

# Modelamiento y pronósticos de tráfico correlacionado

Modeling and forecast of correlated traffic

CÉSAR AUGUSTO HERNÁNDEZ SUÁREZ

Ingeniero electrónico de la Universidad Distrital Francisco José de Caldas. Especialista Servicios Telemáticos e Interconexión de Redes. Magíster en Ciencias de la Información y las Comunicaciones de la Universidad Distrital Francisco José de Caldas y aspirante a Magíster en Economía de la Universidad de los Andes. Docente e investigador de la Universidad Distrital Francisco José de Caldas e integrante del grupo de investigación ARMOS e Internet Inteligente.  
lctsubasa@gmail.com

FREDY HERNÁN MARTÍNEZ SARMIENTO

Ingeniero electricista de la Universidad Nacional de Colombia. Especialista en Gestión de Proyectos de Ingeniería de la Universidad Distrital Francisco José de Caldas y aspirante a Doctor en Ingeniería Sistemas y Computación de la Universidad Nacional de Colombia. Docente e investigador de la Universidad Distrital Francisco José de Caldas y director del grupo de investigación ARMOS (Arquitecturas Modernas para Sistemas de Alimentación) y del proyecto de investigación *Evaluación de estructuras digitales para el diseño de sistemas de control difuso*.  
fhmartinezs@udistrital.edu.co

ANDRÉS ESCOBAR DÍAZ

Ingeniero electrónico de la Universidad Distrital Francisco José de Caldas. Magíster en Electrónica de la Universidad de los Andes. Docente de la Universidad Distrital “Francisco José de Caldas e integrante del grupo de investigación ORCA (Orden y Caos).  
andresed@gmail.com.

Fecha de recepción: 14 de septiembre de 2007

Clasificación del artículo: revisión

Fecha de aceptación: 21 de abril de 2008

**Palabras clave:** arima, autocorrelación, modelo de tráfico, series de tiempo.

**Key words:** arima, autocorrelation, time series, traffic models.

## RESUMEN

Hoy en día la necesidad de establecer una planeación y control eficiente en una red de datos se ha vuelto una tarea de vital importancia. Poder predecir los valores de tráfico de datos futuros es un factor

de gran relevancia en las tareas de planeación y control de redes. Por lo anterior este artículo pretende iniciar al lector en modelos estadísticos con series de tiempo, que permitan estimar pronósticos futuros de tráfico en las redes de comunicaciones

modernas, haciendo uso de la correlación existente entre valores de tráfico anteriores en una red de datos. Este modelamiento en series de tiempo está basado en medidas tomadas de eventos con una base periódica.

#### ABSTRACT

Today the need to establish an efficient planning and control in a data network has become a task

of vital importance. Forecast the values of future data traffic is a factor of great significance in the planning and control networks. For This paper intends to initiate the reader into statistical models with time series, to estimate the forecasting future of the traffic in modern communications networks, using the correlation between values traffic earlier in a data network. This modeling of time series is based on measures of events with a periodic base.

\* \* \*

## 1. Introducción

Las nuevas características en las capacidades y en las demandas de las actuales redes de datos empiezan a develar inconsistencias entre los modelos tradicionales –no correlacionados, como Poisson– y las medidas observadas, especialmente en lo referente a estructuras de correlación, se expanden a lo largo de diferentes escalas de tiempo, lo cual invalida los resultados tradiciones de la teoría de tráfico, las cuales se basaban en modelos no correlacionados, pues el nuevo tráfico sobre las redes es demasiado complejo para ser modelado mediante técnicas desarrolladas para la red telefónica [3].

Los modelos de tráfico no correlacionados –como el proceso de conteo de Poisson– ofrecen una gran tratabilidad matemática, pero no son capaces de capturar muchas de las características más relevantes del tráfico moderno, en particular las asociadas con la autocorrelación observada en el tráfico real [3]. Las series de tiempo tienen como objetivo central desarrollar modelos estadísticos que expliquen el comportamiento de una variable aleatoria que varía con el tiempo, lo que permit estimar pronósticos futuros de dicha variable aleatoria [21 y 35]. Por lo anterior dichos modelos estadísticos se postulan como una variante interesante en los modelos de tráfico para las redes de datos, ya que son beneficiosos para: la planeación de cobertura, reservación de recursos, monitoreo de la red, detección de anomalías y producción de modelos de simulación más exactos, en la medida en que

pueden pronosticar el tráfico en un tiempo de escala determinado [7, 22 y 41].

Desarrollar un modelo basado en series de tiempo no es algo para nada fácil, sin embargo, BOX y JENKINS lograron crear una metodología bastante práctica para estimar los modelos de series de tiempo; dicha metodología es la siguiente:

- Por medio de transformaciones o diferencias se estabiliza la varianza, y se eliminan la tendencia y la estacionalidad de la serie; así se obtiene una serie estacionaria.
- Para la serie estacionaria obtenida se identifica y se estima un modelo que explica la estructura de correlación de la serie con el tiempo.
- Al modelo hallado en la viñeta anterior se le aplican transformaciones inversas que permitan establecer la variabilidad, la tendencia y la estacionalidad de la serie original.
- El modelo estimado se valida a través de la correlación de sus residuales, si éstos llegan a presentar algún tipo de correlación entonces se vuelven a estimar nuevos parámetros, es decir, se regresa al modelo hallado en la segunda viñeta para realizar una nueva iteración. El modelo definitivo se usa para pronosticar [9, 11 y 15].

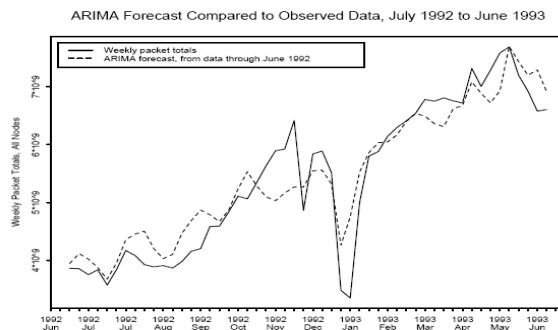
## 2. Estado del arte

Nancy K. Groschwitz y George C. Polyzos en su trabajo *A Time Series Model of Long-Term NSFNET Backbone Traffic* desarrollaron un modelo ARIMA

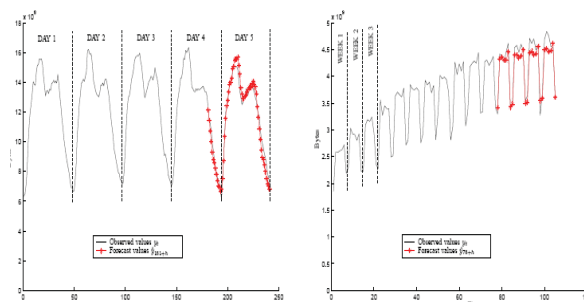
(Autorregresivos e Integrados con Promedios Móviles) para el tráfico del backbone de la NSFNET. Los datos de este experimento consistieron en todos los datos de tráfico de cada día de los nodos que pertenecen al backbone entre agosto 1 de 1988 y el 30 de junio del 1993; los datos fueron recolectados a través del protocolo de gestión de red SNMP. Los datos que se perdieron para algunos meses fueron estimados a través del mismo modelo de series de tiempo que se diseñó. Como la forma de poder modelar a través de Arima es garantizar que la serie sea estacionaria, se diferenció hasta lograrlo, para el modelo identificado (p,q,d) fue (2,2,1). El

modelo desarrollado fue capaz de predecir tráfico con resultados satisfactorios a un año completo (ver gráfica 1) [31].

Lionel Fillatre, en su trabajo *Forecasting Seasonal Traffic Flows*, desarrolló un modelo SARIMA (*Seasonal Auto Regressive Integrated Moving Average*) con base en las series de tiempo para pronosticar tráfico estacional; aquí, nuevamente, los datos se obtuvieron a partir del protocolo SNMP a intervalos de cinco minutos. Las constantes que identifican el modelo son (1,1,1)x(1,1,1), para los cuales el modelo se aproxima bastante al real (ver gráfica 2) [27].



Gráfica 1. Modelo ARIMA(2,2,1) desarrollado (Groschwitz, 2000)



Gráfica 2. Modelo SARIMA desarrollado (Marakov, 1998)

Chandrashekhar G. Dethe Y G. Wakde, en su trabajo *On the Prediction of Packet Process in Network Traffic Using FARIMA Time-Series Model*, desarrollaron un modelo FARIMA (p,d,q) (*Fractionally Integrated Autoregressive Moving Average Process*), aprovechando la dependencia de rango largo y la dependencia de rango corto que muestra el tráfico, al analizar la covarianza, la correlación y

la autocorrelación de los datos. El modelo Farima tiene la siguiente forma:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + Z_t + \sum_{i=1}^q \theta_i Z_{t-i} \quad (1)$$

Los datos en que se basó el modelo fueron tomados de las trazas de tráfico publicadas por Bellcore Net-

works [24]. Por su parte, Yiannis Kamarianakis, en su trabajo *Forecasting Traffic Flow Conditions In An Urban Network: Comparison Of Multivariate And Univariate Approaches*, desarrolló un modelo STARIMA (*single space-time ARIMA*), para el cual los resultados no son muy satisfactorios, debido a la dependencia del modelo con la velocidad de los datos de las muestras de tráfico [39].

María Papadopouli, Haipeng Sheng, Elias Raftopoulos, Manolis Ploumidis, Felix Hernández, en su investigación *Short-Term Traffic Forecasting in a Campus-Wide Wireless Network*, caracterizaron el tráfico de una infraestructura IEEE 802.11, lo que puede ser beneficioso en muchos dominios, incluyendo la planeación, el cubrimiento, la reserva de recursos, el monitoreo de la red y la detección de anomalías [3 y 46].

Este estudio se realizó con datos obtenidos en una escala horaria, es decir, los datos se tomaron cada hora, para distintos AP, los que manejaban mayor tráfico, ubicados en la universidad para los cuales se propusieron varios modelos de tráfico todos basados en una metodología de series de tiempo. La extracción de la serie de tiempo original contó con un total de 488 AP –que pertenecen a tres series distintas de Cisco– ubicados en un área de 729 acres. Este proceso se realizó a través del protocolo SNMP, las solicitudes se enviaron entre las 9:00 a.m. del 29 de septiembre del 2004 y las 12 a.m. del 30 de noviembre del mismo año para un total de 63 días, tomando muestras cada hora para un total de 1.471 datos tráfico de 1.512 posibles; los valores perdidos correspondieron a diversas situaciones como, por ejemplo, la pérdida de señal.

Para algunos de los AP la carga de tráfico mostró una fuerte periodicidad diaria, debido a esto el modelo desarrollado fue un ARIMA con característica de estacionalidad. Es interesante destacar que en todas las investigaciones consultadas, aunque pocas, coinciden en que el comportamiento del tráfico moderno se ajusta muy bien a una serie de tiempo ARIMA, ya sea para una red alámbrica o una inalámbrica. Esto puede ser la causa de la

autosemejanza detectada por muchos autores [12, 29 y 32].

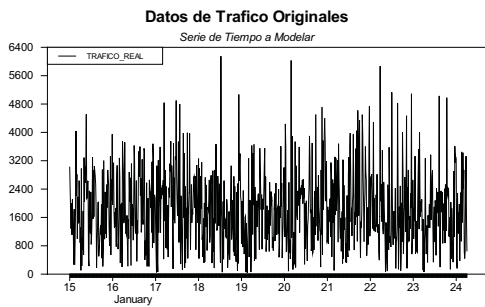
### 3. Desarrollando un modelo de tráfico

El desarrollo de este trabajo fundamenta su metodología en la de Box-Jenkins. El primer paso en el desarrollo de cualquier modelo de tráfico –y de un modelo matemático en general– es tomar una muestra de datos, con los se pueda caracterizar el tráfico de un tipo de red predeterminada [6 y 8]. En el presente estudio se decidió desarrollar un modelo de tráfico para una red con tecnología Wimax, puesto que esta tecnología es la que actualmente está tomando auge en el ámbito de las redes de datos, debido a todas sus características; sin embargo, como lo han confirmado varios estudios de tráfico en redes alámbricas e inalámbricas (como WiFi,) el tráfico actual como el de Internet e incluso el de video, presenta características fuertes de correlación [19, 20, 44 y 46]. Los datos fueron extraídos a través de la herramienta Netflow Analyzer.

Netflow Analyzer es una herramienta de software basada en Web que permite el monitoreo del ancho de banda de cualquier red. Esta herramienta es tan poderosa que es utilizada por otras herramientas como el Netflow de CISCO. Los datos son exportados de los dispositivos de Routing hacia el Analyzer Network Traffic el cual reporta el ancho de banda utilizado en tiempo real a través de la red monitoreada. Debido a su carácter de herramienta de software basada en Web, permite la monitorización remota de cualquier red a través de los dispositivos de red –como Routers– que se encuentren configurados en este software. A pesar de haberse extraído 889 (ver gráfica 3) datos de tráfico los modelos se desarrollarán únicamente con los 672 primeros (correspondientes a siete días), los datos restantes se utilizarán para comparar y evaluar los pronósticos de cada una de las series [27 y 45].

#### 3.1 Seleccionando el modelo adecuado

El concepto de estacionariedad es de suma importancia para el análisis de series de tiempo. En



Gráfica 3. Serie de tiempo del tráfico Wimax extraído.

general, para caracterizar completamente un proceso estocástico es necesario conocer la función de densidad conjunta de sus variables aleatorias; sin embargo, en la práctica no es realista pensar que esto se pueda lograr con una serie de tiempo. Como se mencionó anteriormente, en lo que respecta a la covarianza, no existe dependencia del tiempo, pero sí de la separación ( $k$ ) que hay entre las variables. Lo anterior conduce a pensar que la serie mostrará el mismo comportamiento en términos generales, sin importar el momento en el que se observe. Esto es, si se graficara un cierto número de observaciones contiguas de una serie, la gráfica que se obtendría sería bastante similar a la que se lograría al graficar el mismo número de observaciones contiguas, pero  $k$  periodos hacia delante o hacia atrás de los considerados inicialmente [13 y 34].

De la observación de la serie de tiempo que nos atañe, se puede concluir fácilmente que ésta no es estacionaria, debido a que su media no es cero, aunque su varianza parece constante, dos condiciones imperiosas para que una serie de tiempo sea estacionaria, lo cual se puede comprobar analíticamente a través de la prueba de raíz unitaria de Dickey-Fuller, cuyos resultados son:

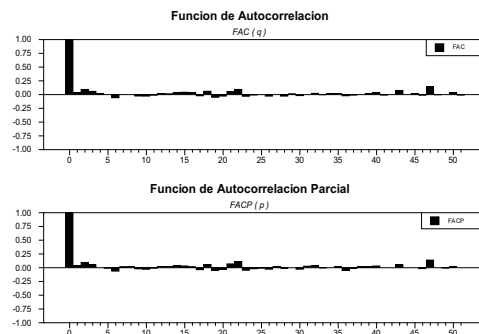
*Dickey-Fuller Unit Root Test, Series TRAFICO\_REAL*  
*Regression Run From 2007:01:15//62 to 2007:01:21//95*  
*Observations 611*  
*With intercept with 60 lags*  
*T-test statistic -1.49785*

*Critical values: 1%= -3.443 5%= -2.867 10%= -2.569*

Según el criterio de Dickey-Fuller, la serie no es estacionaria, ya que el valor absoluto del test es menor al valor absoluto del valor crítico del 5% [26]. De acuerdo con los resultados obtenidos es necesario transformar la serie para volverla estacionaria. La transformación se realizó por medio de las primeras diferencias [10 y 16]. El modelo seleccionado es entonces el ARIMA, el cual, a diferencia de otros como el AR, MA o ARMA, permite realizar transformaciones de la serie de tiempo para volverla estacionaria, con esto la etapa de identificación está terminada [14].

### 3.2. Estimando el modelo seleccionado

Ahora que ya se tiene una serie estacionaria se deben determinar el orden de “ $p$ ” (autoregresivo) y de “ $q$ ” (promedio móvil); para esto se utiliza la función de autocorrelación y la función de autocorrelación parcial [2, 4 y 23]. Estas gráficas de autocorrelación y autocorrelación parcial se obtuvieron con el software RATS y se muestran en la gráfica 4. Estas dos graficas (FAC y FACP) permiten estimar el valor de “ $p$ ” y “ $q$ ” para construir el modelo ARIMA ( $p,d,q$ ) que nos interesa. Entonces de la FAC se obtiene,  $q=22$ , y de la FACP se obtiene,  $p=22$ , como la serie se diferenció finalmente sólo una vez, entonces  $d=1$ . Es decir, que finalmente se tiene un modelo inicial ARIMA (22,1,22) [37].



Gráfica 4. FAC y FACP de la serie de tráfico Wimax.

Según los resultados de la FAC y la FACP se tiene el modelo presentado en la ecuación (2), para el cual aún no se conocen sus coeficientes.

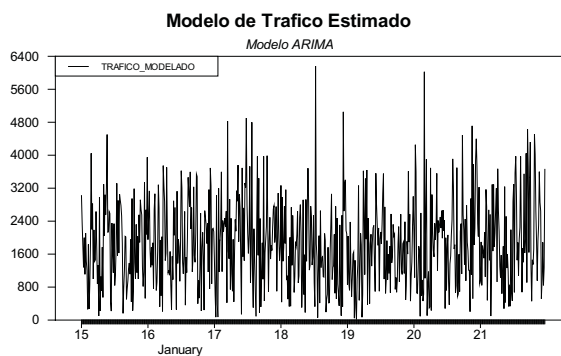
$$Z_t = \phi_1 \times Z_{t-1} + \phi_{12} \times Z_{t-12} + a_t - \theta_1 \times a_{t-1} - \theta_{12} \times a_{t-12} \quad (2)$$

Ahora se tiene un fuerte candidato, hay que estimarlo. En la práctica ésta es una labor de cómputo, aquí se debe escoger el paquete que se va a usar, por la flexibilidad que ofrece; para este estudio se eligió el software RATS (por encima inclusive del Software Eviews) por su gran potencialidad y estimación de máxima verosimilitud. Lo usual es pasar de la estimación inicial al análisis de los residuos; aquí se vuelven a buscar picos, pero ahora en los residuos. Estos picos revelan términos que uno debe incluir en la nueva formulación ARIMA que se volverá a estimar. Este ciclo de re-especificación dinámica termina cuando los residuos ya no presentan corre-

laciones (picos) y se puede decir que son residuos de ruido blanco. [11, 38 y 42].

El procedimiento anterior se lleva a cabo hasta que la FAC y FACP de los residuales demuestren que no existe correlación alguna entre los residuales del modelo estimado, lo cual se logró luego de tres iteraciones adicionales, sin embargo, el número de parámetros obtenidos para el modelo correspondiente fue 18. Un modelo con un número grande de parámetros es un modelo que no presenta una buena parsimonia, por lo que se analizó el nivel de significancia de cada parámetro eliminando aquéllos que presentan un valor mayor al 5%, ya que éstos no resultan significativos para el modelo. Una vez realizado lo anterior es necesario validar nuevamente el modelo y dependiendo del resultado iterar una vez más. [1], [28]. Finalmente, se llega a un modelo definitivo descrito por la ecuación (3) y cuyo comportamiento se muestra en la gráfica 5.

$$Z_t = -0.092467147 \times Z_{t-6} - 0.090197053 \times Z_{t-10} + 0.093022957 \times Z_{t-18} + a_t - 0.978862096 \times a_{t-1} - 0.009496113 \times a_{t-10} - 0.009945630 \times a_{t-18} \quad (3)$$



Gráfica 5. Serie de tráfico del modelo de la ecuación (3).

Como conclusión final se obtiene un modelo ARIMA (18,1,18) de seis parámetros, que está definido por la ecuación (3) y que por inspección visual de la gráficas 5, modela bastante bien los 672 datos de tráfico Wimax.

#### 4. Una aplicación del modelo desarrollado

El modelo desarrollado tiene múltiples aplicaciones, en la medida en que sus pronósticos sean precisos en un rango tolerable; para este caso se

aplicó el modelo desarrollado para encontrar, a partir de éste, un modelo que determine el retardo de los paquetes. Para encontrar un modelo matemático para el retardo promedio de los paquetes en el sistema, es necesario conocer la tasa de salida y de llegada de los paquetes en el sistema. El modelo de tráfico desarrollado a partir de series de tiempo determina el valor de la tasa instantánea de llegada de paquetes, por lo que sólo haría falta determinar la tasa de salida de los paquetes. Para obtener un modelo de retardo instantáneo se realiza un modelo en series de tiempo para los datos de tráfico de salida, correspondientes a los datos de tráfico de entrada con los que se modeló la ecuación (3), el resultado es la ecuación (4) [48 y 47].

$$Z_t = 0.1795532 \times Z_{t-9} + a_t - 0.0162852 \times a_{t-10} \quad (4)$$

Como los modelos para las tasas de llegada y salida de paquetes arrojan datos instantáneos (en bits por segundo) se puede calcular el tiempo de retardo instantáneo para un solo bit y asumir dicho resultado para todo el sistema con un grado de error significativamente bajo.

Denominando  $V_L(t)$  al modelo de tráfico de la ecuación (3), y correspondientemente  $V_S(t)$  al modelo de tráfico de la ecuación (4). Se tiene que el tiempo de llegada  $T_L$  para un único bit es:

$$T_s = \frac{1}{V_s(bps)} \times 1(bit) = \frac{1}{V_s} seg. \quad (5)$$

De la misma forma se tiene que el tiempo de salida  $T_s$  para un único bit es:

$$T_s = \frac{1}{V_s(bps)} \times 1(bit) = \frac{1}{V_s} seg. \quad (6)$$

Ahora se puede determinar el tiempo de retardo instantáneo  $T_R$  para un único bit como:

$$T_r = T_s - T_L = \frac{1}{V_s} - \frac{1}{V_L} \quad (7)$$

Sin embargo, hay que hacer una salvedad importante, si el valor del tiempo de salida es menor al valor del tiempo de llegada, el retardo no queda definido como la diferencia, ya que para ese instante

el retardo en teoría valdría cero. Digo en teoría, pues es posible que en ese momento aún existan paquetes en cola, por lo que los nuevos arribos tendrán que esperar un tiempo  $T_{R^*}$  determinado por el cociente del número de paquetes en cola  $N_R$  entre la diferencia de las velocidades instantáneas de salida y llegada, como se indica en la ecuación (8) [17 y 49].

$$T_{R^*} = \frac{N_R}{V_s - V_L} \quad (8)$$

Ahora el numero de paquetes  $N_{R^*}$  que experimentarían el retardo  $T_{R^*}$  está dado por la ecuación (9), siempre y cuando durante dicho tiempo la diferencia entre la velocidad de salida y la velocidad de llegada permanezca constante y mayor que cero.

$$N_{R^*} = V_L \times T_{R^*} \quad (9)$$

Ahora bien, si lo que se desea es determinar el retardo promedio  $E[T_R]$  en lugar del retardo instantáneo, entonces se toma el estadístico del valor esperado para un número “n” significativo datos de retardo instantáneo, es decir, se realiza el promedio de “n” valores adyacentes de retardo instantáneo.

$$\bar{T}_R = E[T_R] = \sum_{i=1}^n \frac{\left(\frac{1}{V_s}\right)_i - \left(\frac{1}{V_L}\right)_i}{n} \quad (10)$$

## 5. Análisis de resultados

Luego de obtener el modelo de tráfico se realizaron los análisis pertinentes para determinar la calidad del modelo de tráfico y su desempeño en el pronóstico de datos de tráfico futuros. La calidad de ajuste es uno de los factores de desempeño de un modelo y está definido como la suma de los cuadrados de los residuos dividida por el tamaño de la muestra, y su objetivo es medir la habilidad del modelo para reproducir los datos de la muestra, es decir, verifica qué tan parecida es la serie modelada con la real [33 y 25].

La similaridad entre la serie original y la desarrollada se podría evidenciar en una forma cualitativa (algo subjetiva) de la observación de las figuras correspondientes, de las que también se concluye

que el modelo ARIMA es un buen modelo de la serie de tiempo original. Sin embargo, debido a que se buscan criterios cuantitativos que determinen de forma objetiva el mejor modelo desarrollado, se calculó el criterio de calidad de ajuste para el modelo desarrollado el cual arrojó un valor de 171,413; no obstante, este dato no dice mucho por sí solo, si no se puede comparar con otro, además que sanciona severamente al modelo por al menos un dato atípico que se tenga en la serie de tráfico capturada [40 y 5].

En la realidad se puede sostener la tesis de que un modelo es realmente útil sólo en la medida en que anticipa la evolución de la variable. En este sentido se esta aceptando que se espere a que vengan las

futuras observaciones para después analizar la calidad del modelo. Esto se denomina una evaluación ex-post, y es una validación más fuerte en el sentido común, que el análisis de residuos [36 y 18].

Para el modelo desarrollado se pronosticaron 217 ( $217=889-672$ ) datos de tráfico que fueron respectivamente comparados con los datos de tráfico originales, para determinar la exactitud de los pronósticos en función de estadísticos como el error cuadrático medio, el promedio de la desviación estándar en valor absoluto, el coeficiente de correlación y la calidad de ajuste (ver tabla 1), lo cual permite observar detalladamente la eficacia del modelo desarrollado para pronosticar los respectivos datos de tráfico [50].

**Tabla 1.** Evaluación de pronósticos

Modelo Desarrollado	Modelo ARIMA
Error Cuadrático Medio	2,266523%
Desviación Estándar en Valor Absoluto	5,932770%
Coefficiente de Correlación	0,998136212
Calidad de Ajuste	4306,278315

Es realmente interesante analizar la capacidad de predicción que posee el modelo ARIMA ya que tan sólo tiene un 6% de error en promedio de los pronósticos de los datos de tráfico para los dos días siguientes, según la desviación estándar en valor absoluto. Pero ¿qué tanto tiempo a futuro podría mantener tan fantásticos pronósticos? [30 y 43].

## 6. Conclusiones

La ventaja de la determinación de parámetros de calidad de servicio a partir de modelos correlacionados como las series de tiempo y en especial el modelo ARIMA, se fundamenta para el caso de retardo promedio, en que la estructura de correlación presente en el tráfico, induce una estructura de correlación de características similares en los retardos.

No todos los modelos de series de tiempo caracterizan apropiadamente el tráfico Wimax, en especial para la serie de tráfico Wimax de este estudio, las series de tiempo como la de Promedios Móviles (MA) no logra capturar toda la dinámica de la serie; otros modelos, como el AutoRregresivo (AR), proporcionan realizaciones bastante parecidas en magnitud a las originales, pero con un número grande de parámetros que reduce significativamente su parsimonia y, por ende, aumenta el costo computacional; generalmente, las series de tráfico correlacionadas, como la que nos ocupa, exhibe uno o más parámetros de una componente autorregresiva y al mismo tiempo uno o más parámetros de una componente de promedio móviles, sugeridos por las funciones de autocorrelación y autocorrelación parcial, proponiendo como fuerte candidato un modelo ARIMA. La inclusión de ambas componentes tiene



por objetivo reducir significativamente el número de parámetros necesarios para modelar la serie original, tal y como se demostró en el desarrollo de dicho modelo. La evaluación realizada al modelo basado en series de tiempo implicó en su totalidad algún tipo de estadístico alrededor de los residuales (diferencia entre el valor real y el estimado por el modelo), para los cuales el modelo ARIMA obtuvo resultados satisfactorios además de un alto grado de desempeño, confiabilidad y exactitud en sus pronósticos.

Las series de tiempo, en especial los modelos Autorregresivos e Integrados con Promedios Móviles (ARIMA), resultan realmente apropiados para modelar el tráfico moderno con características de correlación fuertes como se demostró en la modelación del tráfico wimax. La evaluación del modelo ARIMA desarrollado y seleccionado finalmente como el más apropiado, exhibe un desempeño bastante alto en relación con la magnitud de sus residuales, los cuales no experimentaron correlación alguna.

---

## Referencias bibliográficas

---

- [1] Akaike, H. (1969) *Fitting Autoregressive Models for Prediction. Annals of the institute of statistical mathematics*. pp. 243-247.
- [2] Akaike, H. (1973) Information theory and an extension of the Maximum Likelihood Principle. Second International Symposium on Information Theory. Budapest, pp. 267-281.
- [3] Alzate, Marco Aurelio. (Junio de 2004) Modelos de tráfico en análisis y control de redes de comunicaciones. *Revista de Ingeniería*, 9 (1), 63-87.
- [4] Anderson, T.W. (1980) *Maximum Likelihood Estimation for Vector Autoregressive Moving-Average Models, Directions in Time Series*. Institute of mathematical statistics, p. 80-111.
- [5] Ansley, C.F. and Kohn, R. (1985) *On the Estimation of ARIMA Models with Missing Values, Time Series Analysis of Irregularly Observed Data*. Editorial Parzen.
- [6] Armstrong, J.S. (2001) *Principles of Forecasting*. Editorial Springer.
- [7] Arvidsson, Å. and Karlsson, P. (June, 1999) On traffic models for TCP/IP. Proc. of 16th international teletraffic congress (ITC). Edinburgh, pp. 457-466.
- [8] Beran, J. (1994) *Statistics for Long-Memory Processes*. Chapman & Hall.
- [9] Bertsekas, Dimitri and Gallager, Robert. (1987) *Data Networks* (2ª edición). New Jersey: Prentice Hall.
- [10] Box, G.E.P. and Cox, D.R. (1964) *An Analysis of Transformations*. Stat. soc..
- [11] Box, G.E.P. and Jenkins, Gwilym M. (1976) *Time Series Analysis: Forecasting and Control* (Revised Edition). Oakland, California: Editorial Holden-Day.
- [12] Box, G.E.P. and Pierce, D.A. (Sin fecha) *Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving-Average Time Series Models*. Stat. assoc.
- [13] Brillinger, D.R. (2001) *Time Series: Data Analysis and Theory*. Universidad de California. Holden-Day. SIAM.
- [14] Brockwell, P.J. (2001) On Continuous-Time ARMA Processes. En: *Handbook of statistics*, 19, 249-276.
- [15] Brockwell, P.J. and Davis, R.A. (2002) *Introduction to Time Series and Forecasting* (2nd edition). New York: Editorial Springer.
- [16] Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods*. New York: Springer Verlag.
- [17] Camerano Fuentes, Rafael. (1997) *Teoría de colas*. Bogota: Fondo de publicaciones Universidad Distrital Francisco José de Caldas.
- [18] Campos, Manuel. (2006) *Estadística actuarial III. Campus de colmenarejo. Departamento de estadística*. Madrid: Universidad Carlos III de Madrid.
- [19] Casilari, E.; Reyes, A.; Lecuona, A.; Díaz Estrella, A. y Sandoval, F. (2002) *Caracterización de tráfico de video y tráfico Internet*. Malaga: Universidad de Malaga, Campus de Teatinos.
- [20] Casilari, E.; Reyes, A.; Lecuona, A.; Díaz Estrella, A. y Sandoval, F. (2003) *Modelado de tráfico telemático*. Malaga: Departamento de Tecnología electrónica, E.T.S.I. telecomunicación. Universidad de Málaga, Campus de Teatinos.
- [21] Correa Moreno, Emilia. (2004) *Series de tiempo: conceptos básicos*. Medellín: Universidad Nacional de Colombia, Facultad de Ciencias, Departamento de matemáticas.
- [22] Couch, L. (2001) *Digital and Analog Communication System*. New Jersey: Prentice Hall.
- [23] Davis, R.A. (1996) Maximum Likelihood Estimation for MA(1) Processes with a Root on or Near the Unit Circle. *Econometric Theory*, 12, 1-29
- [24] Dethe, Chandrashekar y Wakde D.G. (2003) On the Prediction of Packet Process in Network Traffic Using FARIMA Time Series Model. India: Department of Electronics, College of Engineering.
- [25] Devore, Jay L. (2001) *Probabilidad y estadística para ingeniería y ciencias* (5ª edición). México: Editorial Thomson.
- [26] Dickey, D.A. and Fuller, W.A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *J. Amer. stat. assoc.*, 74, 427-431.
- [27] Fillatre, Lionel; Marakov, Dimitry and Vatou, Sandrine. (2003) *Forecasting Seasonal Traffic Flows*. Paris: Computer Science Department, ENST Bretagne.
- [28] Fuller, W.A. (1996) Introduction to Statistical Time Series. A Survey. *Int. stat. Review*, 53, 301-329.

- [29] Ghaderi, M. (2003) *On the Relevance of Self-Similarity in Network Traffic Prediction*. Tech. Rep. School of Computer Science. Waterloo: University of Waterloo.
- [30] Granger C., W. J. (1981) Some Properties of Time Series Data and their Use in Econometric Model Specification. *J. Econometrics*, 16, 121-130.
- [31] Groschwitz, Nancy K. and Polyzos, George C. (2005) *A Time Series Model Of Long-Term NSFNET Backbone Traffic*. San Diego: Computer Systems Laboratory, Department of Computer Science and Engineering, University of California.
- [32] Grossglausser, M. y Bolot, J.C. (1999) On the Relevance of Long-Range Dependence in Network Traffic Source. *IEEE/ACM Trans. Networking*, 7.
- [33] Guerrero Guzmán, Víctor Manuel. (2003) *Análisis estadístico de series de tiempo económicas* (2ª edición). México: Editorial Thomson.
- [34] Hamilton, James D. (1994) *Time Series Analysis*. New Jersey: Princeton University Press.
- [35] Harmantzis, F.C. and Hatzinakos, D. (2005) *Heavy Network Traffic Modeling and Simulation Using Stable FARIMA Processes*. Hooben: Stevens institute of technology, Castle point on the Hudson.
- [36] Harvey, A.C. (1993) *Time Series Models*. Harvester Wheatsheaf.
- [37] Jones, R.H. (1975) Fitting Autoregressions. *J. amer. Stat. assoc.*, 70, 590-592.
- [38] Jones, R.H. (1978) *Multivariate Autoregression Estimation Using Residuals, Applied Time Series Analysis*. New York: Academic Press.
- [39] Kamarianakis, Yiannis and Prastacos, Poulicos. (2003) Forecasting Traffic Flow Conditions in an Urban Network: Comparison of Multivariate and Univariate Approaches. 82nd Transportation Research Board Annual Convention. Paper number 03-4318.
- [40] Leland W.E.; Taquu, M.S.; Willinger, W. and Wilson, D.V. (1994) On the Self-Similar Nature of Ethernet Traffic. *IEEE/ACM Trans. Networking*, 2.
- [41] López Ardao, José Carlos. (2004) *Contribución al análisis del impacto de la correlación en las prestaciones de las redes de alta velocidad*.
- [42] Makridakis, Spyros G.; Wheelwright, Steven C. and Hyndman, Rob J. (1997) *Forecasting: Methods and Applications* (3ª edición). USA: Editorial Wiley.
- [43] Montgomery Douglas, C.; Peck, Elizabeth A. y Vining, G. (2002) Geoffrey. *Introducción al análisis de regresión lineal* (3ª edición). Editorial Continental.
- [44] Olexa, Ron. (2004) *Implementing 802.11, 802.16, and 802.20 Wireless Networks: Planning, Troubleshooting, and Operations*. Editorial Newnes.
- [45] Pajouh, Danech. *Methodology for traffic forecasting*. Arcuel: The French National Institute for Transport and Safety Research (INRETS).
- [46] Papadopouli, María; Sheng, Haipeng; Raftopoulos, Elias; Ploumidis, Manolis and Hernández, Felix. (2004) *Short-Term Traffic Forecasting in a Campus-Wide Wireles Network*.
- [47] Schwartz, Misha. (1994) *Redes de telecomunicaciones, protocolos, modelado y análisis*. Editorial Addison Wesley Iberoamericana.
- [48] Tanenbaum, Andrew S. (2003) *Redes de computadoras* (4ª edición). México: Prentice Hall
- [49] Willinger W.; Paxon V. y Taquu M.S. Self-Similarity and Heavy Tails: Structural Modeling Of Network Traffic. En: *A Practical Guide to Heavy Tails: Statistical Techniques and Applications* (pp. 27-53). Bostón: Birkhauser.
- [50] Yu, Guoqiang and Zhang, Changshui. (2003) *Switching ARIMA Model Based Forecasting for Traffic Flow*. State Key Laboratory of Intelligent Technology and System. Department of Automation, Tsinghua University.



Anónimo. La carrera 6 entre la calle 14 y la Avenida Jiménez de Quesada. fotografía en B/N