Proposal of a congestion control technique in LAN networks using an econometric model ARIMA

Joaquín F Sánchez1, Martha M Cuellar2

1Docente Ingeniería de Telecomunicaciones, Facultad de Ingeniería, San Mateo Educación Superior, Bogotá Colombia
2Docente Negocios Internacionales, Facultad de Negocios Internacionales, San Mateo Educación Superior, Bogotá Colombia

Abstract

The aim of this paper is to propose a methodology to apply to an econometric model which could be the basis for generating a congestion control strategy in a LAN network. Using time series as analysis tool, ARIMA model (Autoregressive Integrated Moving Average) is used in an econometric model, and this same analysis is transferred to the mechanism of congestion control of the TCP protocol. As a result, a transmission rate reduction parameter is generated in a test environment, where congestion control strategy is applied. It is shown that this procedure works and its application could be further expanded in data networks.

1. Introduction

This paper is an extension of work originally presented in conference 2016 8th Euro American Conference on Telematics and Information Systems (EATIS) “Congestion control LAN networks using an econometric model ARIMA”[1]. In Computer Sciences, especially in Telecommunications, several mathematical tools to make analysis and design of the systems allowing to improve their own performance are used. For instance, time series are used on research and industry fields, and reviewing their applications as in [2]-[5], it was found that stock exchange analysis has some similarities with the telecommunications field. After look over the LAN networks traces as viewed in [5] and [6], network trafficking behavior is as similar as the variation graphs of the shares in the stock exchange. Considering this, the present paper aims to reach commonalities between data network trafficking and stock exchange behavior to 1) design a traffic analysis tool in data networks, and 2) propose a model of congestion control based on an econometric model.

Time series are suitable tools to evaluate a given situation of a performance movement of the analyzed variable, and data congestion in the telecommunications networks in this case. This kind of mathematical resources are appropriate to detect immediate trends or errors related to the volume and movement of analyzed data, providing better correctives and trends which are important for data channels management; therefore, time series afford occurrence predictions for decision-making.

Congestion analysis in data networks may carry out using different parameters and trends, getting estimates for the use, control, and application of statistics measures in decision-making. This could guide predictions of the variables in a given model with their historical behavior, statistics, and features to give the chance to value the future occurrence level [8]. In this way, past data may lead to forecast future values movements if data are established and depicted through graphs and specific indicators given by positive or negative variables trends. It is worth mentioning that delay times and LAN networks bandwidth are the variables considered for this paper.

According to the previous concepts, analysis of one share of a given company will be presented. This shall allow to propose the congestion control model of data obtained from a LAN network implementation. In section 2, a theoretical revision of econometric analysis for one share is made and, in section 3, it is discussed why to design a congestion control strategy is worth, explaining some theoretical elements. Proposal of the model and results are presented in sections 4 and 5, and conclusions are displayed in the last section.

2. ARIMA model (autoregressive integrated moving average) to analyze one share

It could be useful to estimate a model capable of making technical analysis related to the behavior of negotiation indicators in the financial market because they might enhance the evaluation of trends, variables, relations of influence, and historical behavior within supply and demand of stock market. In general, the price and the volume of negotiation are the most relevant variables taken...
into account mainly because they determine the movement of the variable: the upward or downward trend, or its resistance to change through considered time opening and closing, like relation of values, maximum and minimum, and volatility in periods of time such as days, weeks, months, and years. In this way, information of global market and its necessities is reflected, determining negotiation trends with patterns of behavior, and allowing that information does exhibit a cyclic and repeated activity according to the own market features.

Looking over time series characteristics in stock markets, basic concepts of one analyzed variable could be established by chartism (analysis of financial graphs) [9]. In fig. 1, the price of one share is depicted. The downward market is the bold outlined circle and the upward market is the light. The bases (dotted lines) delimit the trend in the bearish movement, giving the chance to rebound towards the bullish, meaning that the price demand stands out compared to the supply. On the other hand, in the upward trend of the market (dashed lines), supply is greater than demand.

Fig1: Variation concept [2], [8]

The assessments of prices can progress in a specific time \( t(n) \) to have a suitable measure in a given period of time. Highlighting that a market is named bullish when the price exhibits the trend to be over its own value, the maximum benchmarks come up, (1, 2, 3, 4 in fig. 2) increasingly higher, and minimum benchmarks (1., 2., 3., 4. in fig. 2) are greater than the previous ones. In this way, it is created a trend line which breaks the base of the price trend, giving rise to a new trend, named pull back in the stock market environment. On the other hand, the bearish market exhibits an opposite behavior compared to the bullish market, with the same contextual benchmarks and indicators, associated to the downward trend.

Fig2: Prices trend [2], [8]

ARIMA model can be used to make a Forecasting analysis. When the model is proposed, an estimation of values behavior of the shares is obtained, allowing to make a decision to sell or buy. This model comes from a nonstationary homogenous series.

Given the series \((1 - B)^dZ_t\) which follows a general stationary process \(ARIMA(p, q)\), then:

\[
\phi_p(B)(1 - B)^dZ_t = \theta_q(B)a_t 
\]

(1)

Where the stationary operator \(AR \phi_p(B) = (1 - \phi_1B - \cdots - \phi_pB^p)\) and the operator \(MA \theta_q(B) = (1 - \theta_1B - \cdots - \theta_qB^q)\) share common factors. The \(\theta_0\) parameter is used in different ways when \(d = 0\) and \(d > 0\). When \(d = 0\), the original process is stationary, then \(\theta_0\) is related to the mean of the process. When \(d \geq 1\), the \(\theta_0\) operator is named the term of deterministic trend. The homogenous result of the described model refers to an autoregressive model which integrates a moving average unit for an order \((p, d, q)\) and is denoted \(ARIMA(p, d, q)\).

Fig. 3 depicts an example of ARIMA model functioning. The time series behavior exhibits the variation of a product with the \(ARIMA(2,1,2)\) scheme where:

\[
(1 - \phi_1B - \phi_2B^2)(1 - B)Z_t = (1 + \theta_1B + \theta_2B^2)a_t 
\]

(1)

In other words, \(ARIMA(p, d, q)\) model allows to describe one value as a linear function of previous data and errors which come from the random behavior of that data. It works with a stationary series, and the \(ARIMA(p, d, q)\) model is used to get an estimation or a price forecast, where the time series is put and mean squared error is calculated. To determine the variance of the series we have:

\[
\pi(B)Z_{t+1} = a_{t+b} 
\]

(2)

Where:

\[
Z_{t+1} = \sum_{j=-\infty}^{\infty} \pi_j Z_{t+i-j} + a_{t+i} 
\]

(3)

Parameters must be calculated with those equations to get an estimation of the variance trend of the original series; hence, we have the equations as follows:

\[
e_n(l) = Z_{n+1} - \hat{Z}_n(l) 
\]

(4)

\[
\psi_j = \sum_{l=0}^{l-\infty} \pi_{j-l} \psi_l 
\]

(5)

www.astesj.com
Error is calculated with equation 4, and terms of the estimated series are determined with equation 5.

2.1. Graphical analysis and series analysis

Fig. 4 depicts the result of taking stock exchange behavior and time series into account. It is worth thinking in an ARIMA model which allows to one independent variable (share #1) be related to or compared with other independent variable (share #2) according to its behavior in the stock exchange.

![Graph of one share behavior](image)

Variable Share #1 will take error $e_t(i)$ into account within the constant mean, and its B (risk level) will be reflected, where risk level remains in the trend of its volatility inside of the independent stock market, producing a single one variance. In the same way, behavior of the share #2 will be the same. Relation between these two variables will produce a covariance which shall describe the behavior inside of the stock market, but behavior itself does not have an impact on the independent variable performance within the market.

Then, we have the mean of the independent variable:

$$x = \frac{\sum_{i=1}^{n} x_i}{N} \quad (6)$$

The variance of the independent variable:

$$\sigma^2 = \frac{(x_1-x)^2+(x_2-x)^2+\ldots+(x_n-x)^2}{N} \quad (7)$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-x)^2}{N} \quad (8)$$

Covariance of the two independent variables within the stock market:

$$\sigma_{xy} = \frac{\sum_{i=1}^{n} (x_i-x)(y_i-y)}{N} \quad (9)$$

Beta, risk level:

$$\beta = \frac{\text{COV(Mercado,Asset)}}{\text{VAR(Mercado)}} \quad (10)$$

Having previous equations, ARIMA model is used within the stock market [10] and [11] which works to calculate the price of an asset or a liability. This model allows to establish the behavior projection of the share in a specific period of time, and in this paper the congestion volume of a LAN network is the variable; therefore, this kind of analysis let to make an estimation considering historical behavior of data which are being analyzed in periods of measurement.

It is considered that ARIMA($p, d, q$) model is defined by:

$$Z_t = Z_{t-1} + \theta_0 + a_t \quad (11)$$

$$Z_t = Z_{t-2} + 2\theta_0 + a_t + a_{t-1} \quad (12)$$

$$Z_t = Z_k + (t-k)\theta_0 + \sum_{j=k+1}^{t} a_j \quad (13)$$

Taking this work as a reference [11]:

- $Z_t$ is the time series of data from expected return of the capital over asset $t$.
- $Z_k$ is the time series of historical data.
- $\theta_0$ is the parameter which determines the series trend.
- $\sum_{j=k+1}^{t} a_j$ are the rest of the terms of the data series.

3. Congestion control in data networks

Congestion in data networks has become a proper phenomenon of this technology, and it comes up in several parts of the network architecture, beginning in the sizes of the routers buffers which are responsible for make the routing of the packets. Those buffers are limited, and they can be shaped as a type of queue M/M/1/ N/∞, where the queue size is $N$ which is a limited resource; therefore, when the queue is full, a known consequence starts to come up, denoted packet loss which can be measured as likelihood of blockage.

$$P_B = P_N = \frac{(1-\rho)\rho^N}{1-\rho^{N+1}} \quad (14)$$

In equation 14, $P_B$ represents the likelihood of blockage of requests in a data network, or the likelihood of discarding packets. $\rho$ is the traffic intensity travelling through connections or flow of packets. $N$ is the number of channels or the size of buffers in the switching systems.

Equation 14 develops as follows:

If we have $\lambda$ = arrival rate of requests $y$ $\mu$ = service rate, the next graphic appears:

![Diagram of states for likelihood of blockage](image)

Based on the diagram of state, it is possible to propose equations of balance for different states:

<table>
<thead>
<tr>
<th>State</th>
<th>Output rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$P_0$</td>
</tr>
<tr>
<td>1</td>
<td>$P_1 + P_1ls$</td>
</tr>
<tr>
<td>2</td>
<td>$P_2 + P_2ls$</td>
</tr>
<tr>
<td>$n$</td>
<td>$P_n( + )$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>$P_0 + P_1ls$</td>
</tr>
<tr>
<td>$P_1 + P_2ls$</td>
</tr>
<tr>
<td>$P_n( + )$</td>
</tr>
<tr>
<td>$P_n( + ) + P_{n+1}$</td>
</tr>
</tbody>
</table>

www.astesj.com
Then, we have:

\[
P_1 = \frac{p_0 \lambda}{\mu}
\]

\[
P_2 = p_0 \left(\frac{\lambda}{\mu}\right)^2
\]

\[
P_n = p_0 \left(\frac{\lambda}{\mu}\right)^n
\]

If it is known that \( \sum_{n=0}^{N} P_n = 1 \), then:

\[
P_n = \sum_{n=0}^{N} P_0 \left(\frac{\lambda}{\mu}\right)^n = 1
\]

\[
P_0 = \frac{1}{\sum_{n=0}^{N} \left(\frac{\lambda}{\mu}\right)^n}
\]

\[
P_n = \frac{\left(1 - \frac{\lambda}{\mu}\right)}{1 - \left(\frac{\lambda}{\mu}\right)^{N+1}}
\]

Now, it is possible to say that \( \lambda = \rho \mu \) which is the trafficking intensity of requests or packets for a data network.

Equation 14 represents likelihood of blockage of a used model at the beginning of telephone systems. In mathematical terms, it models the value of the service quality that would be given to a network, depending on the size of the buffer. In the course of evolution of telecommunications networks, specifically data networks, equation 14 has been adapted to measure resources and fulfill demand needs of the users.

However, the size of the buffers is not the only one resource that affects the congestion of data networks, but also the processing level of the network elements, and the capacity of transporting information, such as the connections bandwidth between nodes. Bandwidth is one of the most important resources in data networks that directly affects congestion because it is the main aspect where packet transport becomes tangible and from which is possible to measure service deterioration.

As fig. 6 depicts, congestion can be represented as a downward curve that drops depending on the number of flows that enter the system. If system turns congested, the more flows they enter, the less efficiency of network performance.

To demonstrate this behavior, the limit when \( \rho \) tends to infinity can be calculated; thus, we have:

\[
P_n = \lim_{\rho \to \infty} \frac{(1 - \rho) \rho^n}{1 - \rho^{N+1}}
\]

\[
P_n = \frac{d}{d\rho} \left[\frac{(1 - \rho) \rho^n}{1 - \rho^{N+1}}\right]
\]

\[
P_n = \frac{1}{N + 1}
\]

This expression defines the congestion shape in a data channel. This behavior can be replicated not only in connection flows, but also in different junction sections of data connections; for instance, the formed bottlenecks when there is a transfer from distribution network to access network because the mean of transmission has to change due to the architecture, as is seen in HFC networks (Hibrid Fiber Coaxia) that mix fiber optic with coaxial cables.

3.1. Congestion control with TCP

Congestion control models have been focused on the transport layer protocol TCP [12] and [13], in such a way that flow control by sliding window protocol, AQM technique (Active Queues Management) [14], admission control, and delay time management are the more used techniques. All of this is made by different changes or versions to the TCP protocol, beginning by TCP Tahoe and TCP Reno, which modify retransmission of ACK packets or confirmations.

These TCP modifications have led to study new forms of making congestion control, which have been focused on transmission states that define how algorithm is executed. This sequence can be modified according to necessities of the desired control over data flow. For instance, in [12] the TCP states define how the algorithm of congestion control changes, from slow start state towards congestion avoidance state. During transition, the size of retransmission window is affected by connection state.

The main goal of this paper is to propose a model of control without any change to the protocol because it leads to make modifications in the times of transmissions, confirmations, and retransmissions of the network elements, and these changes could worsen the network behavior; therefore, a model without modifications in the protocol structure is proposed to help to treat the behavior of network elements for their adaptation to the flow of produced data from random behavior of users.

3.2. TCP basis

RFC 793[15] describes TCP functioning, where window of congestion control is not dynamic. First variation implemented on TCP protocol is the fitting of transference rate to the level of current congestion presented in data network. The purpose of making this reduction of transference rate is to diminish the number of lost packets in connections. Next steps of change in TCP protocol lead to TCP Tahoe algorithm which modifies slow start and congestion avoidance algorithms, as well as fast retransmission. In TCP Reno another modifications happens, consisting in TCP Tahoe but with a variation in the fast recovery algorithm [16].

The main functions of slow start and congestion control are the control of transference rates of generated packets and the
determination of available capacity of network connections; however, making a process in congestion window is needed to achieve this. This process is denoted \( cwnd \), and it is defined as the number of transmitted bytes whose confirmation is not required. Another considered element is the recovery window denoted \( rwnd \) which is the number of bytes that receptor is capable of receiving from a given source [17]–[19].

To establish connection between two network elements, activity of information transport can be described as follows:

- Maximum size is assigned to \( cwnd \) by transmitter.
- The transmitter starts a retransmissions timer.
- Slow start algorithm sends a packet or a section if an ACK is received; then, the \( cwnd \) value is augmented in one section.
- After that, two sections can be sent and cause two ACK.
- For each ACK, \( cwnd \) is augmented in one section.
- Growth process of congestion window is exponential and keeps this rhythm as far as the bounded value is reached by \( rwnd \) and established threshold is attained by ssthresh.
- Congestion avoidance algorithm is executed when ssthresh value is obtained and the \( cwnd \) value is augmented in one additional section, but this time, by each ACK and RTT.
- When packets loss happens, the process returns to the slow start: \( cwnd \) is again established to one section and ssthresh to half the window size.

Having knowledge about TCP protocol functioning and how congestion control is made, a congestion control model based on an econometric model denoted ARIMA is pose, which will be fused together with TCP mechanism to measure congestion control of a LAN network.

4. **Approach of the model**

According to the analysis made in section 2.1 and the congestion control functioning explained in section 3, the relation between ARIMA\((p, d, q)\) model and the congestion control is presented as follows:

- A time series with data of packets circulating in the network must be taken.
- \( Z_t \) parameters for this time series must be calculated.
- The function of yield rate for \( Z_t \) must be estimated.
- Transmission rates must be fitted with obtained results, and the congestion control must be verified.

To attain this, data are taken from a working LAN network through obtained traces from the network interface of a Switch gate. Analogy between compilation of LAN network data and econometric model is the way in which statistics of variables that work in the network are represented. Each feature is taken as one share of a product, leading to analyze the behavior of one or more shares to know how market behavior will be. In this case, market is the whole data trafficking through LAN network. Once the time of network congestion is known, the congestion control operation is made.

Variables taken in network behavior analysis are: \( CD \) (*amount of connected devices*), \( PU \) (*percentage of use*), and \( RT \) (*response time*) which lead to following result:

![Fig7: Behavior of variables in LAN network](image)

Fig. 7 depicts behavior of variables in LAN network, each one separately. It is worth noting that each variable is affected by the congestion phenomenon; for instance, if congestion comes up, response time begins to increase the same as percentage of use, and the connected devices can undergo complications in different services and applications. Thus, model is applied to detect congestion which is identified by increases in response time or percentage use. When the number is detected, transmission rates will be modified to alleviate the LAN network problems.

Test environment is a LAN network with hierarchical structure whose *core* layer is the most rugged, and devices are router of Cisco brand. In second layer, i.e. distribution layer, Cisco *switch catalyst* are found, with management capacity. Finally, all devices of the network users are connected to the third layer, i.e. access layer, where congestion control is made through manipulation of the operating system of the *switches* found in this layer. Fig. 8 depicts test environment distribution.

![Fig8: LAN network test environment](image)

To apply the model, the following methodology is used: simulation of one transmission control system for TCP is made (fig. 9). This model has a traffic source that feeds the mechanism of TCP; then, RED (*Random Early Detection*) module is found, where transmission rates are modified; after that, feedback module is found which allows to make the modification to the mechanism.
of TCP with data being transmitted. This model has been based on previous studies [6], [20], [21]

TCP control system functioning in fig. 7 is described as follows:

\[ r_k = \frac{MK}{RTT \sqrt{P_{k-1}}} \]  
\[ q_k = \left( \frac{nr_k}{M} RTT - \frac{C}{M} R_0, B, 0 \right) \]  
\[ \bar{q}_k = (1 - w) \bar{q}_{k-1} + w q_k \]  
\[ p_k = \frac{\bar{q}_k\min_{ch}}{\max_{ch}-\min_{ch}} P_{max} \]

For previous equations we have:

- \( M \) = packet size
- \( RTT \) = Round Trip Time value
- \( p_k \) = likelihood of dismiss
- \( B \) = buffer size
- \( n \) = TCP number of flows
- \( R_0 \) = RTT minimum value
- \( C \) = connection capacity
- \( K \) = constant 1.25

During model execution, the modified variable is \( n \). This number is modified by \( N_t \) value which is calculated with taken values from the LAN network trace.

Fig. 10 shows \( Z_t \) series which represents taken response times to analyze one LAN subnetwork.

### 5. Results of the model

To examine the behavior of the proposed model in the last section, one data series is taken from LAN network. For this data series, \( Z_t \) series is created taking as reference the following values:

- \( \theta_0 \) trend of the series = limit of response times or percentage of use.
- \( Z_k \) market yield = LAN network throughput in conditions of no congestion.
- \( \sum_{f=k+1}^{\infty} a_f \) yield of one risk-free asset = throughput of channel capacity.

Taking those variables about TCP control system, we have:

\[ q_k = \left( \frac{2e^x}{M} RTT - \frac{C}{M} R_0, B, 0 \right) \]

Fig. 11 shows the comparison between behavior of two types of data flows supported by TCP control system. Flow depicted by the continuous line is the normal flow in which congestion control has not been applied, and flow depicted by the dashed line is the flow with \( q_k \) modification which is the variable affected by \( E(r_t) \).

Fig. 12 shows another way to look at congestion control. The throughput measure of network is made, where that network is saturated to observe model’s behavior. In this environment, network is saturated by injecting an amount of traffic, and response times are measured. When those are risen, the transmission rate keeps a constant value of 10 Mbps. On the other hand, when network is no longer saturated, transmission rate rises above 10 Mbps because congestion control is no longer applied.
Fig 12 shows a comparison between different subnetworks of the LAN network when congestion experiment is made to saturate subnetwork under observation. An interesting result is that the operation only affects the network where congestion is generated, which when translated into technical work, affects the transmission rates only in the gate where subnetwork is connected and in which the experiment is done. In this comparison, channels independency is shown, and a good functioning of the whole LAN network is guaranteed.

![Fig 12: Comparison between subnetworks](image)

As a final part of congestion control test on test subnetwork, lost packets are of particular interest because one of the symptoms of congestion is the increase of lost packets in connection. Fig. 13 shows a test subnetwork behavior. At the beginning of the experiment, the number of lost packets rises when the channel is saturated, but then it drops substantially when congestion control is carried out in the switch gate.

![Fig 13: Lost packets in a subnetwork](image)

6. Conclusions

This paper has explained a congestion control proposal, where two central concepts are fused together. On the one hand, ARIMA \((p, d, q)\) econometric model works with time series, calculating parameters of a linear function to estimate future values. On the other, TCP conceptual model has been taken to make the congestion control, based on RED (random early detection). Following the analogy in which a LAN network is like a stock exchange and the stock market shares are like the network variables that are produced with circulating traffic, a model with a \(Z_t\) series is pose, which is the function that represents the estimation for one asset. In this proposal, the modified parameter is \(n\), which is the number of traffic flows in the network. Moreover, three network LAN variables have been taken: RT (response time), CD (connected devices), and PU (percentage of use), and one experiment to saturate LAN network to analyze RT behavior with the congestion control model integrated into the system was carried out.

Results demonstrate that the way in which congestion control was made works, and although only the behavior of one variable (RT) has been illustrated, it would be expected that works with the other variables. On the other hand, the used strategy to make traffic control is the transmission rates variation. In summary, it is promising the way in which concepts from other fields or disciplines are being used, since it is possible to successfully adapt the econometric model. Further tests and measurements in LAN network and other data networks like Internet should be carried out to provide better tools to alleviate the congestion experienced nowadays.

References


