other subsequent experiments, these values were deemed too low, and more features per combination were apparently required to achieve better F1-score results.

5.2. Simulation Setup and outcomes of Proposed Algorithm (3 features UNSW Dataset)

In the previous section, the proposed algorithm was tested and results analyzed for two feature subsets. While this seemed fine for an initial approach, we found the F1-score results to be quite low, and so we now present results for more features per combination. In particular, Fig. 3 shows the 3-feature subset combinations. Afterwards, we present the results for three, four and five feature subset combinations to inspect how the algorithm performs.

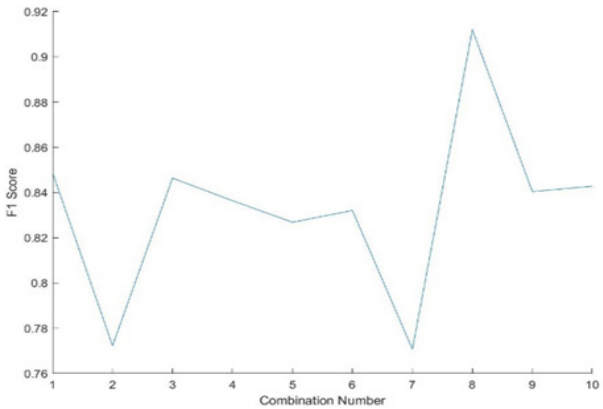


Fig. 3 F1-score of the 2-feature subset combination.

Table 2 3-feature Subset Combination.

#	**Features in this Combination** [19] [20]
1	• The IP address of Source • Service used (HTTP, FTP, SMTP, ssh, DNS, FTP-data, IRC and (-) if not much-used service). • Source packets retransmitted or dropped.
2	• Source TCP window advertisement value • No of connections of the same source IP and the destination IP address in 100 connections according to the uncompressed content size of data transferred from the HTTP service last time • Destination interpacket arrival time (mSec)
3	• Source TCP based sequence Number • Some flows that have methods such as Get and Post in HTTP service. • No. of connections of the same destination address in 100 connections according to the uncompressed content size of data transferred from the HTTP service last time
4	• Destination TCP based sequence Number. • Source TCP based sequence Number. • Mean value of the flow packet size transmitted by the Destination.
5	• Actual uncompressed content size of data transferred from the HTTP service. • TCP base sequence number of destinations. Number of connections is the same destination address & the source port in 100 connections according to the uncompressed content size of data transferred from the HTTP service
6	• Source Jitter (mSec). • Source retransmitted packet. • Destination bits per second.
7	• Source Jitter (mSec) • Mean of the flow packet size transmitted by the destination • Total packet count from Source to destination
8	• Destination bits per the second • Source interpacket arrival time (mSec) • Number for each state dependent protocol according to a specific range of values for source/destination time to live
9	• TCP connection setup time, the time between the SYN_ACK and the ACK packets. • Source bits per second • Actual uncompressed content size of the data transferred from the server's HTTP service.
10	• Destination Interval arrival Time • Source interpacket arrival time (mSec) • Mean of the flow packet size transmitted by the destination.

A list of ten combinations has been shown, and each combination contains three features. The mechanism of the initial approach in the research was that it takes three features at a time and provide an output. Afterwards, it discards the results and tries another randomly selected feature combination. Trying all possible combination implies that it leads to an exhaustive search and takes a huge amount of computational resources and consumes more time. Further, it will more time if a large number of features are considered. However, in the proposed algorithm, the annealing process starts selecting a random set of three features as the first step as it has to begin somewhere randomly. Then the SVM is trained only with these three features, and the generated output is saved. It is considered as a first initial solution. A small random perturbation is made to the current solution changing one or two features because it is assumed that good solutions are generally close to each other, but it is not guaranteed as the best solution. Sometimes the newly generated solution results in a better solution than the algorithm keeps the new solution. If the newly generated solution is worse than the current solution, then the algorithm decides whether to keep or discard the worse solution, which depends on the evaluation of the probability function (7). The higher temperature in the annealing process, it is more likely the algorithm will keep the worse solution. Keeping the worse solution allows the algorithm to explore the solution space and to keep it within the local optima. Also neglecting a worse solution lets the algorithm to exploit a local optimum solution, which could be the global solution for that temperature.

The detection accuracy of the proposed method versus to the feature combination is shown in Fig. 4. In Fig. 3, the combination number denotes which three features were extracted for that combination.

The proposed algorithm achieved a detection accuracy recorded as 98.32% when combination number 8 was selected (Fig. 3). This combination contains three essential features, such as destination bits per second, source interpacket arrival time (mSec), and number of each state dependent protocol according to a specific range of values for source/destination time to live value. The lowest detection accuracy among the given results was recorded as 72.08% when combination number 4 was used. Comparing with the 2-feature subset combination results, we can infer that as another dimension was introduced, the linear classification technique worked better.

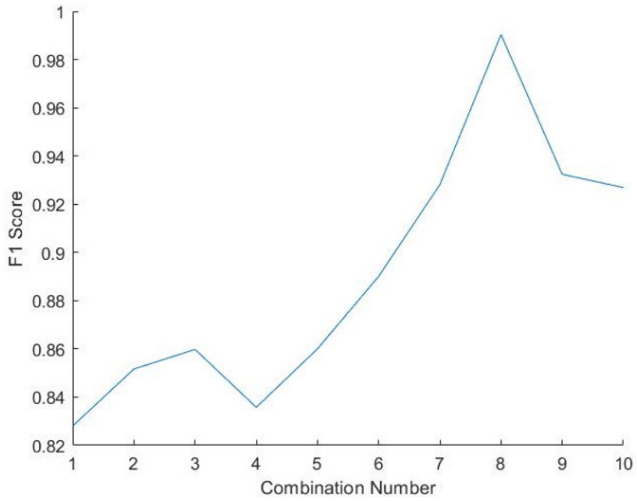


Fig. 4 Detection Accuracy, False Positive and False Negative for 3-feature subset combination.

Furthermore, the performance metrics of the proposed algorithm was investigated. The false positive refers to a situation that the system incorrectly identifies an object as a positive (attack), which, however, is not an attack and is a normal (non-

attack) object. The false negative refers to a situation that the system incorrectly identifies an object as a negative (non-attack), which however is an attack. Also, the percentage of false positives and the percentage of false negatives are shown for the proposed scheme. The false positives and false negatives were recorded as 1.49% and 0.19% respectively, for combination number 8. Therefore, it can be inferred that, if the correlative features are extracted, the false positive and negative rate decreases, it may provide a relatively better solution for intrusion detection.

In Fig. 5 the receiver operating characteristic curves for the proposed schemes are presented.

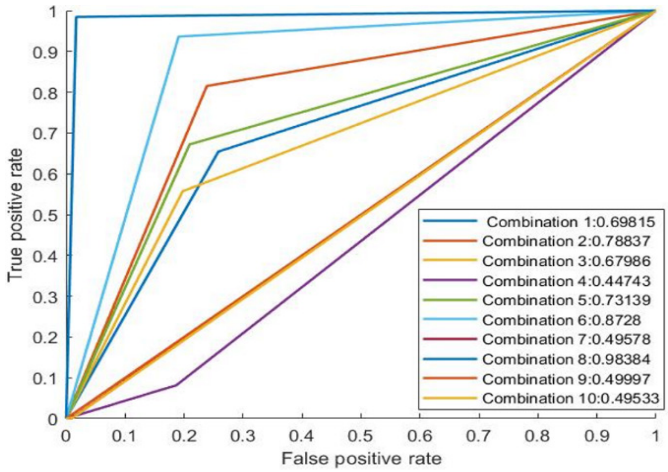


Fig. 5 RoC curve analysis for 3-feature subset.

In a real-world scenario, the misclassifications costs are difficult to determine. In this regard, the ROC curve and its related measures, such as AUC (Area under Curve) can be more meaningful and deemed as vital performance measures. As shown in the figure, it is seen that the performance output is much better than the previous performance with 2-feature combinations. The combination number 8, which contains features such as Destination bits per second, Source interpacket arrival time (mSec), Number of each state dependent protocol according to a specific range of values for source/destination time to live value, provided much better reasonable output. The AUC of that combination is 0.98384, which is closer to one, which presents a better solution in contrast with the other possible solutions.

Fig. 6 F1-score 3-feature subset.

The F1-score of the different combinations is shown in Fig. 6. Note that the combination number 8 achieved the highest F1-score of 0.99.

Fig. 7 demonstrates the detection accuracy of the proposed scheme increases or decreases along as the number of iterations varies.

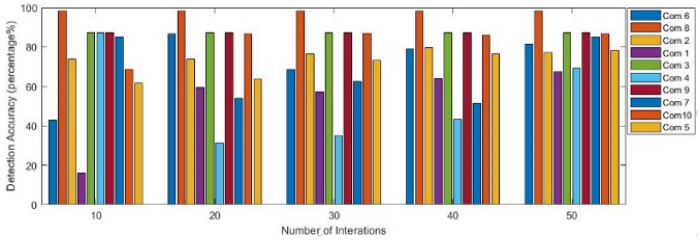


Fig. 7 Detection accuracy difference versus number of iterations.

A cross-validation mechanism was used to assess the predictive performance of the model to evaluate how good the model works on a new independent dataset (using resampling randomly). It is seen that combination 1 (violet) started with very low detection accuracy, but while iterating multiple times the detection accuracy increased, and an average of those accuracies provide a well reasonable solution for that combination.

However, combination 8 (light brown) started with higher accuracy, and while multiple iterations are running, it kept almost the same as it is and an average of those accuracies has been considered. Furthermore, the combination number 4 (sky blue) stared with higher accuracy but while multiple times iterations the accuracy went down and suddenly went up. An average of those accuracies has been considered. Taking averages of the detection accuracies allows the algorithm to be more confident on the provided output. The proposed scheme does not take too long to converge to the local optima (which can be the global optima) but provides a reasonable detection accuracy over short possible time, depending on the system resources.

To evaluate whether the algorithm works as intended, unit testing was performed, and the detection accuracy was measured for different numbers of thermal equilibrium loops, as shown in Fig. 8.

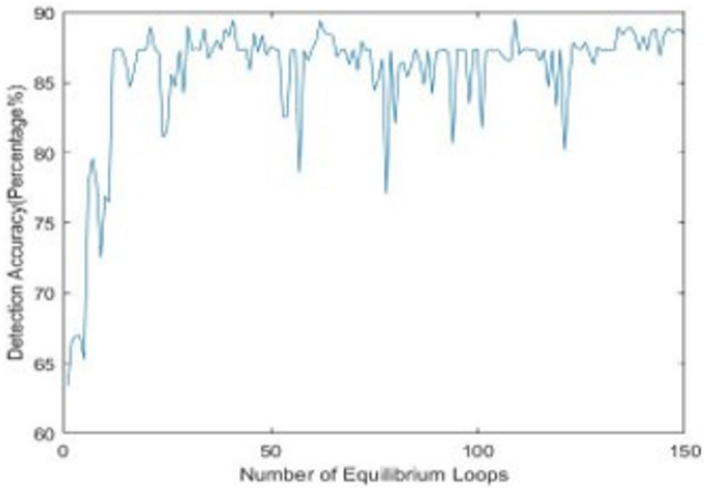


Fig. 8 Performance of the proposed scheme for varying numbers of thermal equilibrium loops.

The annealing process starts at a random direction as it does not know which direction to start with and the first solution is considered as the initial solution. On each iteration, a random perturbation in the solution space is performed, and a new solution is generated. Comparing with the previous solution it is better or worse, it keeps the better solution and marches forward, as it is believed the good solutions may be nearby. Fig. 8 shows that the algorithm starts with a low detection accuracy and marching forward it is going towards higher accuracy. While going forward, it seems that it may have found worse candidate solution (the down hikes on the Fig. 8 but keeping the worse solution allows the algorithm to explore more in the solution space to reach an optimum solution. These observations of the results verify that the algorithm is performing the way it supposed to perform.

As the algorithm starts from a random direction, provides an initial solution, takes random perturbation, and generates another new solution, a scenario is visualized that sometimes the detection accuracy of the old candidate solution and new candidate solution is quite close enough like a slight difference of some percentage. For the time being, it is being trapped on that local optima. So, to save more time and speed up, the algorithm is designed in such way that, if the difference of the accuracy of the old candidate solution and new candidate solution is around ±2% and ten times in a row, the algorithm will terminate that equilibrium loop. Subsequently, the algorithm will go to the next cooling loop (if nCL!=0) and start another inner loop and continue. Fig. 9 shows that inside the equilibrium loop iteration number approximately from 139-150, the detection accuracy stays between 86%-87% for the time being, and it detected an accuracy difference of ±2% ten times in a row. Therefore, the proposed method will terminate the current equilibrium loop, decrease temperature and start the next cooling loop if nCL!=0.

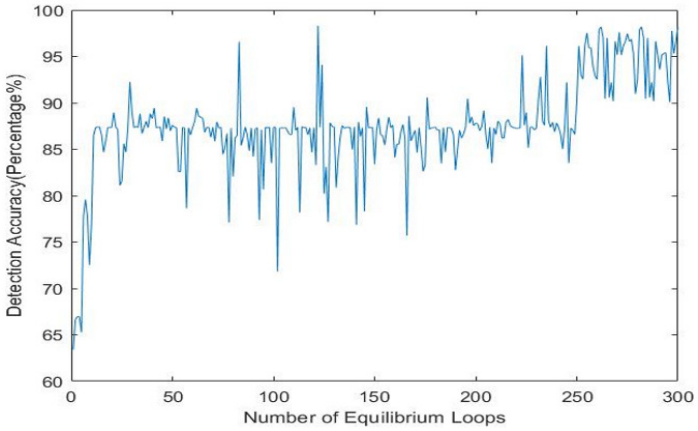


Fig. 9 Detection accuracy of the equilibrium loops.

In a nutshell, the proposed algorithm starts from a random path to find an initial solution (initial 3-feature combination) and takes small random walk in the feature space (changing 1 or more features) and compares with the previous solution that allows the algorithm to converge faster to the local optima which may likely be a local optimum solution.

5.3. Simulation Setup and outcomes of Proposed Algorithm (4 features UNSW Dataset)

Next, the performance of the algorithm was evaluated taking four and five features in a subset. The algorithm will initiate with a four-feature subset combination. Furthermore, the outcomes with be compared with the 3-feature subset combinations. The comparison allows us to determine how the algorithm performs when the number of features increases. The number of training samples and the number of testing samples were kept the same as previous. Table 3 shows the combination of four features.

Table 3 4-feature subset combination.

#	**Features Taken** [19] [20]
1	• Source IP address • Source packets retransmitted or dropped • Destination to the source packet count • No of connections of the same source address (1) and the destination port (4) in 100 connections according to the last time (26).
2	• Source inter-packet arrival time (mSec) • Destination inter-packet arrival time (mSec) • If the FTP session is accessed by user and password then 1 else 0. • No. of connections that contain the same service (14) and source address (1) in 100 connections according to the last time (26)
3	• Source to destination time to live value • Source TCP sequence number • No. of flows that has methods such as Get and Post in HTTP service. • No of flows that have a command in an FTP session.
4	• Destination IP address • Destination TCP window advertisement value • Source TCP sequence number • d. Mean of the flow packet size transmitted by the destination
5	• The content size of the data transferred from the server's HTTP service. • No of connections of the same destination address (3) and the source port (2) in 100 connections according to the last time (26). • No. of connections that contain the same service (14) and destination address (3) in 100 connections according to the last time (26). • d. Source TCP window advertisement value
6	• Source jitter (mSec) • Destination bits per second • No. of connections that contain the same service (14) and source address (1) in 100 connections according to the last time (26). • d. No of flows that have a command in an FTP session.
7	• Source to destination packet count • Mean of the flow packet size transmitted by the destination • The time between the SYN_ACK and the ACK packets of the TCP. • d. No. of flows that has methods such as Get and Post in HTTP service.

8	• No. for each state according to a specific range of values for source/destination time to live value • Destination bits per second • Source bits per second • d. The time between the SYN and the SYN_ACK packets of the TCP.
9	• The content size of the data transferred from the server's HTTP service. • Source bits per second • HTTP, FTP, ssh, DNS., else (-) • Mean of the flow packet size transmitted by the destination.
10	• No. for each state according to a specific range of values for source/destination time to live value • Source bits per second • Destination jitter (mSec) • d. If the source equals to the destination IP addresses and port numbers are equal, this variable takes value one else zero

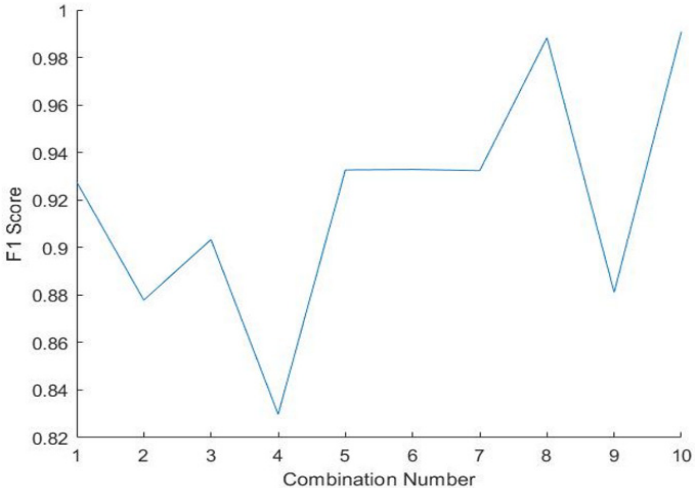


Fig. 10 Detection accuracy for the four-feature combination.

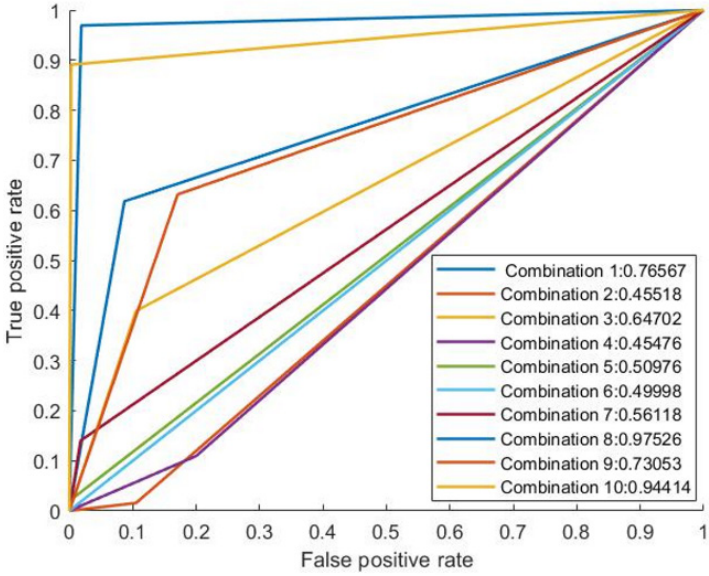


Fig. 11 RoC curve and AUC for the four-feature combination.

Evaluating the above results, we can see that taking four features increases some of the combinations' detection accuracies. For example, comparing 3-feature combination number 8 of Fig. 3 with 4-feature combination number 8 of Table 3, the detection accuracies were 98.32% and 97.97% respectively, resulting in a 0.35% positive difference in detection accuracy. The algorithm kept Destination bits per second and Number of each state dependent protocol according to specific range of values for source/destination time to live value features steady and while taking 4 feature combination it selected two features such as Source bits per second, the time between the SYN and the SYN_ACK packets of the TCP which allowed the algorithm to provide a reasonable solution. Therefore, the 3-feature subset combination 8 is better regarding accuracy, time consumption, low false positive, higher AUC value and higher F1-score compared to 4-feature subset combinations 8.

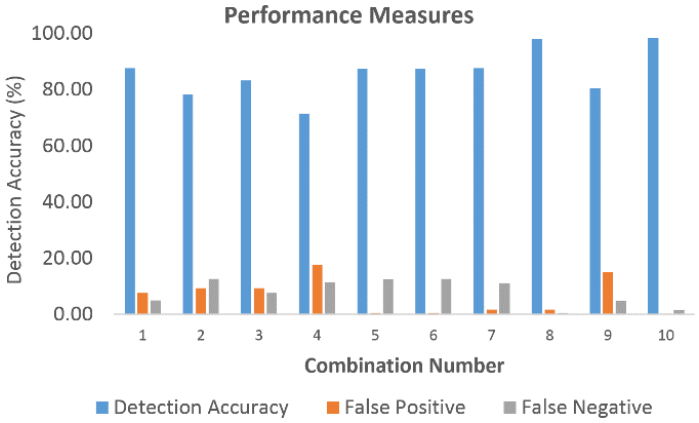


Fig. 12 F1-score for the 4-feature subset combination.

However, if we look at the combination number 10 when 4 features are selected such as No. for each state according to specific range of values for source/destination time to live value, Source bits per second, Destination jitter (mSec) and If source equals destination IP addresses and port numbers are equal-this variable takes value 1 else 0, provided a detection accuracy of 98.39% which is the highest accuracy the algorithm provided. Comparing with the 3-feature combination number 8, the difference of detection accuracy is 0.07% only (3 feature detection accuracy is 98.32%). Increasing dimension (SA) allowed the feature space hyperplane to separate the normal and attack sample more precisely for this particular feature subset compared to 3-feature combination number 8 (Fig. 3). However, regarding AUC, ROC, false negatives and F1-score the 3-feature combination number 8 (Fig. 3) overall was the most reasonable solution here.

For further evaluation of the performance to determine the effect of providing more features, we have analyzed the behavior of the proposed scheme taking five feature subset combinations. It is described in the following section.

5.4. Simulation Setup and outcomes of Proposed Algorithm (5 features UNSW Dataset)

In this setup, five feature subset combination were considered, and the performance of the proposed method was evaluated regarding detection accuracy, false positive and negative, F1-score, ROC characteristics, Area under the curve.

Analyzing the performance of the proposed scheme while considering a 5-feature subset, the detection accuracy of most of the feature combinations reduced significantly. This reduction in detection may be caused because of insignificant and irrelevant dimensions or features in the dataset were considered in a feature subset. The algorithm may not have been trained appropriately due to a large number of dimensions considered in this step. As a result, the algorithm was unable to provide a reasonable solution in most of the feature combinations.

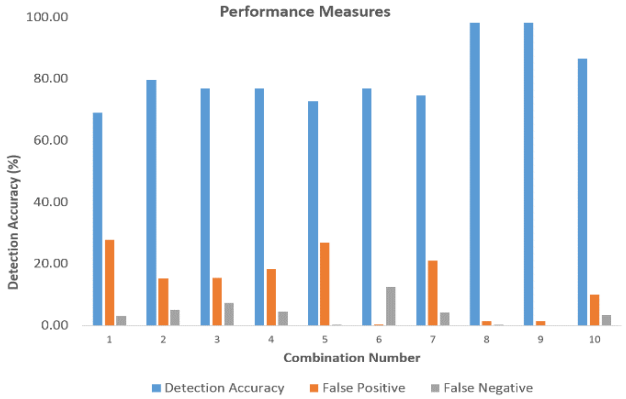


Fig. 13 Detection Accuracy for 5 feature subset combination.

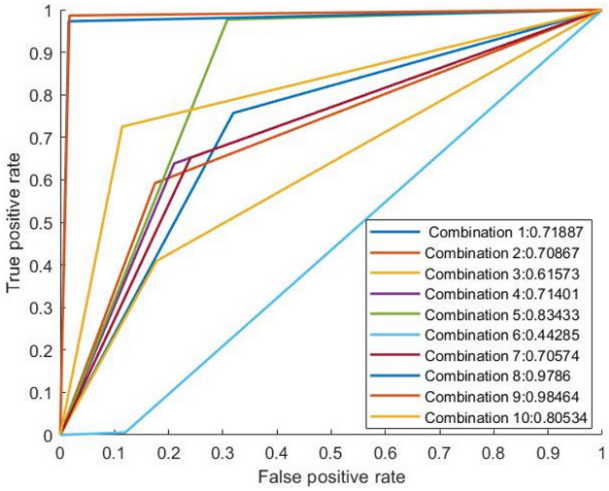


Fig. 14 RoC Analysis of 5-feature subset (UNB Dataset).

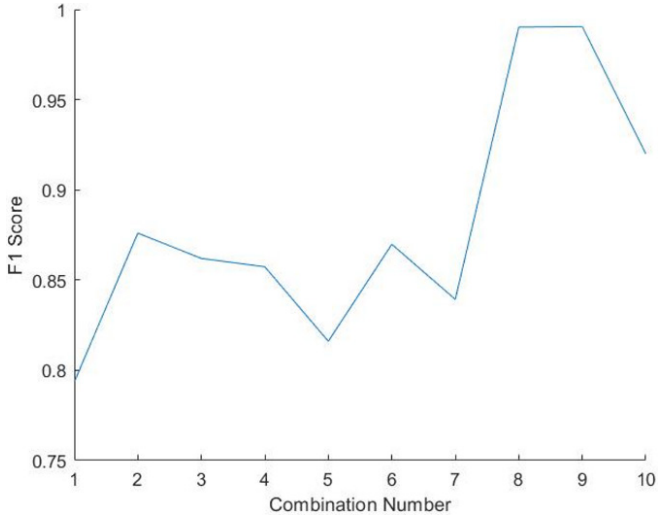


Fig. 15 F1-score for 5 feature subset combination.

However, it is observed that combination number 8 and 9 provided a better detection accuracy compared to other feature subset combinations. The 5-feature subset (combination 9) provided an accuracy of 98.34%, which is only .02% higher than the 3-feature subset (combination 8, Fig. 3). The false positive and false negative rate of combination 3-feature subset (combination 8, Fig. 3) is 1.49% and 0.19%, and 5-feature subset (combination 9) is almost same as 1.49% and 0.17% respectively. There were no significant changes in the parameters. The difference of the AUC value of 5-feature and 3-feature subset is 0.001, which is very negligible. Regarding F1-score, the 3-feature subset combination 8 (Fig. 3) provided better outcomes compared to 5-feature subset combinations. It is obvious that if the 3-feature subset is considered, the algorithm will take a shorter time to provide a solution compared to with that of 5-feature subset combination.

5.5. Simulation Setup and outcomes of Proposed Algorithm (3 features UNB Dataset)

The dataset collected from Canadian Information Security Center of Excellence at the University of New Brunswick (upon request) to analyze the behavior of the proposed scheme. Recently in 2017, they experimentally generated one of the richest Network Intrusion Detection dataset containing 80 network flow features, which were generated from capturing daily network traffic. Full details were provided on [21]. This paper also provided information that what features are very much significant to detect a specific type of attack. This dataset was considered to evaluate that how the proposed scheme will work on an entirely different dataset and how much confidence the algorithm has on its provided outcomes.

This dataset also contained some features similar to the previous UNSW dataset. The dataset contains categorized and continuous variables. It was preprocessed similar way with the feature scaling process equation 12. Detailed experimental results are discussed below.

Table 4 3-feature subset combination for UNB dataset.

#	**Features in this combination** [21]
1	• backward packet length min • total length flow packets • c. flow inter-arrival time min
2	• flow inter-arrival time min • Initial win forward bytes • Flow inter-arrival time std.
3	• flow duration • Active min of bytes • the active mean of flow bytes
4	• backward packet length std • Length of forwarding packets • sub-flow of bytes
5	• avg packet size • Backward packet length std • mean of active packets

6	• flow inter-arrival time min • Backward inter-arrival time means • initial win forward bytes
7	• forward push flags • Syn flag count • c. back packets/s
8	• backward packet length std • Avg packet size • flow inter-arrival time std
9	• forward packet length mean • Total length forward packets • sub-flow forward bytes
10	• initial win forward bytes • Backward packets/s • flow inter-arrival time std

Upon evaluation of the above outcomes, it is observed that the detection accuracy of the algorithm increased in some feature combinations and was able to provide a detection accuracy of 96.58% (Combination 8). The Receiver operation characteristics show that the algorithm was provided with a reasonable AUC value (combination 8) compared to other methods discussed previously. In the research paper [21] Table 3, the author provided a particular set of feature set which is more significant detecting a particular intrusion such as for the detection of a DDoS attack, backward packet length, average packet size and some inter-arrival time of the packets, flow IAT std. Now the proposed scheme selected three features out of the four features (mentioned in that paper by the author) shown on combination number 8 as the evaluation was done taking three features at a time. These three features provided an accuracy of 96.58%, an AUC value of 0.92199 and an F1-score of 0.980416. If the full dataset with all 80 features [21] were available, then the results might get better than the current one.

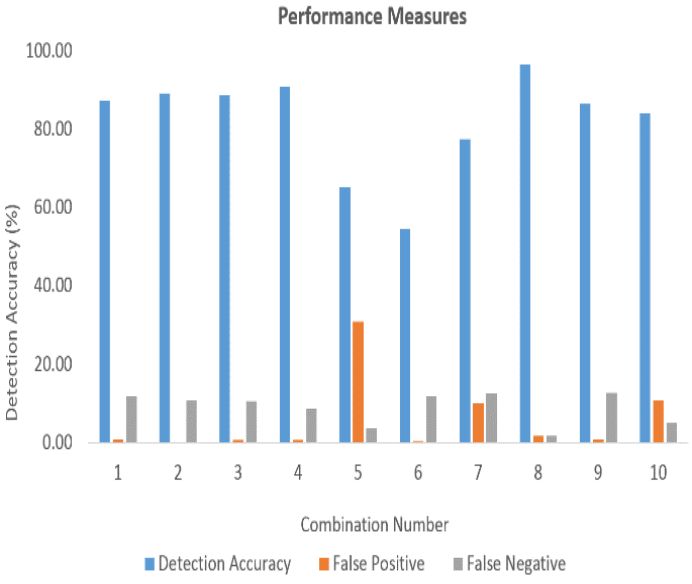


Fig. 16 Detection accuracy for the 3-feature subset combinations (UNB Dataset).

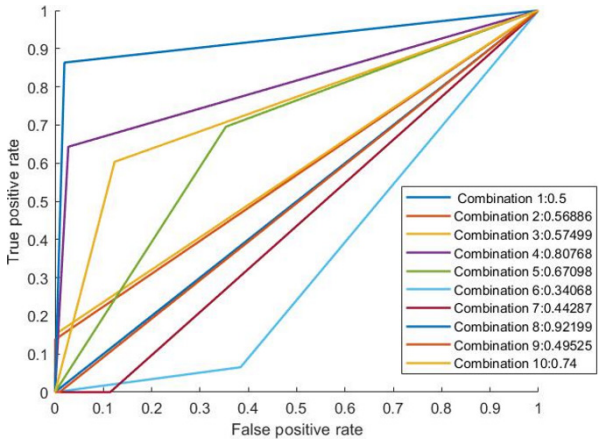
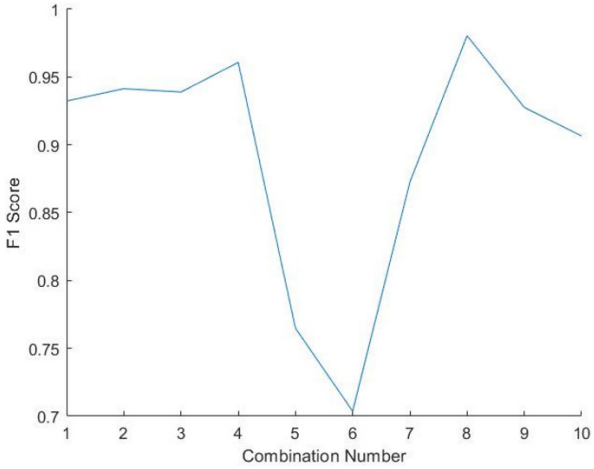


Fig. 17 RoC Analysis of 3-feature subset (UNB Dataset).



F1-score for the 3-feature subset (UNB Dataset).

Upon evaluation of the above outcomes, it is observed that the detection accuracy of the algorithm increased in some feature combinations and was able to provide a detection accuracy of 96.58% (Combination 8). The Receiver operation characteristics show that the algorithm was provided with a reasonable AUC value (combination 8) compared to other methods discussed previously. In the research paper [21] Table 3, the author provided a particular set of feature set which is more significant detecting a particular intrusion such as for the detection of a DDoS attack, backward packet length, average packet size and some inter-arrival time of the packets, flow IAT std. Now the proposed scheme selected three features out of the four features (mentioned in that paper by the author) shown on combination number 8 as the evaluation was done taking three features at a time. These three features provided an accuracy of 96.58%, an AUC value of 0.92199 and an F1-score of 0.980416. If the full dataset with all 80 features [21] were available, then the results might get better than the current one.

5.6. Performance Comparison

In this section, we have analyzed the performances of the proposed method in detail in comparison with general SVM and

Decision tree-based detection method concerning the performance matrices.

In the initial approach of this research [52] was at first the algorithm selects n features using random combination from the N features from the dataset. Then it selects the SVM parameters and trains the classifier. Afterwards, the algorithm rum SVM on the test data samples and tries to identify the positive and negative training examples. It discards the previously selected feature set, takes another random combination of features, and continue the process until all possible combinations (depending on the number of feature input) are completed. This process leads the algorithm to perform an exhaustive search in the whole workspace trying all possible combinations leading to a combinatorial optimization problem. In combinatorial optimization, it is very difficult to determine how many number or which feature combination set will provide a reasonable solution minimizing the cost and decision-making time. When the number of features in a subset increases, the number of combinations also increases exponentially. The performance of the initial algorithm is presented below where the algorithm provides outputs of different random feature combinations and its very random.

We can see from Fig. 18 that the output of different 3-feature subset is very random and the algorithm is performing an exhaustive in the whole workspace to provide an optimal solution. For more substantial dataset form the current one, it will be much exhaustive for the algorithm to provide a solution. Thus, the combinatorial optimization exists in this scenario.

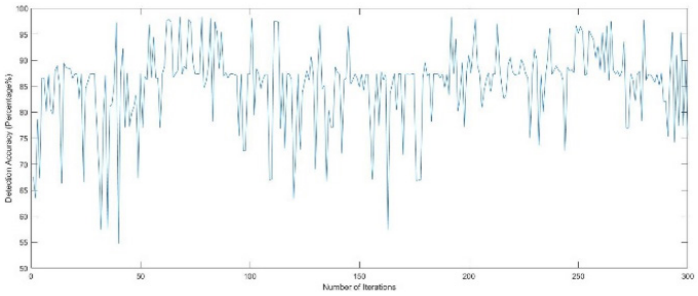


Fig. 18 Performance of the initial approach (exhaustive search).

In contrast with the above discussion, the proposed algorithm does not search the whole workspace for finding a solution rather than it starts from a random direction at the beginning as initially, the algorithm does not know which way to start from. In the outer loop known as known as a cooling loop of SA, depending on the feature input (2,3,4 or more) the algorithm selects a random subset of feature to start with and trains the SVM (cost function of SA) only with the selected features at that temperature. Inside the equilibrium loop, the first solution is considered as the initial one. The cost function determines the goodness of the solution. Afterwards, the algorithm takes a small random perturbation to create a new candidate solution because it is assumed that good solutions are generally close to each other, but it is not guaranteed as the best optimal solution. If the newly generated solution is worse than the current solution, then the algorithm decides whether to keep or discard the worse solution, which depends on the evaluation of the probability function (equation 7). After running the inner loop many times, wherein each loop it takes a new better solution or keeps a worse solution, the algorithm may be viewed as taking a random walk in the solution space to find a sub-optimal solution for the given temperature. The performance inside the equilibrium loop is shown in fig. 18.

If the comparison is made between Fig. 18 and 19, we can see that the proposed method outcome does not show high variance like the Fig. 18 (variance of the Fig. 18 is ± 3.9% and the current is ±1.9%). The initial algorithm is performing an exhaustive search over the whole workspace, finding the global solution whereas the proposed method does not search the whole workspace, and providing the best possible local optimum solution for the given temperature, which may be the global optimum solution. The proposed method consumes less time and less computational cost compared to the exhaustive search method.

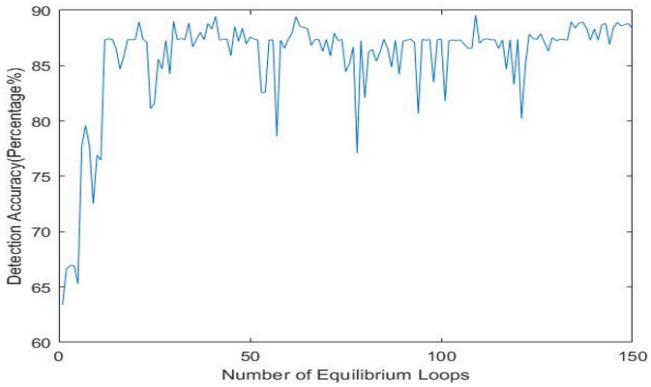


Fig. 19 Performance of the Proposed method.

Sometimes inside the equilibrium loop, it is observed that the difference of the solution outcomes is not that significant. After visual observation of the cost function inside the equilibrium loop, it is observed that the output of the variance of the outcomes lies below 3%. It looks like the algorithm is trapped inside that local optimum solution until the inner loop ends. To save time and speed up the performance, the algorithm is designed in such way that, if the difference of the accuracy of the old candidate solution and new candidate solution is around ±2% factored by ten times in a row, the algorithm will terminate that equilibrium loop. Subsequently, the algorithm will go to the next cooling loop (if nCL! = 0) and start another inner loop and continue. A general standard deviation method was applied to the steps of the algorithm in fig. 20.

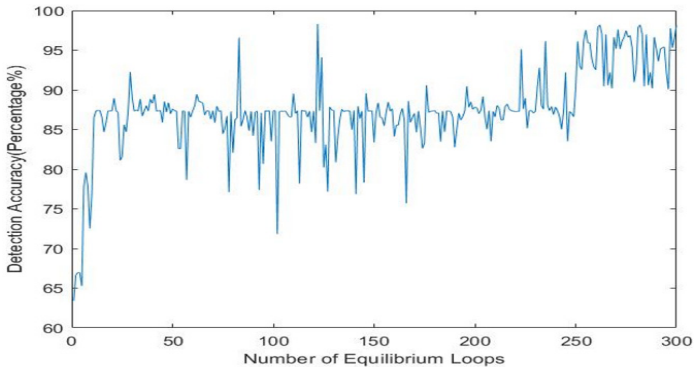


Fig. 20 Termination of equilibrium loop.

If another dataset is considered which contains many features, the number of combinations will increase, and the initial algorithm will take more time, and the computational cost will be more. The proposed method does not try all possible combinations to provide a solution. It provides the best local optimum solution for the given temperature, and it may be the global optimum solution. The proposed methodology can be considered as a general methodology and can be applied t o other sectors to solve the feature extraction combinatorial optimization problem to save time and computational resources. Notably, the proposed method saves time and computational resources and provides a better outcome compared to the general SVM based exhaustive search method.

5.7. Performance comparison with Decision tree-based method

In this section, we have evaluated the performance of a decision tree-based method and compared the outcomes with the proposed method. There are several algorithms based on decision trees like C4.5, CART, and ID3. In this research, we have applied the CART (Classification and Regression Tree) algorithm [44] for classification purposes. For a fair comparison with the proposed method, the same feature set was applied to the decision tree-based method on the same dataset (UNSW dataset). Furthermore, the decision tree-based methodology was applied in the second dataset (UNB dataset) to evaluate how the decision tree algorithm performs on two different types of dataset. We have evaluated the performance of a decision tree-based method on same 3-feature subset combinations (Fig. 3 for UNSW dataset and Table 4 for UNB dataset).

In the Fig. 20, the detection accuracy, false positive and false negative rate were observed when the CART algorithm was applied on the UNSW dataset. The 3-feature subset combinations were kept same as previous. Let us compare the outcomes of the decision tree method with the 3-feature subset combination outcomes of the proposed method.

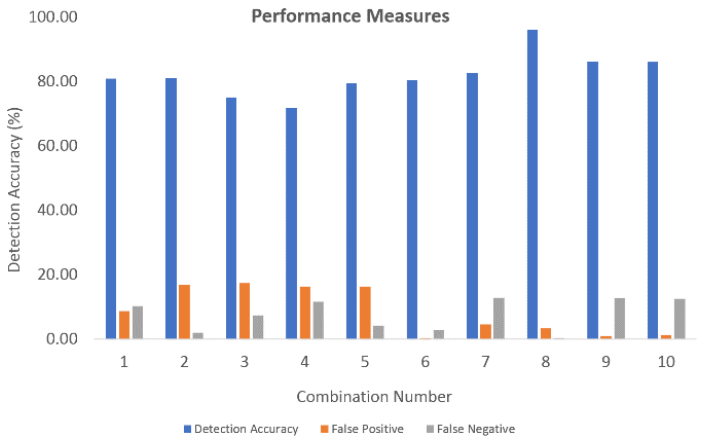


Fig. 21 Performance metrics of the decision tree-based method (UNSW dataset).

Upon evaluation of the performance metrics of the decision tree-based method and the proposed SA-SVM based method we can see that, combination number 8 with three feature subsets such as Destination bits per the second, Source interpacket arrival time (mSec) and No. for each state dependent protocol according to a specific range of values for source/destination time to live value, the detection accuracy of the decision tree is 96.16%, false positive and false negative rate is 3.45% and 0.39%, whereas the proposed method provides a higher detection accuracy of 98.32%, false positive rate of 1.49%, false negative rate of 0.19%. The proposed method provided better results compared to the decision tree-based method in most of the subset combinations. These empirically validated outcomes show that the proposed method outperforms the decision tree-based method.

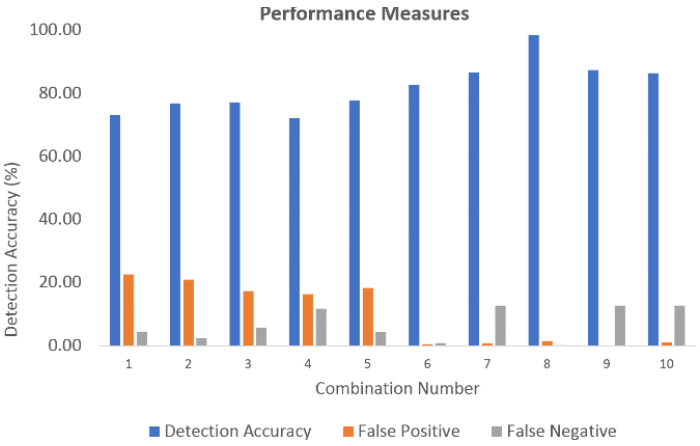


Fig. 22 Performance metrics of the proposed method (UNSW dataset).

Fig. 23 represents the F1-score comparison between the decision tree-based method and the proposed method.

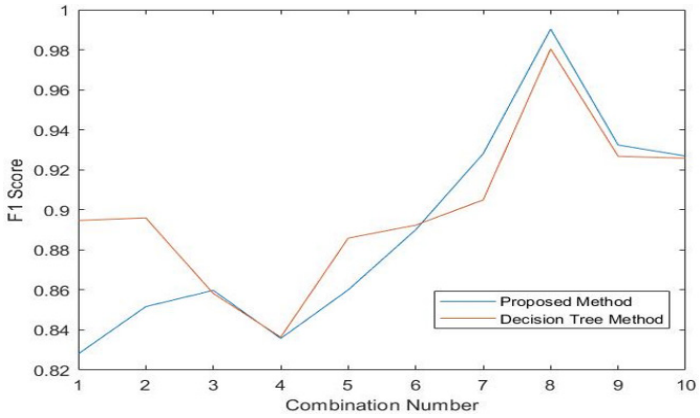


Fig. 23 F1-score comparisons (UNSW dataset).

The F1-score is a statistical measure of tests accuracy that measures how accurate is the classifier; meaning how many instances it classifies correctly. It also tells how robust the classifier is meaning it does not miss a significant number of instances. We can see from Fig. 24 that combination 8 (Fig. 3) the decision tree-based method provided an F1-score of 0.980401 whereas the proposed method provided an F1-score of 0.990522. The proposed method provided higher F1-score compared to the decision tree-based method.

For further performance evaluation, the decision tree-based method was applied on the UNB 2017 dataset also, and the outcomes were compared with the proposed method. The three-

feature subset was kept the same for the evaluation. We will analyze how decision tree performs on a different dataset.

In Fig. 24 the detection accuracy, false positive and false negative rate were observed when the CART algorithm was applied on the UNB dataset. The 3-feature subset combinations were kept same as previous. Let us compare the outcomes of the decision tree method with the 3-feature subset combination outcomes of the proposed method.

Fig. 24 Performance metrics of the decision tree-based method (UNB dataset).

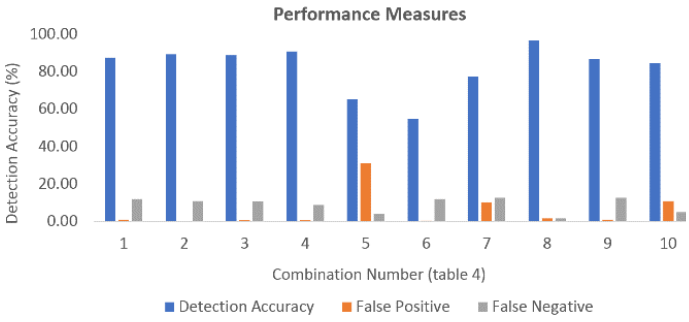


Fig. 25 Performance metrics of the proposed method (UNB dataset).

Upon analysis of the performance metrics of the decision tree-based method and the proposed SA-SVM based method on the UNB dataset, combination number 8 with three feature subsets such as flow inter-arrival time, flow back packet length, flow duration represents DDoS attack [76] the detection accuracy of the decision tree is 92.22%, false positive and false negative rate is 5.42% and 2.36%, whereas the proposed method provides a higher detection accuracy of 96.58%, false positive rate of 1.70%, false negative rate of 1.73%. The proposed method provided better results in contrast to the decision tree-based method in most of the subset combinations when a different dataset was used. These empirically validated outcomes show that the proposed method outperforms the decision tree-based method on UNB dataset also.

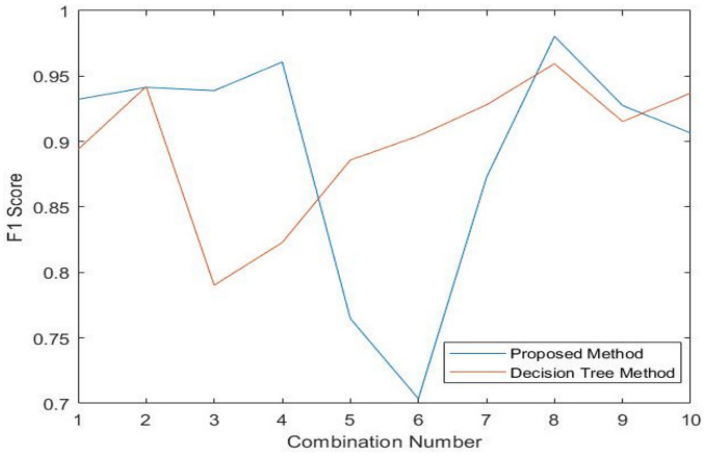


Fig. 26 F1-score comparison (UNB dataset).

6. Conclusions and Future Works

This paper provides a computationally intelligent approach, merging machine learning & combinatorial optimization, to provide a comparatively reasonable solution for optimum feature extraction from a large set of features, which is a difficult combinatorial optimization problem. The experimental results show that this proposed scheme provides better outcomes in comparison with other machine learning methods like Decision Trees and general SVM alone based approaches. As the results showed, this methodology performed effectively and efficiently in contrast with other machine learning methods while applying on different network intrusion datasets. Empirically validated outcomes show that the proposed method outperformed the Decision Tree and general SVM based solution concerning the performance metrics. The proposed algorithm provided an average detection accuracy of 98.32% with false positives and false negative rates of 1.49% and 0.19% respectively when a particular 3-feature subset was considered (Fig. 3). The receiver operating characteristics (ROC) shows a better view of the outcome regarding the AUC value of 0.98384, which is closer to one, an F1-score of 0.9905. This proposed scheme provides a reasonable solution for the feature extraction combinatorial optimization problem and does not require any specific hardware configuration.

Future works can be extended to enhance the performance of the algorithm and evaluate the effectiveness of the proposed scheme. The following are a few proposed extensions of future work:

a) In the proposed scheme Simulated Annealing is used for finding the best feature combination without trying all possible feature combination and SVM is used as the cost function. In this case, instead of Simulated Annealing, Genetic algorithms can be introduced to evaluate the performance of the proposed scheme. Also, Artificial Neural Networks can also be considered to differentiate the performances so we can be more confident about the outcomes that which algorithm performs better.

b) In this research, one of the feature columns was discarded

named attack category, which specifies which type of intrusion it is. In the UNSW and UNB dataset, all the intrusions are labelled as one (regardless of their category), and regular traffic is labelled as logical 0. The proposed scheme can detect the intrusion with improved precision compared to other detection methods discussed but is unable to determine what type of intrusion it is. So, future works can include considering the attack categories so the algorithm will learn more and provide a particular set of significant features to detect a particular type of network attack.

c) To pre-process the data samples, the linear feature scaling process were considered. However, future work can introduce standardization method to preprocess the data samples and analyze the performances to differentiate how good/bad the algorithm performs if different scaling process is used.

Acquiring latest, real-time and effective network intrusion datasets is difficult. The proposed method is only tested on Network intrusion detection datasets from University of New South Wales, Australia and University of New Brunswick, Canada as these two institutions have generated the most recent intrusion detection upgraded datasets, which are being considered as one of the richest datasets. This new detection scheme may be used in other sectors like finance & economy, portfolio management, health analysis and fraud detection. Therefore, as future work, this proposed method can be applied in different sectors and evaluate the performances to establish that this method is a generalized method and may be used towards any dataset to provide a reasonable solution for combinatorial optimization problem compared to other machine learning methods.

Conflict of Interest

The authors declare no conflict of interest.

References

[1] "Challenges of Big Data analysis," *National Science Review,* vol. 1, no. 2, p. 293–314, June 2014.

[2] Alexandra L'Heureux, Katarina Grolinger, Hany F. Elyamany and Miriam A. M. Capretz, "Machine Learning with Big Data: Challenges and Approaches," *" IEEE Access Journal,* vol. 5, pp. 7776 - 7797, April 2017.

[3] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu and F. Kojim, "Big Data Analytics, Machine Learning, and Artificial Intelligence in Next-Generation Wireless Networks," *IEEE Access Journal,* May 2018.

[4] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proc. Natl. Acad. Sci. USA,* vol. 81, 1984.

[5] F. S. Girish Chandrashekar, "A survey on feature selection methods," *Computers & Electrical Engineering,* vol. 40, no. 1, pp. 16-28, January 2014.

[6] P. M. M. A. Pardalos, Open Problems in Optimization and Data Analysis, 2018.

[7] Miloš Madi, Marko Kovačević and Miroslav Radovanović, "Application of Exhaustive Iterative Search Algorithm for Solving Machining Optimization Problems," *Nonconventional Technologies Review Romania,* September 2014.

[8] Cui Zhang , Qiang Li and Ronglong Wang , "An intelligent algorithm based on neural network for combinatorial optimization problems," in *2014 7th International Conference on Biomedical Engineering and Informatics*, Oct. 2014.

[9] M. Chakraborty and U. Chakraborty, "Applying genetic algorithm and simulated annealing to a combinatorial optimization problem," in *Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications*, 1997.

[10] X. Liu, B. Zhang and F. Du, "Integrating Relative Coordinates with Simulated Annealing to Solve a Traveling Salesman Problem," in *2014 Seventh International Joint Conference on Computational Sciences and Optimization*, July 2014.

[11] M. Gao and J. Tian, "Network Intrusion Detection Method Based on Improved Simulated Annealing Neural Network," in *2009 International Conference on Measuring Technology and Mechatronics Automation*, April 2009.

[12] R. D. Brandt, Y. Wang, , A. J. Laub and S. K. Mitra, "Alternative networks for solving the traveling salesman problem and the list-matching problem," in *Proceedings of IEEE International Conference on Neural Networks*, 1998.

[13] SUN Jian, YANG Xiao-guang, and LIU Hao-de, "Study on Microscopic Traffic Simulation Model Systematic Parameter Calibration," *Journal of System Simulation,* 2007.

[14] R. M. V. H. R. Z. W. P. M. Azencott, "Automatic clustering in large sets of time series," in *Partial Differential Equations and Applications*, Springer, Cham, 2019, pp. 65-75.

[15] Ping-Feng Pai and Wei-Chiang Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting," *Energy Conversion & Management,* vol. 46, no. 17, Oct 2005.

[16] J. S. Sartakhti, Homayun Afrabandpey and Mohammad Saraee, "Simulated annealing least squares twin support vector machine (SA-LSTSVM) for pattern classification," *Soft Computing, Springer Verlag,* 2017.

[17] P. J. Herrera, Gonzalo Pajares, María Guijarro and José Jaime Ruz, "Combining Support Vector Machines and simulated annealing for stereovision matching with fish eye lenses in forest environments," *Expert Systems with Applications, Elsevier,* July 2011.

[18] K. Murugan and Dr. P.Suresh,, "Optimized Simulated Annealing Svm Classifier For Anomaly Intrusion Detection In Wireless Adhoc Network," *AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES,* vol. 11, no. 4, pp. 1-13, March 2017.

[19] Moustafa Nour and Jill Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network dataset)," in *Military Communications and Information Systems Conference (MilCIS), IEEE*, 2015.

[20] Moustafa Nour and J. Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective,* 2016.

[21] Iman Sharafaldin,, Arash Habibi Lashkari, and and Ali A, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," in *4th International Conference on Information Systems Security and Privacy (ICISSP)*, Jan 2018.

[22] Julia Neumann, Christoph Schnorr And Gabriele Steidl, "Combined SVM-Based Feature Selection," in *Springer Science & Business Media, Inc.*, 2005.

[23] Leonel Santos, Carlos Rabadan and Ramiro Gonçalves , "Intrusion detection systems in the Internet of Things: A literature review," in *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, June 2018.

[24] A. Le, J. Loo, Y. Luo and A. Lasebae, "Specification-based IDS for securing RPL from topology attacks," *Molecular Diversity Preservation International,* May 2016.

[25] C. Liu, J. Yang, Y. Zhang, R. Chen and J. Zeng, "Research on immunity-based intrusion detection technology for the Internet of Things," in *Natural Computation (ICNC), 2011 Proceedings of the Seventh International Conference*, 2011.

[26] E. Cho, J. Kim and C. Hong, "Attack model and detection scheme for botnet on 6LoWPAN," *In Management Enabling the Future Internet for Changing Business and New Computing Services, Springer,* pp. 515-518, 2009.

[27] Doohwan Oh, Deokho Kim and Won Woo Ro, "A Malicious Pattern Detection Engine for Embedded Security Systems in the Internet of

Things," *Senors (basel), PMC Journals,* p. 24188–24211, Dec 2014.

[28] F. Kuang, W. Xu and S. Zhang, "A novel hybrid KPCA and SVM with GA model for intrusion detection," *Applied Soft Computing, ACM,* vol. 18, no. c, pp. 178-184 , May 2014 .

[29] J. Zhang, M. Zulkernine and A. Haque, "Random-forests-based network intrusion detection systems," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 38, pp. 649-659, 2008.

[30] S.S. Sindhu, S.G. Sivatha and A. Kannan, "Decision tree based lightweight intrusion detection using a wrapper approach," *Expert Systems with Applications, Elsevier,* vol. 39, pp. 129-141, 2012.

[31] T. Lee, C. Wen, L. Chang, H. Chiang and M. Hsieh, "A lightweight intrusion detection scheme based on energy analysis in 6LowPAN," *In Advanced Technologies, Embedded and Multimedia for -centric Computing, Springer,* pp. 1205-1213, 2014.

[32] GuoruiLi, Jingsha He and Yingfang Fu, "Group-based intrusion detection system in wireless sensor networks," *Computer Communications, Elsevier,* vol. 31, no. 18, pp. 4324-4332, Dec 2008.

[33] Julio Barbancho, Carlos León, F.J.Molina and Antonio Barbancho, "Using artificial intelligence in routing schemes for wireless networks," *Computer Communications, Elsevier,* vol. 30, no. 14-15, pp. 2802-2811, Oct 2007.

[34] Edith Ngai, Jiangchuan Liu and Michael R. Lyu, "An efficient intruder detection algorithm against sinkhole attacks in wireless sensor networks," *Computer Communications, Elsevier,* vol. 30, pp. 2353-2364, Sep 2007.

[35] J. C. Platt, "Fast training of support vector machines using sequential minimal optimization," *Advances in kernel methods, ACM Digital Library,* pp. 185-208 , 1999.

[36] Y. LeCun, L. D.Jackel, L. Bottou, A. Brunot, C. Cortes, J.S.Denker, H.Drucker, I. Guyon, U. A. Müller, E. Säckinger, P. Simard and V. Vapnik, "Comparison of learning algorithms for handwritten digit recognition," in *International Conf. Artif. Neural Networks*, 1995.

[37] E. Osuna, R. Freund and F. Girosit, "Training support vector machines: an application to face detection," in *IEEE Computer Society Conference on Computing, Viison and Pattern Recognition*.

[38] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna and T. Poggio, "Pedestrian detection using wavelet templates," in *Computer Vision & Pattern Recognition, IEEE Computer Society*.

[39] J. R. Quinlan, "Learning decision tree classifiers," *ACM Computing Surveys (CSUR),* vol. 28, no. 1, pp. 71-72, Mar. 1996.

[40] Prashant Gupta, "Towards Data Science (Sharing concepts)," [Online]. Available: https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052.

[41] Lior Rokach and Oded Maimon, "Top Down Induction of Decision Tree Classifier-A Survey," *IEEE Transaction on System, Man and Cybernetics Part ,* vol. 1, 2002.

[42] J. R. Quinlan, "Induction of decision trees," *Machine Learning, Springer,* vol. 1, no. 1, p. 81–106, 1986.

[43] J. R. Quinlan, "C4.5: Programs for Machine Learning," in *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, 1993.

[44] Leo Breiman, Jerome Friedman, Charles J. Stone and R.A. Olshen , Classification and Regression Trees, Taylor & Francis Group LLC, 1984.

[45] M. M. Homayounpour, M. H. Moattar and B. Bakhtiyari, "Farsi Text Normalization using Classification and Regression Trees and Multilayer Perceptrons," in *International Conference on Computer & Communication Engineering* , 2006.

[46] Tuncay Soylu, Oguzhan Erdem, Aydin Carus and Edip S. Guner, "Simple CART based real-time traffic classification engine on FPGAs," in *in proceedings of 2017 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, Cancun, Mexico, Dec. 2017.

[47] "Decison Tree Leaning, Wikipedia," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Decision_tree_learning.

[48] Asry Faidhul, Ashaari Pinem and Erwin Budi Setiawan, "Implementation of classification and regression Tree (CART) and fuzzy logic algorithm for intrusion detection system," in *2015 3rd International Conference on Information and Communication Technology (ICoICT)*, Nusa Dua, Bali, May 2015.

[49] Sebastian, "sebastianraschka," 2013. [Online]. Available: https://sebastianraschka.com/faq/docs/decision-tree-binary.html. [Accessed 1 12 2018].

[50] S. Soheily-Khah, P.-F. Marteau and N. Béchet, "Intrusion Detection in Network Systems Through Hybrid Supervised and Unsupervised Machine Learning Process: A Case Study on the ISCX Dataset," in *2018 1st International Conference on Data Intelligence and Security (ICDIS)*, April 2018.

[51] Md Nasimuzzaman Chowdhury and Ken Ferens , "A Computational Approach for Detecting Intrusion in Communication Network Using Machine Learning," in *International Conference on Advances on Applied Cognitive Computing ACC'17*, Las Vegas, Nevada, USA, 2017.

[52] M. N. Chowdhury and K. Ferens, "Network Intrusion Detection Using Machine Learning," in *International Conference on Security & Management, SAM'16*, Las Vegas, USA, 2016.

[53] S. Russenschuck, "Application of Lagrange multiplier estimation to the design optimization of permanent magnet synchronous machines," *IEEE Transactions on Magnetics ,* vol. 28, no. 2, pp. 1525 - 1528, Mar. 1992.

ASTES

Advances in Science, Technology and Engineering Systems Journal Vol. 4, No. 3, 278-286 (2019)

www.astesj.com

Special Issue on Advancement in Engineering and Computer Science

ASTESJ

ISSN: 2415-6698

Factors Influencing the Integration of Freight Distribution Networks in the Indonesian Archipelago: A Structural Equation Modeling Approach

Windra Priatna Humang, Sigit Pranowo Hadiwardoyo*, Nahry

Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Depok 16242, Indonesia

A R T I C L E I N F O

Article history:
Received:15 March, 2019
Accepted:13 June, 2019
Online: 30 June, 2019

Keywords:
Network integration
SEM
Archipelago countries
Freight distribution
Shipping

A B S T R A C T

The main problem of the distribution of freight in archipelago countries such as Indonesia is how to ensure that the outlying and outermost islands are served optimally, with low freight costs and optimal frequency of vessel stops at ports. There are three types of vessels that are subsidized and have the duty of public service obligation (PSO) from the government, namely Sea Tollway, Pelni, and Pioneer shipping. Each vessel has a different route and is not mutually integrated so that its services are not optimal. Integration of the distribution systems of the three types of vessels is needed, thus the distribution costs and the round voyages of vessels are expected to be more optimal and services can be more competitive. The high cost of freight distribution must be minimized so that the government's burden on PSO subsidies can be reduced. This study aims to determine the parameters of the variables that influence the development of an integrated sea transport network model for freight distribution in a region consisting of many islands. The method used was Structural Equation Modeling (SEM) and the variables used were time, cost, freight, port characteristics, vessel characteristics, government policies, and environmental factors (waves and weather). The data used were the results of a questionnaire from 238 respondents, a sample consisting of regulators, shipping operators, academics, and distributors. The results of the conformity testing of the SEM model indicate that all variables have a significant effect on the process of integration of sea transportation networks in the region of the archipelago.

1. Introduction

The distribution of freight in the Indonesian archipelago is constrainted by the big number of small islands (i.e., 2.342 islands, which makes the 12.38% of the total number of islands in Indonesia) that must be served. This condition causes the high cost of transportation and the low frequency of vessel stops at ports. In this case, the role of the government that provides subsidies and public service obligation (PSO) becomes important. As a continental region, population and economic conditions influence the amount of freight demand in the archipelago region. The types of vessels used to service these islands vary depending on the type and size of the freight to be transported [1]. In Indonesia, there are types of vessels that receive subsidies from the government (i.e. Sea Tollway, Pelni, and Pioneer shipping), and vessels that are not subsidized (i.e. Pelra and private shipping). The difference between a subsidized and non-subsidized vessel is that a subsidized vessel has a shipping route regulated by the government, specifically to serve remote, outermost and border areas, whereas non-subsidized vessels are privately owned vessels and the shipping routes are not regulated by the government.

Sea Tollway is a vessel that navigates certain routes regularly and is scheduled from west to east of Indonesia. The Sea Tollway route is served by vessels sized 3000 - 3650 DWT with an estimated capacity of 115 Teus or 2,600 tonnes. Meanwhile, the Pelni vessel is basically a vessel to carry passengers. However, along with the reduced number of passengers, there are several Pelni vessels that are modified so that they can be used to transport passengers, freight and vehicles. Pelni's freight capacity reaches 3,084 passengers, 500 tonnes and 98 Teus, whereas the Pioneer vessel is devoted to serving areas that are difficult to access by large-sized vessels because the area is on small islands, a border area with other countries and has a limited port infrastructure [2]. The size of the Pioneer vessels ranges from 500 - 1,000 DWT/1,200 - 2,000 GT, and is capable of carrying freight up to 1,000 tonnes.

*Sigit Pranowo Hadiwardoyo, Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, sigit@eng.ui.ac.id

Previous research on the sea transportation services of Pioneer in Indonesia showed that this type of service had not been effective from aspects of continuity, safety, vessel capacity, frequency, transport costs and integration [2]. Therefore, it is necessary to enhance a sea transportation network system that integrates a network of vessels that are subsidized: Sea Tollway, Pelni, and Pioneer. Integration of the network is expected to reduce the cost of freight distribution so that the government's burden on PSO subsidies can be reduced. Besides the distribution costs that need to be minimized, this network integration also needs to pay attention to the frequency of round voyage of vessels and schedule arrangements, so that services are more competitive and continuity of freight is maintained.

The process of network optimization is strongly influenced by vessel cost efficiency, energy efficiency (fuel) and the level of safety during the sailing process [3]. Other variables that influence the optimization of sea transportation networks are shipping frequency, sailing time and vessel capacity [4]. The length of time at the port (which consists of loading and unloading time and berthing time) also affects the integration of service networks. Other variables that determine network integration are distance, number of transit ports, vessel size, port facilities, loading and unloading equipment, and economic policy [5,6].

In addition to transportation costs and costs at the port, inventory costs also affect distribution costs [7]. Inventory costs are divided into two, namely marine inventory costs and inventory costs at the port. Inventory costs are positively correlated with freight demand, freight value, and length of sailing time or length of storage time at the port [6]. In Indonesia, the length of the sailing boat and the length of dwelling time at the port has caused inventory costs to increase, which is one of the causes of high selling prices. Several studies have sought to develop an information system to detect the position of the freight so that the owner can reduce inventory costs due to the length of sailing time on the sea [8].

In addition to these variables, there is a factor that is uncertain which can affect the cost and time of delivery, namely repositioning empty containers [9,10]. The imbalance of the position of empty containers between one port and another port will result in significant costs. In Indonesia, environmental constraints such as weather conditions and high waves often affect shipping routes, especially in the sea transport networks of Pioneer. Research by Walter et al [3] explains that weather conditions and high waves are uncertain variables and can be modeled as nonlinear optimization problems or discrete optimization problems with time functions.

In another paper, it is explained that the system of sea transportation networks during certain weather conditions can be modeled by looking at the behavior of the vessel. Behaviors that can be assessed include vessel motion, the speed that can be achieved, fuel consumption and emissions [11].

In the distribution system of PSO subsidies in Indonesia, freight demand is fluctuative and the amount is difficult to ascertain. This condition often causes the carrying of only a small amount of freight by the vessel, which results in increased distribution costs. According to Sumaleeet al. [12], stochastic optimization of routes should pay attention to weather/wave conditions and freights because those are the factors that determine the success of the model.

This research is part of the main research that aims to build an optimization model for the integration of sea transportation networks for freight distribution in the Indonesian archipelago. The objective function to be used is profit maximization and optimization of stop frequency at the destination port. In the first part of this research, we will determine the variables that influence the development of network integration models and the magnitude of the influence (both certain and uncertain), then it will be input into the network integration model that will be created.

2. Methodology

2.1. Conceptual Model

The structural equation modeling (SEM) method is a multivariate statistical technique that combines factor analysis, path analysis, and regression analysis. SEM has been applied to transportation research related to the environment [13], bus network parameters [14], perceptions of pedestrian services [15], accessibility and modal connectivity [16], and train services [17]. An examination is conducted to determine the variables that influence the modeling of the integration of sea transportation networks for freight distribution. The initial hypothesis is that distribution network integration in the archipelago region can be efficient if the network settings also contain uncertain variables, such as freight demand and weather/wave conditions, in addition to time, cost, freight, port characteristics, vessel characteristics, and government policies. This hypothesis will be tested and the influencing variables and the magnitude of the effect will be determined.Some software used for SEM modeling includes LISREL, STATA, and EQS. This study used AMOS software.

To test the effect of each variable (table 1) on network integration, the SEM model was tested with the following steps [18]:

a) Develop structural equation models based on causality and theory-based connection, where a variable change is assumed would create changes in other variables;

b) Arrange causality with path diagrams

c) Arranging structural equations. To compile a conceptual model of SEM, there are two steps that must be performed, namely analyzing the measurement model and analyzing the structural model through the following equation [14]:

$$\text{Structural equations} \rightarrow \eta = \beta\eta + \Gamma X + \zeta \quad (1)$$

$$\text{Measurement equations} \rightarrow Y = \Lambda y + e \quad (2)$$

where: $\boldsymbol{\eta}$ is latent dependent parameters; $\boldsymbol{\beta}$ is coefficients of the η parameters; $\boldsymbol{\Gamma}$is coefficients of the X variables in the structural relationship; $\mathbf{X}$ is exogenous observed parameters; and $\boldsymbol{\zeta}$ is errors in the structural relationship between η and X; $\mathbf{Y}$ is dependent parameters; $\boldsymbol{\Lambda y}$ is coefficients of y on η; $\mathbf{e}$ is errors in the structural relationship Y.

The structural model illustrates the relationship between constructs that have a causal relationship. The structural model consists of independent and dependent variables. Unlike the measurement model that does not recognize dependent

variables and independent variables, the structural model assumes all construct variables as independent variables.

d) Testing goodness of fit (GOF). The structural model results will be tested with the standard criteria for goodness of fit (GOF) [18, 19], namely;

- The good Sig-Probability is ($p \geq 0.05$), indicates that the zero hypothesis is accepted and the predicted matrix input is not different from the statistical method.
- RMSEA (Root Mean Square Error of Approximation) measure the deviation of the parameter values of a model with the population covariance matrix. The RMSEA value is <0.05, means the close fit model, while the value of 0.05 <RMSEA <0.08 shows the good fit model.
- GFI(Goodness of Fit Index) is a measure of the accuracy of the model in producing observed covariance matrix. This GFI value must range from 0 - 1. If the value closer to the number 1, the model will be more better.
- AGFI (Adjusted Goodness Of Fit Index) is same as GFI, but IT adjusts the effect of the degree of freedom on the model. The accepted AGFI measure is ≥ 0.90
- CMIN/DF (The Minimum Sample Discrepancy Function) isa chi-square value divided by the degree of freedom. Value ratio < 2 means a good measure.
- TLI (Turker Lewis Index) is an incremental index that compares a model tested to baseline model, where the recommended value is ≥ 0.95 and a value close to 1 indicates a very good fit.
- CFI (Goodness of Fit Index) is a non-statistical measure which value ranges from 0 to 1. The higher value indicates better fit. GFI value ≥ 0.95 indicates that the tested model has good compatibility.

2.2. Data Processing and Verification

The number of SEM samples is usually greater than 200 [20]. Two hundred and fifty respondents were included to this study, consisting of regulators (transportation ministries and trade ministries, port organizing units), shipping operators (Pelni, Sea Tollway and Pioneer), academics, distributors and traders. Data retrieval was performed by distributing questionnaires directly and Google Docs (online). Respondents were given a statement stating that certain variables had an effect on network integration for freight distribution services in the archipelago region. A five level Likert scale was used to represent the respondents' answers, namely strongly agree (5), agree (4), quite agree (3), disagree (2) and disagree (1).

The results of the questionnaire were verified through validity tests, reliability tests, normalitiy test, outliers test, and multicollinearity. Validity test is used to measure the validity of a questionnaire. A questionnaire is said to be valid if the question in the questionnaire is able to express something that will be measured by the questionnaire. Validation test used the person product moment formula as follows:

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(n\sum x^2 - (\sum x)^2\right)\left(n\sum y^2 - (\sum y)^2\right)}} \quad (3)$$

Where : r = correlation coefficient, ∑X = number of item scores, ∑Y = number of total item scores and n = number of respondents.

Reliability test is used to measure a questionnaire which is an indicator of variables or constructs. A questionnaire is said to be reliable if a person's answer to a statement is consistent or stable over time. Reliability Test used the Cronbach formula

$$\alpha = \left[\frac{k}{k-1}\right]\left[1 - \frac{\sum S_i^2}{S_x^2}\right] \quad (4)$$

where : k is the number of questionnaire; S_i is the single variance; S_X is the whole variance; $\alpha < 0.40$ = low reliability, $0.40 < \alpha < 0.60$ = moderate reliability, $0.60 < \alpha < 0.80$ = high reliability, and $\alpha > 0.80$ = very high reliability.

Normality test with Skewness and Kurtosis is done by comparing the value of Skewness Statistics divided by Error Skewness Standard or Kurtosis Statistical value divided by Error Kurtosis Standard. If the value is between -2 and 2 then the data distribution is normal. Outliers test is done to bring extreme values out. The outliers can affect the test for normality, linearity, and homogeneity of variance and lead to errors in research conclusions from statistical tests results. The multicollinearity test is done by looking at the VIF (variance inflation factor) value below 10.00 and the tolerance value of more than 0.100, then it is concluded that the regression model has no multicollinearity.

The results of the validity test, reliability, normality test, outliers test, multicollinearity in section 3.2 will determine whether the design variables in table 1 can be continued for the manufacture of SEM models.

2.3. Variable Design

Literature search and in-depth interviews with respondents found eight latent variables that influence the integration of sea transportation networks for distribution in the archipelago region. Each latent variable has a manifest variable, as shown in Table 1.

Table 1. The matrix of Operational Variables of SEM Concept

No	Latent Variables	References	Notation	Manifest Variable	Variable Type
1	Port	[7, 21, 22, 23]	P1	Port facilities	certain
			P2	Loading and unloading equipment	
			P3	Pool depth of port	
			P4	Distance between ports	
2	Vessel	[1, 4, 7, 22-22]	S1	Number of vessels	certain
			S2	Speed of vessels	
			S3	Capacity of vessels	
3	Government's policies		K1	Provision of subsidies	uncertain

No	Latent Variables	References	Notation	Manifest Variable	Variable Type
		Interview stakeholder (2018)	K2	Shipping Instruction	
			K3	Delivery order system	
4	Time	[1, 4, 6, 8, 21, 22- 24]	T1	Sailing time	certain
			T2	Loading and unloading time	
			T3	Berthing time	
			T4	Anchoring time	
			T5	Docking time	
5	Cost	[1, 6- 8, 21, 22, 24- 26]	C1	Sailing costs	certain
			C2	Loading and unloading costs	
			C3	Inventory costs	
			C4	Storage costs	
			C5	Container yard costs	
			C6	Terminal Handling Charge	
6	Freight	[5, 8, 9, 10, 21, 22, 23, 27, 28]	M1	Freight unloaded	uncertain
			M2	Freight loaded	
7	Environment	[3, 11, 12]	L1	Wave	uncertain
			L2	Weather	
8	Network integration	[6]	IJ1	Time at the port	certain
			IJ2	Distance between ports	
			IJ2	Freight/container transport costs	

2.4. Hypothesis

Based on the relationship between variables, several hypotheses were formed, as follows:

- H_1 : costs at sea and at the ports affect the integration of freight transportation networks in the archipelago region
- H_2 : time at sea and at the port affects the integration of freight transportation networks in the archipelago region
- H_3 : environmental conditions (waves and weather) affect the integration of freight transportation networks in the archipelago region
- H_4 : the condition of facilities, loading and unloading equipment, the number of ports and the distance between ports affect the integration of freight transportation networks in the archipelago region
- H_5 : government policies affect the integration of freight transportation networks in the archipelago region
- H_6 : vessel characteristics affect the integration of freight transportation networks in the archipelago region
- H_7 : the number of freight items affects the integration of freight transportation networks in the archipelago region

3. Result and Discussion

3.1. Characteristics of Respondents

Only 238 of the total 250 questionnaires were valid to be continued to the data analysis stage. Characteristics of respondents based on gender, age, education and work experience. This can give an idea of the respondent's ability to provide subjective justification and opinion to problem. The characteristics of the respondents from 238 samples can be seen in Table 2.

Table 2. Characteristics of Respondents (N=238)

Characteristics		Total	Percentage
Gender	Male	155	65%
	Female	83	35%
Age (year)	18 – 25	29	12%
	26 – 35	57	24%
	36 – 45	90	38%
	46 – 55	50	21%
	> 55	12	5%
Education	High school	14	6%
	Diploma	43	18%
	Bachelor	143	60%
	Master	33	14%
	Doctoral	5	2%
Length of work (year)	<5	14	6%
	5 – 9	79	33%
	10 – 14	95	40%
	15 – 19	38	16%
	≥ 20	12	5%

The characteristics of the respondents showed that stakeholders involved in goods distribution activities with sea transportation, both regulators, operators shipping, distributors and academics, were dominated by 65% male respondents. From education, 94% of the respondents are graduate from higher education (diploma, bachelor, master and doctoral degree). From the aspect of work experience, 61% of respondents have worked in the field of sea transportation for > 10 years. This indicates that the respondent's experience is enough to provide justification for freight distribution activities by sea transportation and is able to explain the influence of each variable on distribution network integration.

3.2. Statistic Test

Statistical tests were conducted to prove that questionnaires and variables were feasible. The results of the validation test showed that 28 variables were valid, while the reliability test results showed a value of $\alpha = 0.69$ (> 0.60), which indicated that the questionnaire was reliable as a measuring instrument. The results of the normality test with the criteria of critical ratio skewness value were ± 1.98 at the significance level of 0.01. The data have a normal distribution because the critical ratio skewness and kurtosis ratio were between the values of -2 to +2 [18]. The outliers test results showed the value of the Mahalanobis distance. The chi-square value in the degree of freedom 28 (i.e. the number of manifest variables) was at a significance level of $p < 0.001$. Observations using Mahalanobis values that were greater than Chi-square tables could be called as outlier observations. In this study, there were no outlier observations. Furthermore, the multicollinearity test shows that the variance inflation factor (VIF) value is 1.890 (<10) and tolerance is more than 0.520 (> 0.100), so it can be concluded that there is no multicollinearity problem between the independent variables for the data to be analyzed.

3.3. SEM Model

Before making a full structural model, a confirmatory unidimensionality design with confirmatory factor analysis (CFA) is conducted. CFA is used to test whether the manifest variable in a construct is valid and correct in forming a construct so that the construct becomes homogeneous or unidimensional. The requirements of the CFA measurement use the value of the convergent validity factor (CVF). The criterion for good convergent validity is ≥ 0.7 [18], although convergent validity of 0.5-0.6 is still acceptable. If the manifest variable value is lower than 0.5, then the variable is considered to be dimensionless, similar to the other manifest variables in explaining a latent variable. CFA was performed on each latent variable in the model. In this study, there were eight latent variables, seven of which were exogenous constructs, namely ports, vessels, government policies, time, costs, freight demand and environment and one endogenous construct, namely network integration. The CFA test results of latent variables on manifest variables that have a significant effect are shown in Table 3 below.

Table 3. The CFA significance test value of the variables against exogenous constructs

Manifest Variable		Latent Variable	Estimation CVF	Test Results
P1	←	Port	0.748	valid
P2	←	Port	0.913	valid
P3	←	Port	0.666	valid
P4	←	Port	0.516	valid
S1	←	Vessel	0.809	valid
S2	←	Vessel	0.746	valid
S3	←	Vessel	0.726	valid
K1	←	Policy	0.574	valid
K2	←	Policy	0.560	valid
K3	←	Policy	0.532	valid
T1	←	Time	0.981	valid
T2	←	Time	1.054	valid
T3	←	Time	0.558	valid
T4	←	Time	0.872	valid
T5	←	Time	0.788	valid
C1	←	Cost	0.673	valid
C2	←	Cost	0.995	valid
C3	←	Cost	0.700	valid
C4	←	Cost	0.650	valid
C5	←	Cost	0.517	valid
C6	←	Cost	0.527	valid
M1	←	Freight	0.545	valid
M2	←	Freight	0.824	valid
L1	←	Environment	0.972	valid
L2	←	Environment	0.654	valid
IJ1	←	Network Integration	0.601	valid
IJ2	←	Network Integration	0.780	valid
IJ3	←	Network Integration	0.644	valid

The results of the CVF analysis in Table 3 showed that all latent variables (i.e. ports, vessels, government policies, time, cost, freight, environment and network integration) have a significant effect on the manifest variable. After performing the CFA analysis, the next step performed was to estimate the full structural model that only includes manifest variables (all variables have been tested using CFA). Full structural models wouldprovide relationships between constructs that have been determined in SEM. To determine whether a model was good or not, a goodness of fit (GOF) test was carried out with the standard criteria shown in Table 4 (standard cut-off value).

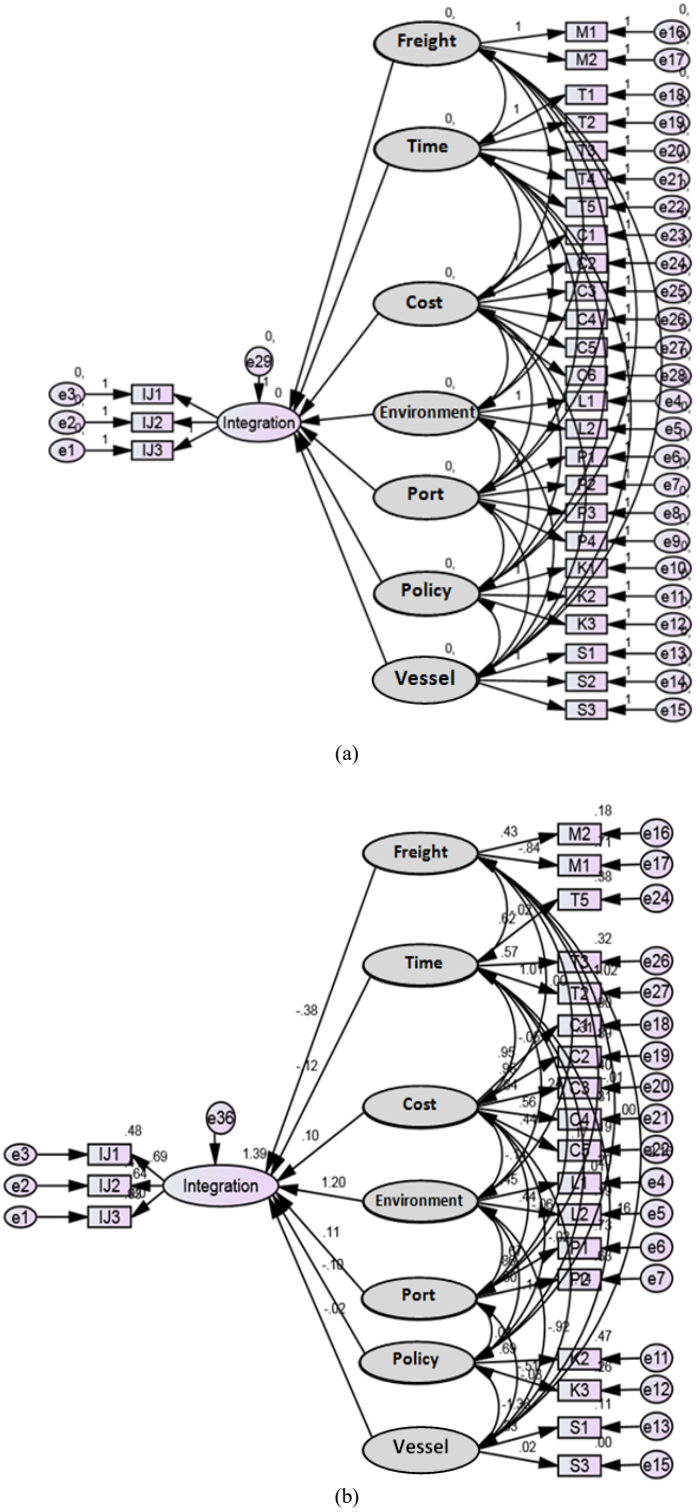


Figure 1. The initial hypothesis of the SEM structural model (a), the estimation results of the structural model (b)

The initial hypothesis of the structural model was that all variables (i.e. freight, time, cost, environment, port, policy, and vessel) affect network integration (Figure 1a). Then, the initial model was tested using GOF criteria. The results of the initial test

of the structural model with AMOS 23.0 showed that the negative variables for network integration were freights of -0.38, time of -0.12, policies of -0.10 and vessels of -0.02, while the positive variables were the cost of 0.10, the environment of 1.20, and port of 0.11. The GOF test showed the sig-probability = 0.003 (not fit), RMSEA = 0.108 (not fit), GFI = 0.082 (not fit), AGFI = 0.455 (not fit), CMIN/DF = 1.813 (good fit), TLI = 0842 (marginal fit) and CFI = 0.879 (marginal fit). The results of the initial estimation of the structural model are shown in Figure 1b.

The results of the test for goodness-of-fit (GOF) criteria for the overall results of the initial estimation of the structural model are shown in Figure 1b and indicate that the model was not good because the only GOF criterion that fulfills the requirements is CMIN/DF (≤ 2.00). Therefore, it is necessary to modify the model. The process of modifying a model is basically similar to repeating the processof testing and estimating the model. The purpose of the modification is to see whether the modifications can reduce the chi-square value so that the GOF standard can be fulfilled. Modifications are performed in several ways, such as changing paths between latent variables to other latent variables, connecting latent variables to other latent variables, or connecting the manifest variable to other manifest variables, until the best model is obtained.

Modification of the structural model has been performed several times to obtain a model that has the best criteria for goodness-of-fit. From several experiments, we found a model that had the best criteria for goodness-of-fit, as shown in Figure 2. Modification of the SEM structural model (Figure 2) was carried out by changing the variable of the vessel path because it affected both network integration and freight, While the charge variable affected the variable of time. The results of the modification of the previous structural model assumed that the seven latent variables (i.e. freight, time, cost, environment, port, policy, and vessel) have a direct effect on network integration; however, changes occurred based on the modification results. Latent variables that affect directly the network integration were time, cost, environment, port, policy, and vessel, while the freight variable did not affect directly.

This modification of the structural model caused the elimination of two manifest variables, namely subsidy (K1) and terminal handling charge (C6). The provision of subsidies for Sea Tollway, Pelni, and Pioneer shipping was assumed to have no significant effect on network integration, because the function of subsidizing was to reduce transportation costs rather than to improve service networks directly. The terminal handling charge (THC) in the modification of the structural model was eliminated because of THC only being applied to the main port and import freights.

The results of the modification of the structural model in Figure 2 show that there was a change in the value of the latent variable which previously had a negative change to positive for network integration for the distribution of goods. Cost variable is 0.08, environment variable is 0.96, port variable is 0.24 and policy is 0.04, vessel is 0.08 and time is 0.02.

The modified structural model showed that the goodness-of-fit criteria fulfilled the requirements, where the Sig-Probability value = 0.061 (good fit), RMSEA = 0.077 (good fit), GFI = 1.034 (good fit), AGFI = 0.984 (good fit), CMIN/DF = 1,920 (good fit), TLI = 0.901 (marginal fit), and CFI = 0.882 (marginal fit). The comparison of results from running the initial model and modified models can be seen in Table 4.

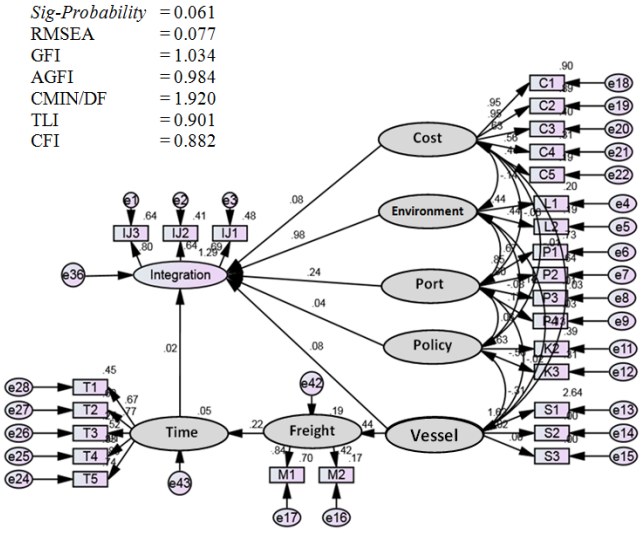


Figure 2. The Best Modification Results of the Full Structural Model

Table 4. Test results on the initial structural model and modified structural model

Test Criteria	Standard cut-off value of test criteria	GOF test results		Remarks (Modified results)
		Initial model	Modified model	
Sig-Probability	≥ 0.05	0.003	0.061	Good fit
RMSEA	≤ 0.08	0.108	0.077	Good fit
GFI	≥ 0.90	0.082	1.034	Good fit
AGFI	≥ 0.90	0.455	0.984	Good fit
CMIN/DF	≤ 2.00	1.813	1.920	Good fit
TLI	≥ 0.95	0.842	0.901	Marginal fit
CFI	≥ 0.95	0.879	0.882	Marginal fit

Tests on the suitability of the model indicated that the modified model was good, because sig-probability, RMSEA, GFI, AGFI, and CMIN/DF values have fulfilled the predetermined requirements. Although there were variables whose values were below cut off value, such as TLI and CFI (marginal fit), this model could still be accepted because the range of values was still close to the cut-off value. According to Ghozali [18], if two or more of all the test criteria used have shown a good fit, it can be called as a good model.

3.4. SEM Model Interpretation

From the results of model suitability testing, it was concluded that all variables significantly influence the integration of sea transportation networks for freight distribution in the Indonesian archipelago. The value of the test results can be seen in Table 5.

The results of the hypothesis test (Sig-Probability) in Table 5 show that all exogenous variables have an effect on an endogenous variable (network integration), thus, all hypotheses can be accepted. Estimation value shows that the time variable can be explained by sailing time indicators, loading and unloading time, berthing time, anchoring time and docking time of 87.2% while the other 12.8% is explained by indicators not contained in this paper; The cost variable can be explained by indicators of sailing costs,

loading and unloading costs, inventory costs, storage costs, and container yard costs of 71.3%; Port variables can be explained by indicators of port facilities, loading and unloading equipment, pool depth of port, and distance between port of 45.8%; Vessel variables can be explained by indicators of the number of vessel operating, vessel speed and vessel capacity of 58.8%; Environmental variables can be explained by wave height indicators and weather conditions of 41.6%; Government policy variables can be explained by the shipping instruction indicator and the delivery order system of 67.2%; The freight variable does not directly affect network integration, but has a significant effect on the time of loading and unloading at the port. Same with vessel variables, in addition to influencing integration, it also has a significant effect on freight.

Table 5. Suitability Test Results of the Integration Model of the Sea Transportation Network in the Archipelago Region

Variable	Estimation	S.E	C.R	Sig-Probability	Remarks
Freight ←Vessel	0.395	0.070	11.341	0.000	Significant
Time ←Freight	0.530	0.067	12.318	0.000	Significant
Integration ← Time	0.872	0.158	6.147	0.000	Significant
Integration ← Cost	0.713	0.064	14.233	0.000	Significant
Integration ← Environment	0.416	0.043	2.720	0.007	Significant
Integration ← Port	0.458	0.048	9.502	0.011	Significant
Integration ← Policy	0.672	0.054	16.071	0.000	Significant
Integration ←Vessel	0.588	0.060	13.229	0.000	Significant

The results of the SEM model testing on the influence of each variable against the integration of sea transportation networks in the Indonesian archipelago can be interpreted as follows:

- The variable of the vessel has a significant relationship to the freight. The capacity of vessels which are currently operating tends not to be proportional to the amount of the freight carried. The average freight only reaches 40-45%, even lower. One of the reasons is the lack of interest of distributors to use Sea Tollway, Pelni, and Pioneer vessels because of the high sailing time (due to multi-port systems), in contrast to the faster private shipping liners (due to the direct shipping system) [2]. In addition, there is very little freight loaded from the destination port because it still relies on the freight of natural resources and there are no industrial products. The number of vessels operating is also not comparable with the number of islands that must be served;one vessel only serves one route. However, according to the law of the provision of PSO subsidies, the government is obliged to serve all communities fairly. Therefore, it is necessary to optimize the sea transportation service network to ensure freight distribution services in the archipelago region.
- The variable of freight has a significant relationship with time, especially sailing time, loading and unloading time, and berthing time. The sailing time is determined by the distance between the port and the speed of the vessel. In addition, it is also indirectly affected by the amount of freight. The age of the vessel used is mostly >20 years old. With the problem of age and the amount of freight, the speed of the vessel will not be maximized, so the sailing time will be longer. Berthing time at the port is the accumulation of loading and unloading time and berthing time for other activities, such as filling fuel and water. The amount of freight causes a high loading and unloading time due to the limited loading and unloading equipment at the port. Nowadays, the loading and unloading process still relies on vessel cranes, while the general freight loads of Pioneer vessels still use loading and unloading labor.
- The variable of time has a significant relationship to network integration. The main constraints of network integration are time (i.e. sailing time, loading and unloading time, berthing time, anchoring time, and docking time) [4,6,8]. Long sailing time is caused by the distance between the ports that are far away, also due to weather and wave conditions that occur in the western season [29]. Pioneer vessels sized of 500 DWT are often not allowed to sail because of the weather conditions and high waves up to 3-5 meters. The loading and unloading process at the port is considered to be ineffective because of the lack of loading and unloading equipment and still relying on labor. This condition also resulted in long berthing time at the port and high berth occupancy ratio (BOR). In certain periods, the vessel must be docked to guarantee marine reliability within 1-3 months depending on the condition of the vessel. With a limited number of vessels, the time used for the docking process will reduce distribution service performance.
- The variable of cost has a significant relationship to network integration. This is related to the economic level of each route. The chosen route must minimize distribution costs. The distribution network system in Indonesia is considered not efficient [1]. Although the distance is close, the distribution costs are not necessarily cheap, and vice versa. This condition is caused by different port costs [6]. The government, through PSO subsidies, has minimized transport costs (sailing costs), but the costs of loading, warehousing, stacking and inventory costs are still borne by the freight owner. The tendency of distributors is to choose routes that are low in cost, and they tend to ignore the long distribution time.
- Environmental variables (weather and waves) have a significant relationship to network integration. The routes of Pioneer vessels have a risk of weather conditions and high waves. The higher waves on each route, will increase the potential for accident risk [11]. Forty-eight routes of Pioneer vessels exist in eastern Indonesia, most of which are on the region that has high waves (3 to 5 meters), thus it is quite vulnerable in terms of shipping safety, especially in the west wind season.
- The variable of the port has a significant relationship to network integration. Optimizing the operation of port facilities and the productivity of loading and unloading equipment are directly related to service time. Therefore, the operational performance of port facilities must be maximized so that the time used can be minimized [7]. The operation of port facilities is represented by berth occupancy ratio (BOR), shed occupancy ratio (SOR) and yard occupancy ratio (YOR), while the productivity of loading and unloading equipment is represented by the productivity of container gantry cranes, mobile cranes, forklifts, vessel cranes, and labor. The depth of the port pond is related to the ability of the vessels to dock.

Nowadays, several ports have a large demand, but the depth of the pond is not sufficient, such as for a Sea Tollway vessel with a draft of ≥ 5.5 meters, so that it can only be used by Pioneer vessels. Furthermore, network integration will be better if the service distance between ports is not too far away. Current conditions show that the distance between the ports served is classified as a long-distance track (i.e. Sea Tollway and Pelni with an average of 300-1000 sea miles, Pioneer with an average of 50-200 sea miles) thus, clustering is needed to minimize service distance.

- The variable of government policy has a significant relationship to network integration. The policy of shipping instruction and the order delivery system of Sea Tollway and Pelni vessels were still perceived as a burden on the shipper (freight owner) in the process of ordering freight space on the vessels. Nowadays, there is still a monopoly on ordering freight space and the administrative processes require a long time. The government has tried to improve the implementation of an online information system of vessels freight space, but it has not been implemented properly.
- The variable of the vessel has a significant relationship to network integration. The government, through the provision of subsidies and the PSO, commissioned the Sea Tollway, Pelni, and Pioneer vessels to serve small islands, with or without freight, routinely. However, the number of vessels operating is still not proportional to the number of ports that must be served. On the other hand, the vessels with low freight are not profitable to ship operators (i.e. Sea Tollway, Pelni, and Pioneer). If changes are made by operating a small vessel (with small load capacity and higher speed), it will be constrained by high waves.

The magnitude of the influence of each variable on the sea transport network integration illustrates that the implementation of freight distribution services in the Indonesian archipelago must pay attention to these variables. This finding is in accordance with the real conditions, in which freight distribution is often late due to weather conditions and high waves. Pioneer vessels in the western season will be difficult to sail. In addition, policy issues related to port administration and ordering procedures of freight space, and loading systems on vessels also contribute to the long waiting times for freight at the port. Another problem is the small amount of freight and sometimes there is no freight demand at one port point.

4. Conclusions

The integration of the sea transportation network to obtain a better freight distribution system in the Indonesian arcipelago, which includes the network composed of Sea Tollway, Pelni and Pioneer shipping, is directly affected by several variables. These variables can be categorized as certain and uncertain variables. The effects of the certain variables are:

- Time variables have a significant effect on network integration. 87.2% of the time variable can be explained by indicators of sailing time, loading and unloading time, berthing time, anchoring time and docking time;
- Cost variables have a significant effect on network integration. A total of 71.3% of the cost variable can be explained by indicators of sailing costs, loading and unloading costs, inventory costs, storage costs, and container yard costs;
- Port variables have a significant effect on network integration. 45.8% of port variables can be explained by indicators of port facilities, loading and unloading equipment, pool depth of port and distance between ports;
- Vessel variables have a significant effect on network integration. 58.8% of ship variables can be explained by indicators of the number of vessel operating, vessel speed and vessel capacity;

While the uncertain variables that influence network integration are:

- Environmental variables have a significant effect on network integration. 41.6% of the environmental variables can be explained by wave height indicators and weather conditions;
- Government policy variable have a significant effect on network integration. 67.2% of government policy variables can be explained by shipping instruction indicators and delivery order systems;
- Freight variable, which is loading and unloading, basically does not directly affect network integration, but has a significant effect on the time of loading and unloading at the port.

This paper has elaborated on the important aspect of variables, and further research activities can focus on defining new mathematical models to smoothly integrate sea transportation networks. Such model is expected to be able to finally optimize the freight distribution service in Indonesia.

Acknowledgments

Thank you to the Ministry of Research, Technology and Higher Education for the doctoral dissertation research grant, in order to complete doctoral studies at the Universitas Indonesia.

References

[1] Sonny, I., Hadiwardoyo, S. P., Susantono, B., & Benabdelhafid, A.. "The Development of a Freight Distribution Model for Connecting Inter-Island Freight Transport". *International Journal of Technology*, *5*, 743-750. 2015.

[2] Humang, W. P., Hadiwardoyo, S. P. & Nahry. "The Effectiveness of Pilot Marine Transport Services for Goods Distribution in Maluku Island Group, Indonesia". In 2017 6th IEEE International Conference on Advanced Logistics and Transport (ICALT) (pp. 165-169). IEEE. 2017. 10.1109/ICAdLT.2017.8547034

[3] Walther, L., Rizvanolli, A., Wendebourg, M., & Jahn, C. "Modeling and optimization algorithms in ship weather routing." International Journal of e-Navigation and Maritime Economy 4: 31-45. 2016. https://doi.org/10.1016/j.enavi.2016.06.004Arnone,

[4] M., Mancini, S., & Rosa, A. "Formulating a mathematical model for container assignment optimization on an intermodal network". Procedia-Social and Behavioral Sciences, 111, 1063-1072. 2014. https://doi.org/10.1016/j.sbspro.2014.01.141

[5] Hsu, C. I., & Hsieh, Y. P. "Direct versus terminal routing on a maritime hub-and-spoke container network". Journal of Marine Science and Technology, 13(3), 209-217. 2005. doi：10.6119/JMST

[6] Hsu, C. I., & Hsieh, Y. P. "Routing, ship size, and sailing frequency decision-making for a maritime hub-and-spoke container network". Mathematical and Computer Modelling, 45(7-8), 899-916. 2007. https://doi.org/10.1016/j.mcm.2006.08.012

[7] Gelareh, S., & Pisinger, D. “Fleet deployment, network design and hub location of liner shipping companies”. Transportation Research Part E: Logistics and Transportation Review, 47(6), 947-964. 2011. https://doi.org/10.1016/j.tre.2011.03.002

[8] He, J., Huang, Y., & Chang, D. “Simulation-based heuristic method for container supply chain network optimization”. Advanced Engineering Informatics, 29(3), 339-354. 2015. https://doi.org/10.1016/j.aei.2014.08.001

[9] Agarwal, R., & Ergun, Ö. “Ship scheduling and network design for cargo routing in liner shipping”. Transportation Science, 42(2), 175-196. 2008. https://doi.org/10.1287/trsc.1070.0205

[10] Ye, H. Q., Yuan, X. M., & Liu, X. A “Tactical Planning Model for Liner Shipping Companies: Managing Container Flow and Ship Deployment Jointly”. School of Business, National University of Singapore. 2007

[11] Vettor, Roberto, and C. Guedes Soares. "Development of a ship weather routing system." Ocean Engineering 123: 1-14. 2016. https://doi.org/10.1016/j.oceaneng.2016.06.035

[12] Sumalee, A., Uchida, K., & Lam, W. H. “Stochastic multi-modal transport network under demand uncertainties and adverse weather condition”. Transportation Research Part C: Emerging Technologies, 19 (2), 338-350. 2011. https://doi.org/10.1016/j.trc.2010.05.018

[13] Ülengin, F., Kabak, Ö., Önsel, Ş., Ülengin, B., & Aktaş, E. “A problem-structuring model for analyzing transportation–environment relationships”. European Journal of Operational Research, 200(3), 844-859. 2010. https://doi.org/10.1016/j.ejor.2009.01.023

[14] Das, T., Apu, N., Hoque, M. S., Hadiuzzaman, M., & Xu, W. “Parameters Affecting the Overall Performance of Bus Network System at Different Operating Conditions: A Structural Equation Approach”. Transportation research procedia, 25, 5059-5071. 2017. https://doi.org/10.1016/j.trpro.2017.05.206

[15] Zhou, J., Guo, Y., Dong, S., Zhao, L., & Yang, R. "Structural equation modeling for pedestrians’ perception in integrated transport hubs." Procedia engineering 137, 817-826, 2016. https://doi.org/10.1016/j.proeng.2016.01.321

[16] Papaioannou, D., & Martinez, L. M. “The Role of Accessibility and Connectivity in Mode Choice. A Structural Equation Modeling Approach”. Transportation Research Procedia, 10, 831-839. 2015. https://doi.org/10.1016/j.trpro.2015.09.036

[17] Eboli, L., & Mazzulla, G.. “Structural equation modelling for analysing passengers’ perceptions about railway services”. Procedia-Social and Behavioral Sciences, 54, 96-106. 2012. https://doi.org/10.1016/j.sbspro.2012.09.729

[18] Ghozali, I.. “Model persamaan struktural: konsep dan aplikasi dengan program Amos 16.0”. Badan Penerbit Universitas Diponegoro. 2013.

[19] Sarwono, Y.. Pengertian dasar structural equation modeling (SEM). Ilmiah Manajemen Bisnis, 10(3). 2010

[20] Lei, P. W., & Wu, Q. “Introduction to structural equation modeling: Issues and practical considerations”. Educational Measurement: issues and practice, 26(3), 33-43. 2007. https://doi.org/10.1111/j.1745-3992.2007.00099.x

[21] Vasconcelos, Adriano D., Carlos D. Nassi, and Luiz AS Lopes. "The uncapacitated hub location problem in networks under decentralized management." Computers & Operations Research 38.12: 1656-1666. 2011. https://doi.org/10.1016/j.cor.2011.03.004

[22] Gelareh, S., Maculan, N., Mahey, P., & Monemi, R. N. “Hub-and-spoke network design and fleet deployment for string planning of liner shipping”. Applied Mathematical Modelling, 37(5), 3307-3321. 2013. https://doi.org/10.1016/j.apm.2012.07.017

[23] Hao, C., & Yue, Y. “Optimization on combination of transport routes and modes on dynamic programming for a container multimodal transport system”. Procedia engineering, 137, 382-390. 2016. https://doi.org/10.1016/j.proeng.2016.01.272

[24] Nikolaou, Michael. "Optimizing the logistics of compressed natural gas transportation by marine vessels." Journal of Natural Gas Science and Engineering 2.1: 1-20. 2010. https://doi.org/10.1016/j.jngse.2010.02.001

[25] Pedersen, M. B., Crainic, T. G., & Madsen, O. B.. “Models and tabu search metaheuristics for service network design with asset-balance requirements”. Transportation Science, 43(2), 158-177. 2009. https://doi.org/10.1287/trsc.1080.0234

[26] Nahry, “Pengembangan Model Optimasi Sistem Distribusi Komoditsa Untuk Meningkatkan Efisiensi Sistem distribusi BUMN-PSO”, Ph.D Thesis, Universitas Indonesia. 2010.

[27] Meng, Q., Wang, T., & Wang, S. “Short-term liner ship fleet planning with container transshipment and uncertain container shipment demand”. European Journal of Operational Research, 223(1), 96-105. 2012. https://doi.org/10.1016/j.ejor.2012.06.025

[28] Güner, A. R., Murat, A., & Chinnam, R. B. “Dynamic routing for milk-run tours with time windows in stochastic time-dependent networks”. Transportation Research Part E: Logistics and Transportation Review, 97, 251-267. 2017. https://doi.org/10.1016/j.tre.2016.10.014

[29] Rutz, W. O., & Coull, J. R. “Inter-island passenger shipping in Indonesia: development of the system: Present characteristics and future requirements”. Journal of Transport Geography, 4(4), 275-286. 1996. https://doi.org/10.1016/S0966-6923(96)00028-2