

Desarrollo de un Modelo Predictivo para la Estimación del Comportamiento de Variables en una Infraestructura de Red

Danilo A. López*, Nancy Y. García, Jhon F. Herrera

Universidad Distrital "Francisco José de Caldas", Carrera 7 No. 40 – 53, Bogotá-Colombia.
(e-mail: dalopez@udistrital.edu.co, nygelvezg@udistrital.edu.co, jfherrerac@udistrital.edu.co)

* Autor a quien debe ser dirigida la correspondencia.

Recibido Abr. 16, 2015; Aceptado Jun. 8, 2015; Versión final Jun. 24, 2015, Publicado Oct. 2015

Resumen

El artículo presenta los resultados encontrados al modelar la variable de red *perdida de paquetes* (para un flujo de datos de señales de televisión) y *rendimiento* (para una secuencia de tráfico de video digital) en una topología de red inalámbrica. Esto se hace para estimar su comportamiento futuro haciendo uso de modelos autorregresivos integrados con promedios móviles (ARIMA). Todo esto permite identificar la robustez de este tipo de series de tiempo para ser aplicado en las redes interconectadas con el fin de mejorar su rendimiento. El desarrollo de la propuesta se basó en el uso de la metodología Box-Jenkins. Este tipo de series de tiempo univariadas son un buen candidato para generar pronósticos de comportamiento en las redes de telecomunicaciones.

Palabras clave: series de tiempo, ARIMA, modelos predictivos, Box-Jenkins

Developing a Predictive Model to Estimate the Behavior of Variables in a Network Infrastructure

Abstract

The article presents the results found by modeling the variable network packet loss (for a data stream of television signals) and throughput (for a digital video traffic sequence) in a wireless network topology. This is done to estimate its future behavior using autoregressive models integrated with moving averages (ARIMA). All this allows identifying the strength of this type of time series to be applied in internetworking with the purpose of improving its performance. The proposal development is based on the use of the Box-Jenkins methodology. This type of univariate time series is a good candidate to generate behavior predictions in telecommunications networks.

Keywords: time series, ARIMA, predictive models, Box-Jenkins

INTRODUCCIÓN

Desde hace varios años, la sociedad en general viene dependiendo de las ciencias de la información y las comunicaciones como medio de transmisión y divulgación de información de manera rápida y eficiente; y en este apogeo que viven actualmente las redes de telecomunicaciones, la cantidad de información que deben procesar los nodos de los proveedores de servicios de Internet (López et al., 2014) para enviar eficientemente el tráfico circundante es muy elevado, lo que hace cada vez más imperativo se comiencen a proponer posibles soluciones a las múltiples problemáticas existentes en lo relacionado al transporte eficiente de datos (Haghighi et al., 2014), dentro de los que se encuentra la generación de mecanismos de enrutamiento más eficientes que incluyan como variable de selección, poder llegar a estimar el comportamiento futuro del tráfico y de las múltiples variables (throughput, pérdida de paquetes, jitter, etc) asociadas a un flujo de datos determinado, con el fin de generar rutas más óptimas a destinos a partir del pronóstico del comportamiento del tráfico y/o variables de red (García et al., 2015). Es por ello, que el modelamiento de tráfico en las internetworking, se está convirtiendo en una clara línea de investigación donde el objetivo principal es proponer y/o construir modelos o metodologías que pronostiquen el comportamiento de variables de red, que conduzcan a la toma de decisiones más acertadas en la fase de enrutamiento

De las diversas opciones existentes para pronosticar datos, una de las más interesantes es la basada en modelos estadísticos, pues representan de manera óptima procesos estocásticos haciendo uso de paradigmas probabilísticos, además de que permiten tratar la no continuidad del tráfico de datos (paquetes, ráfagas y sesiones) (Chen, 2007). En este sentido, el presente paper propone hacer uso de las series de tiempo para estimar el comportamiento futuro (de manera independiente) para las variables de red *throughput* y *pérdida de paquetes* en una estructura de red inalámbrica. Para ello, se comienza por definir el concepto de series de tiempo, seguidamente se describe la metodología aplicada en la investigación para generar el modelo de tráfico; finalmente se presentan los resultados encontrados y las respectivas conclusiones.

En principio, se puede definir una serie temporal como una colección de observaciones o datos tomada a lo largo de cierto tiempo, cuyo objetivo principal es describir, explicar, predecir y controlar algún proceso. Las observaciones están ordenadas respecto al tiempo, donde las sucesivas observaciones son generalmente dependientes; característica que es de vital importancia a la hora del análisis de la serie (Correa, 2004), (Yanhua et al., 2010). De forma general, las series de tiempo por si solas se reducen a una simple organización temporal de muestras que permiten la descripción de algunos parámetros básicos de un proceso estocástico. No obstante lo anterior, lo que realmente representa un gran aporte en el estudio de este tipo de series son las metodologías y los modelos matemáticos que se han desarrollado en torno a la descripción y predicción de variables aleatorias (Romani et al., 2013), en especial los modelos autorregresivos, de medias móviles y todos los que han derivado de la combinación de estos dos (Min et al., 2004).

MODELADO DE SERIES DE TIEMPO

La descripción de algunos de los estándares más representativos en el modelado de series de tiempo requiere de notación específica, como sigue (Martínez y López, 2013):

Y_t . Serie de tiempo que será analizada.

u_t . Función de blanco con promedio cero y varianza constante.

d y D . Grados de diferenciación normal y estacional.

$\Phi_p(L)$. Polinomio de orden p del componente autorregresivo.

$\Phi_P(L)$. Polinomio de orden P del componente estacional autorregresivo.

θ_q . Polinomio de orden q del componente de medias móviles.

$\theta_Q(L)$. Polinomio de orden Q del componente estacional de medias móviles.

S . Período de la función si presenta estacionalidad.

μ . promedio de la función original sin diferenciar.

Modelo Autorregresivo (AR) p

Con este modelo, descrito en la ecuación (1), se expresa el valor actual de la serie como una función de los valores que tomó la misma en las p muestras anteriores ponderadas por un factor φ_i y de una perturbación aleatoria presente.

$$\Phi_p(L)Y_t = u_t \quad (1)$$

Modelo de medias móviles (MA)_p

Considera que el valor de la serie estacionaria se desplaza alrededor de un valor medio μ . Además supone que el desplazamiento de μ en el tiempo presente t es ocasionado por infinitas perturbaciones ocurridas en el pasado, ponderadas por un factor θ_i que mide la influencia de cada una de ellas en el valor presente de la serie. Su descripción matemática se plasma en (2).

$$Y_t - \mu = \theta_q(L)u_t \quad (2)$$

Modelo ARMA

Como se muestra en la ecuación (3), este modelo representa una serie de tiempo como una combinación de los dos modelos anteriores:

$$\Phi_p(L)Y_t = \theta_q(L)u_t \quad (3)$$

Modelo ARIMA

Otro tipo de serie temporal es ARIMA, donde se asume que el comportamiento del proceso estocástico Z_t se puede explicar a través de sus valores pasados y de una suma ponderada de errores pasados y presentes. Una característica relevante de este tipo de modelos es que una o más de las raíces de la ecuación característica definidas por el factor autorregresivo del modelo, es mayor o igual a la unidad, tendiendo la serie a un valor no estacionario; comportamiento contrario al que se aprecia con ARMA (ecuación 3), donde las raíces se encuentran dentro del círculo unitario, garantizando que la secuencia de datos es estacionaria. Esto implica que algunas series (como ARIMA) deben ser diferenciadas para eliminar tendencias o varianzas cambiantes y así obtener funciones estacionarias.

Específicamente, el modelo ARIMA hace referencia a un sistema ARMA que se ha aplicado a una serie diferenciada. La ecuación 4 representa su forma matemática, que a diferencia de la ecuación 3, presenta el término de tendencia determinística θ_o (equivalente a la media del proceso cuando la serie es estacionaria (valor que se suele omitir a menos que sea realmente necesario)) y la serie diferenciada $(1 - L)^d$, de acuerdo con el orden de integración de la misma.

$$\Phi_p(L)(1 - L)^d(Y_t - \mu) = \theta_o\theta_q(L)u_t \quad (4)$$

ESTIMACIÓN DE LAS SERIES DE TIEMPO*Metodología de simulación*

Para las estimaciones en series de tiempo de las variables que se abordan en esta investigación, se empleó la metodología Box-Jenkins (Hernández, 2006) debido a que solo involucran sumas de términos. El modelo consta de 4 etapas que incluyen: la identificación del modelo, estimación de los parámetros implícitos del modelo, verificación de supuestos y uso del modelo. La Fig. 1 resume la metodología descrita. Específicamente este método consiste en proponer un modelo ARIMA que describa la manera más óptima posible acerca de la realidad de un proceso, de una variable en un proceso estocástico; es decir de un proceso aleatorio, todo esto en función del tiempo pasado y presente de la misma variable.

Red implementada para la obtención de los datos

Como el modelo de estimación propuesto, se enfoca en la predicción del comportamiento del *throughput* de red para una secuencia de señales de televisión y el pronóstico del comportamiento de la variable *pérdidas* para una transmisión de video digital haciendo uso de series tipo ARIMA; la estructura de red de donde se obtuvieron los datos es la mostrada en la Fig. 2; donde la fuente de los datos se encuentra almacenada en el "Servidor CECAD" y el destino de los mismos en el equipo de cómputo identificado como "Cliente". Para ello se recurrió al uso de la herramienta Wireshark con el fin de monitorear la red y hacer un seguimiento al flujo de datos objetivo y así poder monitorear las variables de trabajo. También se utilizó la aplicación StreamAnalysis, la cual permite extraer información como el nivel de pérdidas de paquetes durante una transmisión, o el *throughput* existente en la red. Cabe destacar que se hizo necesario en Wireshark realizar un filtrado de datos por número de puerto, con el fin de disminuir la captura de paquetes a las secuencias objetivo en aquellos instantes de tiempo donde existía más de un flujo simultáneo.

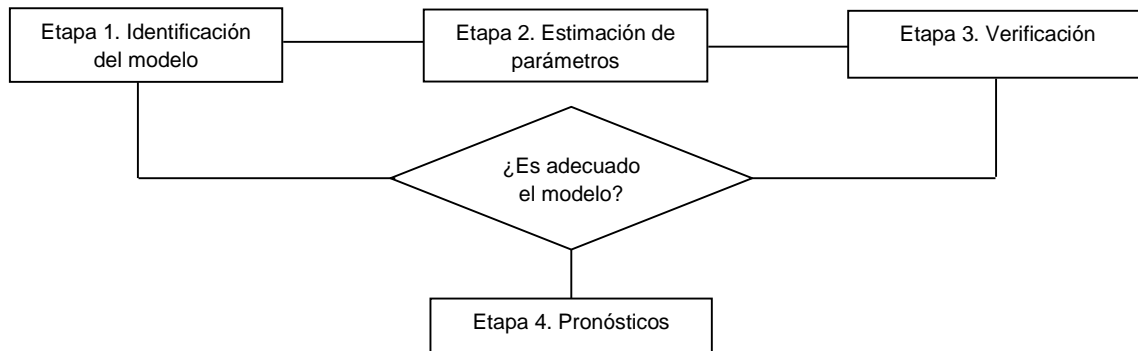


Fig. 1: Metodología de Box-Jenkins (Hernández, 2006).

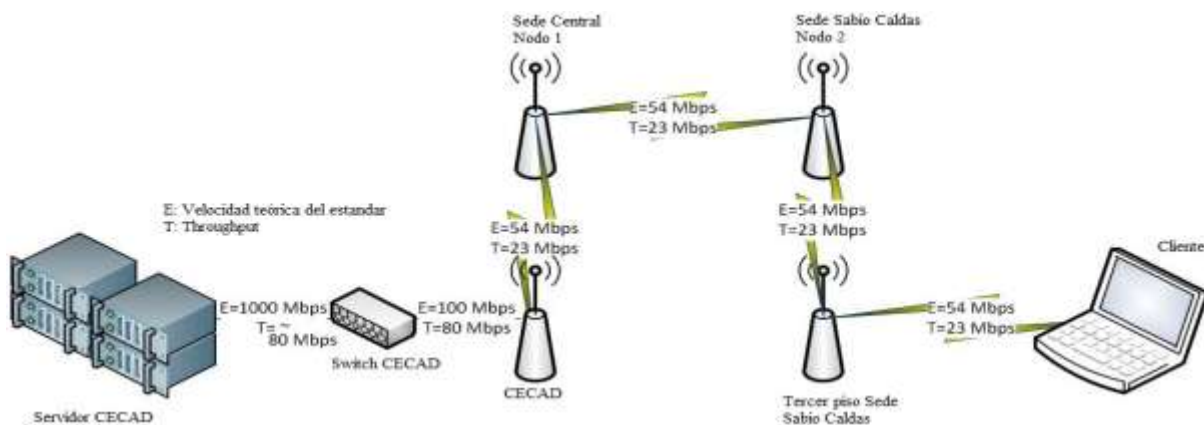


Fig. 2: Estructura de la Red de Prueba.

DESARROLLO DE LOS MODELOS

Dado que el análisis propuesto se enfoca en la *pérdida de paquetes* y estimación del *throughput* (variables que fueron estudiadas independientemente) se eligieron las series temporales univariadas para la representación de los datos, teniendo en cuenta que permiten analizar el comportamiento de la variable en sí misma sin pretender explicar los factores que influyen en esta. A continuación se presentan detalladamente los pasos que se siguieron en la obtención de los modelos ARIMA (Martínez y López, 2013).

Organización de los datos

Para poder hacer una buena representación, se debe escoger un intervalo de tiempo que capture de alguna forma un comportamiento descriptivo para el patrón que se desea analizar. En econometría por ejemplo, este problema depende generalmente de la periodicidad con la que se obtengan los datos (índices mensuales, anuales, trimestrales, etc). En el tráfico de datos cada paquete transmitido porta encabezados que brindan información de manera casi continua, por lo cual se debe escoger un intervalo que arroje datos globales. En general, intervalos de tiempo muy cortos arrojarán comportamientos más puntuales, mientras que periodos muy largos harán énfasis en comportamientos más globales (Martínez y López, 2013), (Haghighi et., 2014).

Para la primera variable de estudio (throughput), se analizaron las capturas obtenidas de una de las transmisiones de televisión; y de los 1900 datos obtenidos se graficaron los primeros 1500, dejando los últimos 400 para realizar la validación del modelo. Para la segunda variable de análisis (perdida de paquetes), se seleccionaron las capturas de una de las transmisiones de video digital y se tomaron solo 200 muestras (dejando para validación las últimas 50) ya que en dicho momento de la captura de la información el canal no se encontraba saturado.

Seleccionados los datos, se procede a la graficación de los mismos para observar su evolución a lo largo de la variable independiente (tiempo). En la Fig. 3 y Fig. 4, aparecen las secuencias de la serie original para cada caso.

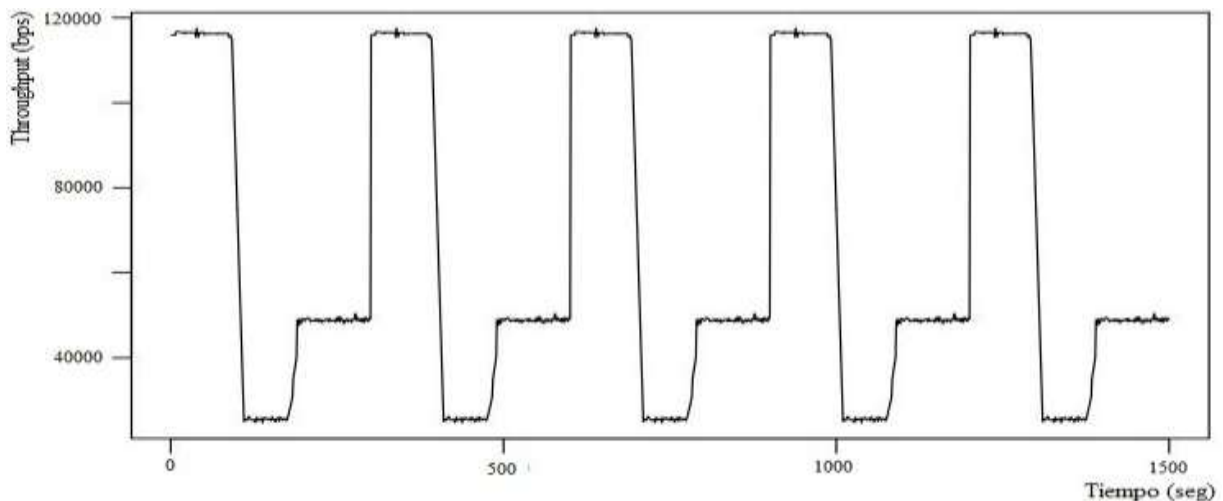


Fig. 3: Serie de tiempo para la variable Throughput (Gómez y Ramírez, 2013).

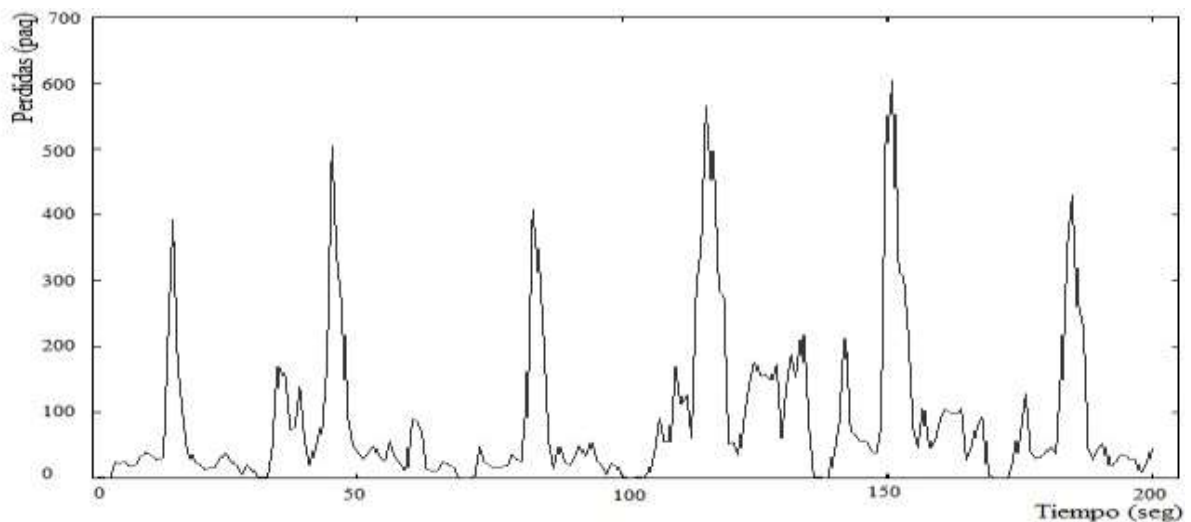


Fig. 4: Serie de tiempo para la variable perdida (Martínez y López, 2013).

Estacionariedad de las series

Caso A. Serie Throughput

Con el fin de determinar la estacionariedad de la primer serie (variable throughput) se genera en la Fig. 5, la función de autocorrelación simple (ACF) y parcial (PACF) (Taj y Akil, 2011). A simple vista y a partir de la definición de estacionariedad (que exige una media igual a cero y una varianza constante) se puede intuir que la serie no es estacionaria. Para verificar cuantitativamente la anterior tesis se realiza la prueba de la Raíz Unitaria de Dickey –Fuller usando el software R.

De acuerdo con el test de raíz unitaria en la muestra se determina que la muestra no es estacionara debido a que el valor del test es menor en magnitud al valor crítico del 5%; razón que hace necesario diferenciar la serie (teniendo en cuenta que cada diferenciación le resta información a la señal original, generando errores en la secuencia). Para este caso la serie se diferenció una vez, quedando la misma con el aspecto mostrado en la Fig. 6. Al graficar nuevamente la función de autocorrelación se encuentra lo mostrado en la Fig. 7.

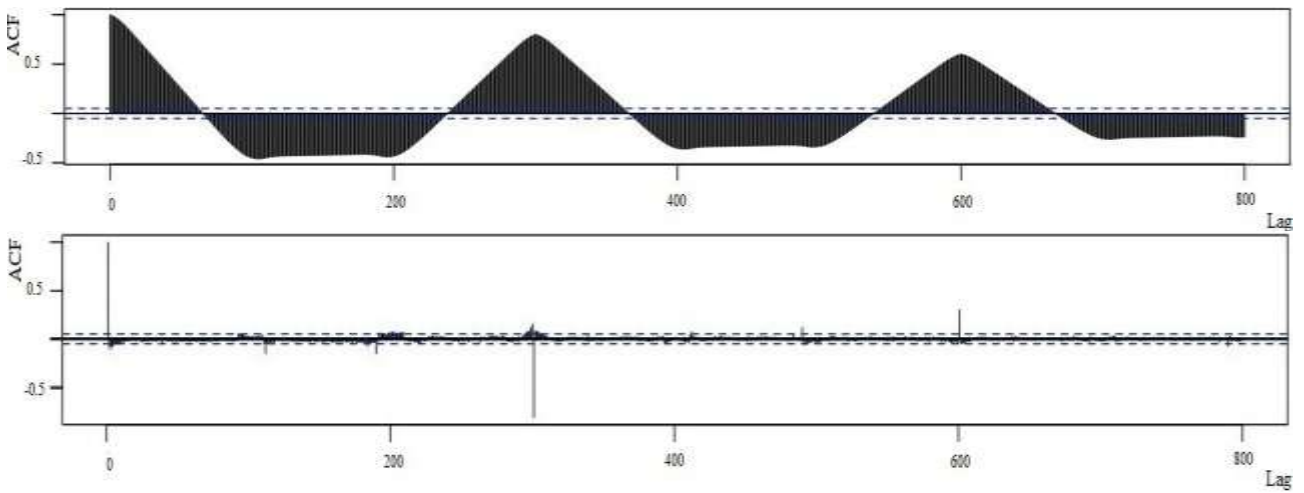


Fig. 5: Autocorrelación simple y parcial para la serie (Gómez y Ramírez, 2013).

Los resultados obtenidos son:

Value of test-statistic is: -1.3286

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

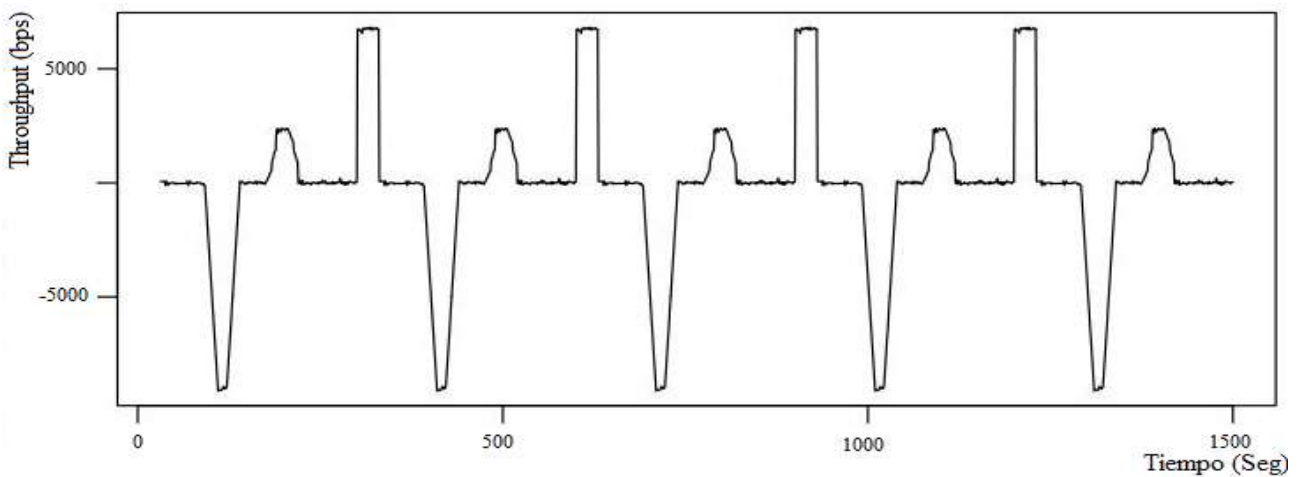


Fig. 6: Diferenciación de la serie de tiempo original.

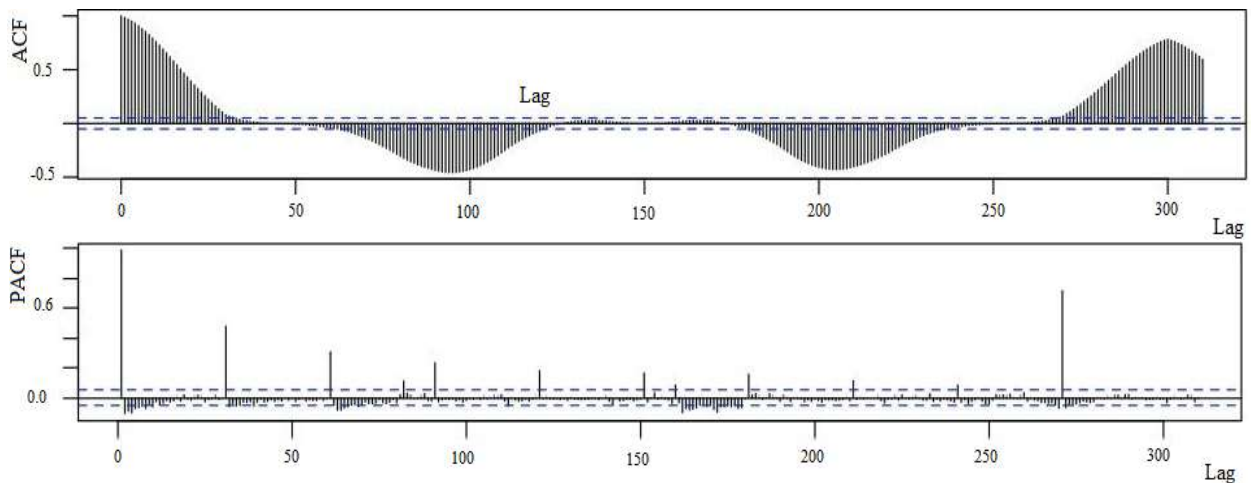


Fig. 7: Autocorrelación simple y parcial después de diferenciada (Gómez y Ramírez, 2013).

Aplicando nuevamente el test de raíz unitaria, los resultados son:

Value of test-statistic is: -3.4543

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

A partir de la Fig. 7, y del test Dickey-Fuller, se encuentra que su valor es aceptable para poder calcular el modelo ARIMA ya que la muestra cumple con el criterio de estacionariedad.

Caso B. Serie pérdida paquetes

Siguiendo las pautas aplicadas en el caso anterior, de la Fig. 8, se observa que no existe una tendencia (creciente o decreciente) o media constante, lo que implica que la serie no es estacionaria.

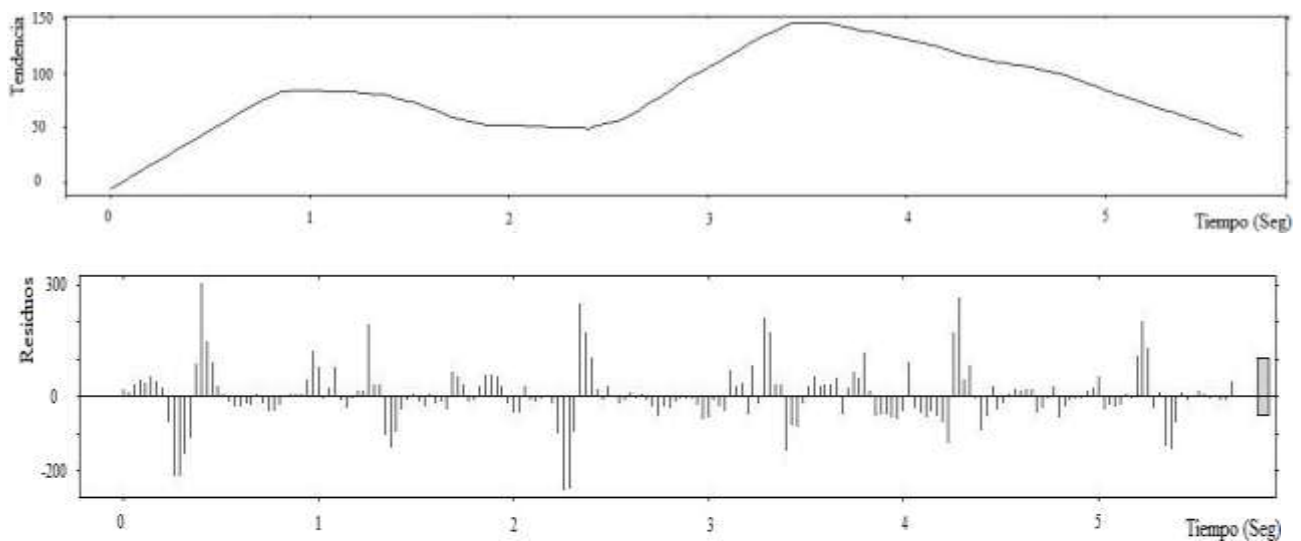


Fig. 8: Análisis de la serie original (Martínez y López, 2013).

Una vez diferenciada la serie, la estacionariedad de la nueva secuencia se verifica con las funciones ACF Y PACF en la Fig. 9, donde se ve claramente la existencia de un carácter periódico en la autocorrelación parcial (grafica inferior). Como se puede observar, estas funciones arrojan valores significativos en los retardos cercanos a 34 y demás múltiplos de la frecuencia, lo cual muestra el comportamiento periódico de la serie.

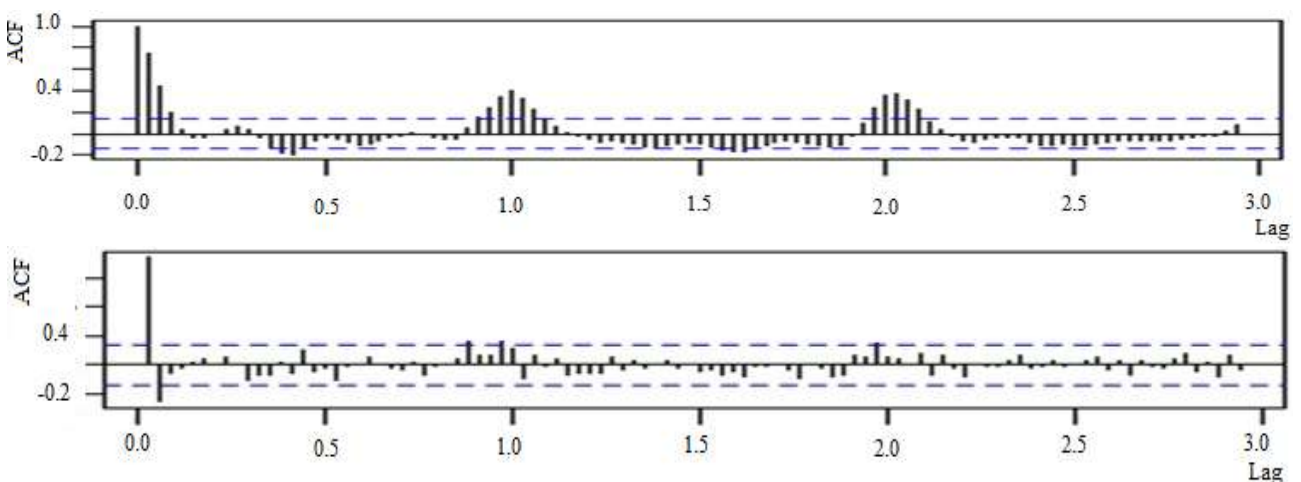


Fig. 9: Funciones de autocorrelación y autocorrelación parcial de la serie.

Una vez las muestras cumplen con el criterio de estacionariedad, la generación y selección de un buen modelo va a depender del análisis de los coeficientes: factor de ajuste y nivel de predicción; valores de selección que son diferentes pero complementarios entre sí, los cuales se analizan seguidamente para cada estudio de caso.

AJUSTE DEL MODELO ARIMA

Al proponer la generación de modelos, es frecuente encontrarse con múltiples alternativas que pudieran estimar el comportamiento de la variable estudiada, y se debe decidir cuál es la mejor opción para determinar la aproximación (no predicción) más acertada para la serie tratada (Brockwell y Davis, 1986) (Martínez y López, 2013). Dentro de las metodologías existentes se encuentran los criterios de: parsimonia, calidad de ajuste y el error cuadrático medio (RMSE). En la investigación, se hizo uso del ajuste; valor obtenido mediante el uso del criterio de información de Akaike (AIC), que hace referencia a una medida de la calidad relativa del modelo estadístico, basando su análisis en la entropía a través de la estimación relativa de la información perdida.

Caso A. Serie Throughput

Un modelo robusto en series temporales, utiliza el histórico de los datos para poder reproducir apropiadamente el comportamiento ya conocido (a partir del aprendizaje de la forma más adecuada de hacerlo) con el fin de obtener un comportamiento futuro muy cercano a la realidad (Hernández, 2006); en este sentido y mediante la metodología de Akaike, se analizaron un total de 100 modelos mediante la variación de los parámetros p y q , de los cuales el 18% presentaron errores de convergencia. En la tabla 1, aparecen los valores encontrados para los tres modelos más acertados.

Tabla 1: Calidad de ajuste a partir del criterio de Akaike.

Modelo de la serie temporal	Coficiente AIC
ARIMA(1,1,4)	1644,1794
ARIMA(34,1,0)	1657,0031
ARIMA(3,1,4)	1647,9041

Los resultados sugieren que el modelo ARIMA(1,1,4) es el que tiene un menor valor, condición que garantiza que es la mejor opción para reproducir la serie de datos para la variable estimada. Es importante destacar que el ajuste del modelo se hizo sobre parte de la serie original (primer 79% de las muestras), teniendo en cuenta que la parte restante es la que se utilizó para generar el modelo de predicción (Hirotsugu, A, 1974).

Caso B. Serie pérdida de paquetes

Mediante un procedimiento análogo al caso A, se llegó a que los resultados más representativos para los modelos fueron los de la tabla 2, teniendo en cuenta que el ajuste se aplicó sobre el 75% de la secuencia.

Tabla 2: Calidad de ajuste a partir del criterio de Akaike.

Modelo de la serie temporal	Coficiente AIC
ARIMA(1,1,3)	1650,2403
ARIMA(0,1,36)	1678,2622
ARIMA(3,1,3)	1653,6172

El prototipo que presento el mejor ajuste fue el ARIMA(1,1,3), valor muy por debajo del segundo y tercer piloto mostrado en la tabla 2.

PREDICCIÓN DE LA SERIE

Esta etapa incluyó para cada variable de red, la predicción de un conjunto de datos con el fin de determinar el grado de acierto de los modelos para estimar los datos reales (Martínez y López, 2013). Cada uno de los modelos se ajustaron sobre parte de las muestras (ver ítem anterior: Ajuste del modelo ARIMA), y la predicción se hizo sobre los datos restantes.

Caso A. Serie Throughput

A partir de los periodos de estacionalidad de la Fig. 8, 9, y de la aplicación de las reglas establecidas en (Escrinche, 2014) se obtuvieron los valores de las variables p y q para la serie ARIMA(p,d,q). La tabla 3, resume algunos de los modelos encontrados junto con el error cuadrático medio (RMSE) presentado por cada predicción frente a los valores reales de la serie original.

Tabla 3: Valores RMSE para diferentes modelos.

Modelo	RMSE
ARIMA(0,1,3)	133,779280
ARIMA(34,1,0)	Error
ARIMA(4,1,4)	125,738997
ARIMA(3,1,34)	130,278409

Se puede verificar que para valores de $p=4$ y $q=4$, la predicción encontrada en el modelo ARIMA (4,1,4) fue la que presento el error medio cuadrático más bajo, lo que redundará en mejores predicciones, ya que los valores estimados están en promedio más cerca de la serie original. A partir de estos valores y con la ayuda del software R se calcularon los coeficientes del modelo (ver tabla 4)

Tabla 4 : Coeficientes de la serie ARIMA.

ar1	ar2	ar3	ar4	ma1	Ma2
0.0701	0.0055	-0.0859	0.0321	0.0064	0.0241
Ma3	Ma4				
0.0034	0.0255				

De acuerdo a estos coeficientes, finalmente se tiene el modelo de predicción ARIMA (4,1,4) de la ecuación 6.

$$Z_t = 0.0701 * Z_{t-1} + 0.0055 * Z_{t-2} - 0.0859 * Z_{t-3} + 0.0321 * Z_{t-4} + a_t - 0.0064 * a_{t-1} + 0.0241 * a_{t-2} - 0.0142 * a_{t-3} - 0.1225 * a_{t-4} \quad (6)$$

Caso B. Serie pérdida de paquetes (Martínez y López, 2013)

Análogamente, algunos de los valores RMSE para la variable *perdida de paquetes*, se muestran en la tabla 5. De ella se concluye que el modelo más acertado tiene los parámetros $p = 3$ y $q = 34$, llegando al modelo de predicción ARIMA (3,1,34) de la ecuación 7.

Tabla 5: Valores RMSE para diferentes modelos.

Modelo	RMSE
ARIMA(0,1,3)	130,72478
ARIMA(36,1,0)	Error
ARIMA(0,1,36)	134,99471
ARIMA(3,1,34)	127,000812

$$(1 - 0,6514L - 0,0213L^2 + 0,0377L^3)(1 - L)(Y_t) = (1 - 0,7874L + 0,0962L^2 + 0,1728L^3 + 0,0352L^4 - 0,1124L^5 + 0,0191L^6 - 0,0216L^7 + 0,1104L^8 - 0,3373L^9 + 0,019L^{10} + 0,0978L^{11} + 0,2585L^{12} - 0,101L^{13} + 0,1744L^{14} - 0,4003L^{15} + 0,253L^{16} - 0,1597L^{17} + 0,1166L^{18} - 0,0387L^{19} - 0,1555L^{20} + 0,2901L^{21} - 0,1394L^{22} + 0,0818L^{23} - 0,1859L^{24} + 0,2466L^{25} - 0,3429L^{26} + 0,1492L^{27} - 0,0591L^{28} + 0,4379L^{29} - 0,5549L^{30} - 0,1444L^{31} + 0,1461L^{32} - 0,071L^{33} + 0,3321L^{34})a_t \quad (7)$$

RESULTADOS

Con los coeficientes ARIMA se procedió a realizar las predicciones. Para la serie *throughput* el nivel de predicción se graficó y/o proyectó para los siguientes 400 datos. Los resultados se muestran en la Fig. 10. Como se puede observar el nivel de predicción es bastante acertada; este hecho es corroborado cuantitativamente por los criterios estadísticos de la tabla 6.

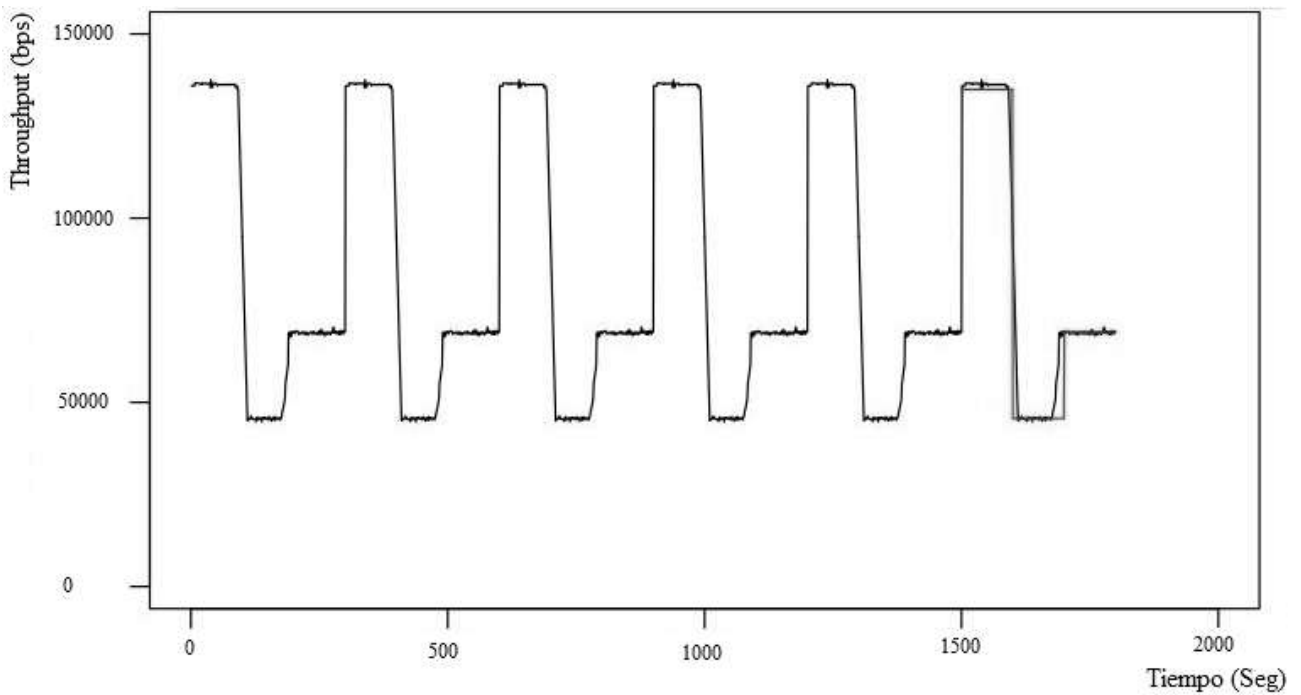


Fig. 10: Predicción del modelo ARIMA(4,1,4) (Gómez y Ramírez, 2013).

Tabla 6: Criterios estadísticos del pronóstico ARIMA.

ARIMA	Caso 1
Error absoluto	0,124972788%
Desviación estándar	0,140485556%
Coefficiente de correlación	0,999999992%

Los resultados gráficos de la estimación futura de la serie para la variable *perdida de paquetes* y los respectivos valores estadísticos se sintetizan en la Fig. 11 y la tabla 7:

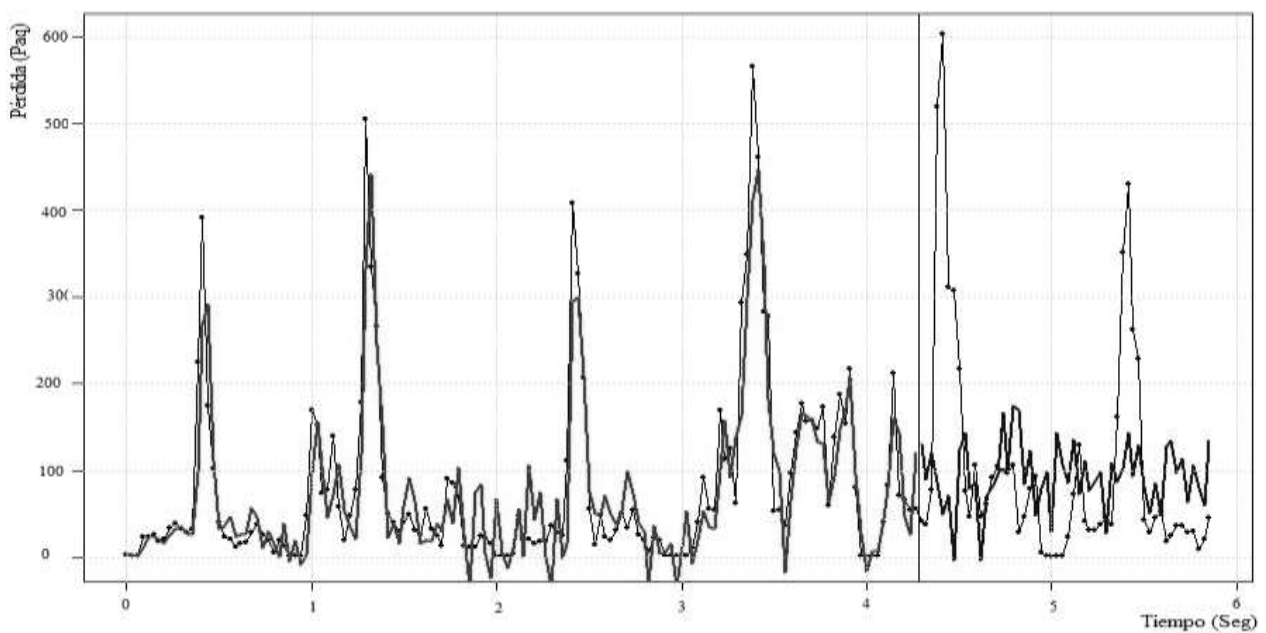


Fig. 11: Predicción del modelo ARIMA(3,1,34) (Martínez y López, 2013).

Tabla 7: Criterios estadísticos del pronóstico ARIMA.

ARIMA	Caso 2
Error absoluto	0,118571229%
Desviación estándar	4,137492547%
Coefficiente de correlación	0,989977967%

DISCUSIÓN

Es lógico plantear la hipótesis de que un modelo de predicción es útil solo cuando éste es capaz de anticiparse al comportamiento futuro real de la variable estudiada. Lo anterior implica, poder evaluar la exactitud de cada modelo, haciendo uso por ejemplo del método de evaluación ex-post (Hernández, 2006); metodología que fue utilizada en este artículo pronosticando 400 datos para la variable *perdida* y 50 para el *throughput*, los cuales fueron comparados con las secuencias originales para generar las tablas 6 y 7 que miden el nivel de pronóstico a partir del error absoluto, de la desviación estándar y del coeficiente de correlación. Analizando, por ejemplo el valor de la desviación estándar, se puede establecer que la capacidad de estimación del modelo ARIMA para cada caso es en su orden aproximadamente del 99 y 95 %. De esta manera queda demostrado que los modelos además de presentar una buena parsimonia, tiene una alta exactitud en sus mediciones a nivel de estimación de las variables de red analizadas.

CONCLUSIONES

A partir de los porcentajes de predicción encontrados (99% para el throughput y 95% para la variable perdida), se puede afirmar que el uso de modelos autoregresivos tipo ARIMA son apropiados para modelar las variables de red que inciden en el desempeño de las redes de telecomunicaciones ya que su nivel de estimación es muy cercano a la realidad.

A pesar de que paradigmas como ARIMA y sus variaciones, fueron diseñados para estudiar y evaluar modelos económicos; la aplicación de estos, en otros campos de la ciencia como la ingeniería de telecomunicaciones resultan ser prometedores, pues a partir de su nivel de pronóstico podría ser de mucha utilidad para robustecer los algoritmos que se encargan de enrutar y balancear el tráfico en las networking; aunque antes se hace necesario evaluar el consumo en cuanto a procesamiento para procesar grandes volúmenes de información con series temporales.

AGRADECIMIENTOS

Este proyecto fue realizado gracias a la financiación del Centro de Computo de Alto Desempeño de la Universidad Distrital "Francisco José de Caldas" en el marco del proyecto de investigación diseño e implementación de una plataforma experimental de televisión en una red IP (IPTV) para la Universidad Distrital. También se agradece al programa de Jóvenes Investigadores de Colciencias, y en general a todos los que se hicieron partícipes de este trabajo.

REFERENCIAS

- Brockwell, P., R. Davis, Time Series: Theory and Methods, New York: Springer (1986)
- Chen, T, Network traffic modeling, The Handbook of Computer Networks, 1st ed. vol. 3, Editorial: Wiley (2007)
- Correa, E, Series de tiempo: conceptos básicos, Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de matemáticas, Medellín (2004)
- Escrinche, S, Predicción de tráfico en redes IP. [En línea], consultado en Septiembre 21 de 2014: <http://upcommons.upc.edu/pfc/bitstream/2099.1/11278/1/memoria.pdf>
- García, N., N. Vera, Nelson., D. López, Análisis del Funcionamiento del Algoritmo de Balanceo de Carga LCM en Redes de Conmutación de Etiquetas Multiprotocolo Generalizado (GMPLS), doi: 10.4067/S0718-07642015000100005, Revista Información Tecnológica, (en línea), 26 (5), 41-54 (2015)

- Gómez, R., J. Ramírez, Generación de una propuesta de diseño para la implementación de servicios de televisión multicast IPv6, Trabajo de grado de pregrado, Universidad Distrital, Bogotá (2013)
- Haghighi, M., Y. Xiang., V. Varadharajan., B. Quinn, A stochastic time-domain model for burst data aggregation, Computers, IEEE Transactions on IEEE 802.15.4 Wireless Sensor Networks, 627-639 (2014)
- Hernández, C, Desarrollo de un modelo estadístico que permita estimar pronósticos futuros de tráfico en redes WiMax a través del modelamiento en series de tiempo Universidad Distrital "Francisco José de Caldas", Maestría en Teleinformática, Colombia (2006)
- Hirotsugu, A, A new look at the statistical model identification, IEEE Transactions on Automatic Control, 19(6), 716-723 (1974)
- López, D., N. García., P. Figueroa, Estimación y optimización del balanceo de carga usando ingeniería de tráfico adaptativa MATE en redes de conmutación de etiquetas multiprotocolo generalizado (GMPLS), doi: 10.4067/S0718-07642014000200006, Revista Información Tecnológica, (en línea), 25 (2), 47-56 (2014)
- Martínez, C., D. López, Modeling video transmission losses using SARIMA and ARIMA models, Revista Tecnura, 17 (37), 53-63 (2013)
- Min, H., X. Jianhui., X. Shiguo., Y. Fu-Liang, Prediction of chaotic time series based on the recurrent predictor neural network, IEEE Transactions on Signal Processing, 56(12), 3409-3416 (2004)
- Romani, L., A. de Ávila., D. Chino., J. Zullo, A new time series mining approach applied to multitemporal remote sensing imagery , IEEE Transactions on Geoscience and Remote Sensing, 51 (1), 140-150 (2013)
- Taj, M., M. Akil, Cognitive radio spectrum evolution prediction using artificial neural networks based multivariate time series modelling, Wireless Conference: Sustainable Wireless Technologies (European Wireless), 11th European, Austria (2011)
- Yanhua, Y., J. Wang., M. Song., J. Song, Network traffic prediction and result analysis based on seasonal ARIMA and correlation coefficient, International Conference Intelligent System Design and Engineering Application (ISDEA), IEEE, 980-983 (2010)