

# A Study on the Short Term Internet Traffic Forecasting Models on Long-Memory and Heteroscedasticity

H. G. Sohn<sup>a</sup> · S. Kim<sup>a,1</sup>

<sup>a</sup>Department of Applied Statistics, Chung-Ang University

(Received November 26, 2013; Revised December 24, 2013; Accepted December 24, 2013)

---

## Abstract

In this paper, we propose the time series forecasting models for internet traffic with long memory and heteroscedasticity. To control and forecast traffic volume, we first introduce the traffic forecasting models which are determined by the volatility and heteroscedasticity of the traffic. We then analyze and predict the heteroscedasticity and the long memory properties for forecasting traffic volume. Depending on the characteristics of the traffic, Fractional ARIMA model, Fractional ARIMA-GARCH model are applied and compared with the MAPE(Mean Absolute Percentage Error) Criterion.

Keywords: Fractional ARIMA, Fractional ARIMA-GARCH, internet traffic, forecasting, Bps.

---

## 1. 서론

IT기술의 발전에 따라 유무선 통신기술은 국내산업의 중추신경과 같은 존재로 관련 산업들의 발전에 많은 기여를 하고 있다. 인터넷 보급에 따라 인터넷 사용자 수와 트래픽 사용량은 급격히 증가 하였으며, 최근들어 스마트 기기나 P2P 응용으로 인하여 트래픽 양은 급격히 증가하였으며, 이에 따라 통신망의 안정을 위하여 네트워크의 효율적 관리의 중요성이 부각되고 있는 실정이다.

과거 포트기반 방식의 단순한 트래픽에서 현재는 다양한 방식의 트래픽이 발생되고 있어 탐지가 복잡하고 어려워졌다. 이에 따라 실시간으로 정확한 트래픽 탐지 방법의 개발이 필요하다. 실시간 예측을 위하여 시계열 모형을 적용한 예측이 필요하며, 이를 통하여 조기 이상침입 탐지 및 동적인 대역폭 공급이 가능하다. 인터넷 트래픽의 특성을 파악하고 예측하기 위하여 재무 시계열 자료의 예측을 위한 방법론을 참조하는 것은 좋은 예가 될 것이다. Baillie 등 (1996)은 재무 시계열 자료의 예측을 위하여 FARIMA(Fractional Autoregressive Integrated Moving Average)-GARCH 모형을 사용하였다. 한편 Basu 등 (1996)은 인터넷 트래픽 예측을 위하여 처음으로 시계열 모형을 도입하였다.

Liu 등 (1999)은 인터넷 트래픽의 장기기억 특성을 고려한 FARIMA 모형을 이용하여 트래픽을 예측하였다. Shu 등 (2005)은 계절형 ARIMA를 이용하여 무선 트래픽을 예측하였다.

---

<sup>1</sup>This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2013-008277).

<sup>1</sup>Corresponding author: Professor, Department of Applied Statistics, Chung-Ang University, Dongack-Gu, Seoul 156-756, Korea. E-mail: sahm@cau.ac.kr

트래픽 예측의 시계열을 이용한 최근 연구를 살펴보면 Jiang과 Papavassiliou (2006)는 주파수 도메인에서 트래픽 데이터를 분석하고 예측하는 새로운 방법론을 처음으로 도입하였다. 이 분할 전략의 결과로서, 트래픽은 기본 요소와 단기 요소란 두가지 주요 부분으로 나누어서 ARIMA 모형을 기반으로 ERAN 알고리즘을 이용하여 예측을 하였다. Tikunov와 Nishimura (2007)는 모바일 트래픽의 예측을 위하여 Holt-Winters 방법을 적용하였다.

Kim (2007)은 Engle (1982)이 제안한 자기회귀이분산성(ARCH; Autoregressive Conditional Heteroskedasticity) 모형을 이용하여 트래픽자료를 예측하는데 있어 AR 모형보다 성능이 우수함을 보였고, Kim (2011)은 시간별 자료를 통하여 Seasonal AR-GARCH 모형이 Seasonal ARIMA와 AR-GARCH 모형보다 트래픽 자료를 예측하는데 성능이 우수함을 보였다. 본 연구에서는 5분 간격으로 측정되는 단기 트래픽 자료를 예측하는데 있어 장기 기억특성을 고려한 수정된 Fractional ARIMA 방법과 분산의 이분산성을 설명할 수 있는 AR-GARCH와 ARMA-GARCH 모형을 이용하며, Fractional ARIMA 방법과 GARCH 모형을 결합한 Fractional ARIMA-GARCH 모형을 이용하여 트래픽의 예측 성능을 비교하고자 한다.

## 2. 시계열 모형의 소개

### 2.1. ARMA-GARCH 모형

환율이나 주식과 같이 변동성이 큰 시계열 자료에서는 이분산성이 많이 발생하는데, Engle (1982)과 Bollerslev (1986)는 이분산 문제점을 다루는 방법으로 ARCH와 GARCH 모형을 제안하였다. 일반적으로 ARMA( $p, q$ )-GARCH( $p', q'$ ) 모형은 다음과 같은 형태로 제안된다.

$$\begin{aligned}\phi_p(B)Z_t &= \theta_q(B)\epsilon_t, \\ \epsilon_t &= e_t\sqrt{h_t}, \\ h_t &= \alpha_0 + \sum_{j=1}^{p'} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{q'} \beta_j h_{t-j},\end{aligned}\quad (2.1)$$

여기서  $\phi_p(B)$ 와  $\theta_q(B)$ 는  $p$ 차 AR 작용소와  $q$ 차 MA 작용소를 의미하며,  $e_t \sim \text{iid } N(0, \sigma^2)$ 이다. 만일 GARCH(1, 1) 모형이라면,  $\alpha_0 > 0$ ,  $\alpha_1 \geq 0$ ,  $\beta_1 \geq 0$ ,  $\alpha_1 + \beta_1 < 1$ 을 따른다.  $p'$ 는 ARCH 부분의 차수이고,  $q'$ 는 GARCH 부분의 차수이다.

### 2.2. Fractional ARIMA(FARIMA) 모형

정상시계열 자료  $Z_t$ 의  $k$ 시차의 ACF  $\rho_k$ 의 절대합이 유한일 경우 시계열 자료  $Z_t$ 는 단기 기억(short memory)를 갖는다고 하며, 식 (2.2)와 같이 표현가능하다.

$$\sum_{k=0}^{\infty} |\rho_k| < \infty. \quad (2.2)$$

만약 시계열 자료  $Z_t$ 의  $k$ 시차의 ACF  $\rho_k$ 의 절대합이 유한하지 않으면 시계열 자료  $Z_t$ 는 장기 기억(long memory)을 가지며, 식 (2.3)과 같이 표현가능하다.

$$\sum_{k=0}^{\infty} |\rho_k| = \infty. \quad (2.3)$$

Liu 등 (1999)은 인터넷 트래픽 자료의 장기기억 특성을 보여주었다.

Fractional ARIMA(FARIMA)는 장기기억 특성을 고려한 모형으로서 표준 ARIMA( $p, d, q$ ) 모형의 일종으로  $d$ 가 자연수가 아닌 실수를 가지는 모형으로 FARIMA( $p, d, q$ )에 대해 시계열  $Z_t$ 는 식 (2.4)와 같이 정의된다.

$$\phi_p(B)\Delta^d Z_t = \theta_q(B)\epsilon_t, \quad (2.4)$$

여기서  $\epsilon_t$ 는 백색잡음이며 모형의 정상성을 위해  $d \in (-0.5, 0.5)$ 이며,  $\delta^d$ 는 분수차분연산자로 식 (2.5)와 같이 정의된다.

$$\Delta^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k, \quad (2.5)$$

여기서  $\binom{d}{k} = d!/(d-k)!k!$ 으로 FARIMA 모형은 일반적인 ARMA 모형이 된다. FARIMA 모형은 Liu 등 (1999)이 트래픽 자료의 장기기억 특성을 고려한 모형으로 제안되었다.

### 2.3. Fractional ARIMA-GARCH 모형

Fractional ARIMA-GARCH 모형은 Fractional ARIMA 모형과 GARCH 모형의 결합된 형태로 식 (2.6)의 Fractional ARIMA( $p, d, q$ ) 모형의 분산에 대하여 식 (2.7)에 GARCH( $p', q'$ )가 결합된 형태이다.

$$\begin{aligned} \phi_p(B)\Delta^d Z_t &= \theta_q(B)\epsilon_t, \\ \Delta^d &= (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k, \end{aligned} \quad (2.6)$$

$$\epsilon_t = e_t \sqrt{h_t},$$

$$h_t = \alpha_0 + \sum_{j=1}^{p'} \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^{q'} \beta_j h_{t-j}. \quad (2.7)$$

이는 재무시계열 자료에서 Baillie 등 (1996)이 Fractional 차분과 ARMA-GARCH 모형을 결합한 것으로, 인터넷 트래픽이 장기 기억 특성과 이분산성 동시에 가짐에 따라 적용 가능성이 높다고 할 수 있다.

## 3. 인터넷 트래픽 예측 성능 비교

본 연구에서 사용한 자료는 2013년 5월 1일부터 27일까지 4주간, 외부에서 중앙대학교 서버로 들어온 트래픽 Bps(Bit per second)와 교내 서버로부터 외부로 보내지는 트래픽 Bps를 이용하여 분석을 하였다. Bps는 초당 비트수로 1초 동안에 몇 개의 비트를 전송할 수 있는가를 나타내는 단위로서 1초간에 1비트를 전송하면 1bps로 표시된다.

본 연구에서는 5분 간격으로 측정된 평균 Bps를 이용하였으며, 앞에 3주간의 6048개의 자료를 이용하여 모형을 적합하고, 이후 6일 동안의 1728개에 대하여 예측하였다. 예측력을 비교하기 위하여 1~12까지의 다단계 예측 실시하여 MAPE에 따라서 모형의 성능을 살펴보았다.

자료에 분산의 안정화를 위하여 자연로그변환을 실시하며, 이 자료에 대하여 모형 적합과 정상성을 만족시키기 위하여 차분을 실시한다. 이에 따라 식 (3.1)과 같이 로그 차분을 통한 변수변환을 실시하여 모

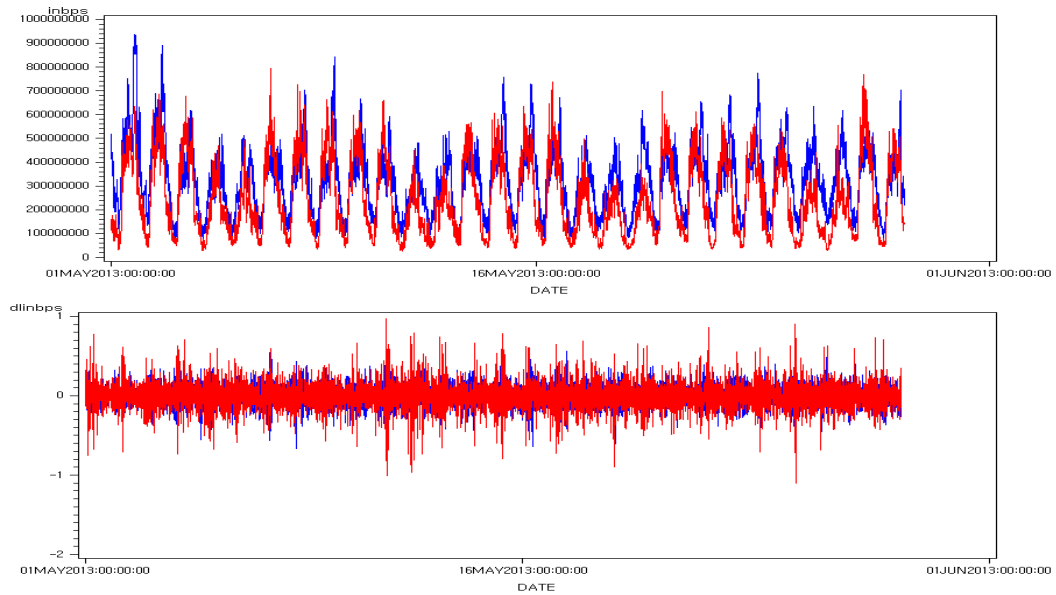


Figure 3.1. Time plot for original data and log differenced data

Table 3.1. LM test

	Order	LM	Pr>LM	LM	Pr>LM	
In Bps	1	204.7577	<.0001	Out Bps	454.0231	<.0001
	2	228.7081	<.0001		502.009	<.0001
	3	247.5398	<.0001		529.7845	<.0001
	4	268.9593	<.0001		537.3998	<.0001
	5	309.3175	<.0001		547.2179	<.0001
	6	316.9673	<.0001		559.5853	<.0001
	7	325.1313	<.0001		587.1071	<.0001
	8	339.6123	<.0001		594.7143	<.0001
	9	351.4673	<.0001		611.3511	<.0001
	10	360.4408	<.0001		616.7781	<.0001
	11	373.5931	<.0001		625.2178	<.0001
	12	380.4327	<.0001		627.9829	<.0001

수를 추정하였다.

$$Y_t = \log \left( \frac{Z_t}{Z_{t-1}} \right). \tag{3.1}$$

Figure 3.1에서 보면 원자료와 로그차분을 한 자료의 시계열도표를 나타낸 것이다. 원자료의 시계열 도표가 일정한 주기를 가지고 같은 패턴이 반복 되고 있으나, 로그 차분을 한 자료의 시계열도표를 살펴보면 주기적 패턴이 사라짐을 알 수 있다. 이는 인터넷 트래픽 자료가 매우 측정주기를 가진 데이터로서 전 시점 대비 증감을 나타내는 로그차분을 한 자료에서는 계절주기적 영향이 사라짐을 알 수 있었다.

본 연구에서는 로그차분을 통하여 정상성을 만족시킨 자료에 AR-GARCH 모형과 ARMA-GARCH 모

**Table 3.2.** Parameter estimates of AR-GARCH

In Bps: AR(2)-GARCH(1, 2)					In Bps : AR(2)-GARCH(1, 2)				
	Estimate	Standard error	<i>t</i> -value	<i>p</i> -value		Estimate	Standard error	<i>t</i> -value	<i>p</i> -value
$\phi_1$	-0.2772	0.0133	-20.7802	<.0001	$\phi_1$	-0.1928	0.0143	-13.5035	<.0001
$\phi_2$	-0.1815	0.0131	-13.9011	<.0001	$\phi_2$	-0.1545	0.0135	-11.4561	<.0001
$\alpha_0$	5.7832	1.4478	3.9946	<.0001	$\alpha_0$	16.5162	2.5116	6.5761	<.0001
$\alpha_1$	0.0668	0.0120	5.5563	<.0001	$\alpha_1$	0.1543	0.0150	10.2760	<.0001
$\beta_1$	0.6219	0.1965	3.1651	0.0015	$\beta_1$	0.4557	0.0843	5.4033	<.0001
$\beta_2$	0.2706	0.1833	1.4764	0.1398	$\beta_2$	0.3269	0.0762	4.2878	<.0001

**Table 3.3.** Parameter estimates of ARMA-GARCH

In Bps: ARMA(2, 1)-GARCH(1, 2)					In Bps : ARMA(2, 1)-GARCH(1, 2)				
	Estimate	Standard error	<i>t</i> -value	<i>p</i> -value		Estimate	Standard error	<i>t</i> -value	<i>p</i> -value
$\phi_1$	0.2670	0.0352	7.5905	<.0001	$\phi_1$	0.3729	0.0415	8.9910	<.0001
$\phi_2$	-0.0593	0.0180	-3.2847	0.0010	$\phi_2$	-0.0639	0.0181	-3.5237	0.0004
$\theta_1$	-0.5764	0.0330	-17.4552	<.0001	$\theta_1$	-0.5895	0.0397	-14.8652	<.0001
$\alpha_0$	5.9810	1.4898	4.0146	<.0001	$\alpha_0$	15.4060	2.3863	6.4562	<.0001
$\alpha_1$	0.0667	0.0117	5.7095	<.0001	$\alpha_1$	0.1537	0.0149	10.2844	<.0001
$\beta_1$	0.5833	0.1860	3.1353	0.0017	$\beta_1$	0.4634	0.0813	5.7015	<.0001
$\beta_2$	0.3071	0.1738	1.7664	0.0773	$\beta_2$	0.3240	0.0735	4.4109	<.0001

**Table 3.4.** Parameter estimates of FARIMA

In Bps: FARIMA(2, <i>d</i> , 1)					In Bps : FARIMA(1, <i>d</i> , 1)				
	Estimate	Standard error	<i>t</i> -value	<i>p</i> -value		Estimate	Standard error	<i>t</i> -value	<i>p</i> -value
<i>d</i>	0.1555				<i>d</i>	0.1634			
$\phi_1$	0.2450	0.0673	3.6399	0.0001	$\phi_1$	0.3409	0.0219	15.5826	<.0001
$\phi_2$	-0.0420	0.0421	-0.9950	0.0100	$\theta_1$	-0.7519	0.0938	-49.0240	<.0001
$\theta_1$	-0.7273	0.0938	-7.8239	<.0001					

형, Fractional ARIMA(FARIMA) 모형과 FARIMA-GARCH 모형을 적용시킨다. 먼저 이분산성 모형에서 이분산성이 존재하는지에 대해 LM test를 실시하며, Table 3.1과 같이 나타난다. LM test 결과 이분산성이 존재하는 것으로 나타났다. 이에 대하여 GARCH 모형을 고려하였으며, 식 (3.1)과 같이 로그 차분한 자료를 이용하여 모형 선택 기준은 BIC를 기준으로 선택하였다.

예측된 모형은 외부에서 교내 서버로 들어온 Bps인 In Bps와 교내에서 외부로 나간 Bps에 대한 Out Bps에 대하여 AR-GARCH 모형은 두 자료 모두 AR(2)-GARCH(1, 2)가 선택되었으며, ARMA-GARCH 모형 또한 두 자료 모두 ARMA(2, 1)-GARCH(1, 2)가 선택되었다. 장기기억특성을 고려한 Fractional ARIMA(FARIMA) 모형은 In Bps에서는 FARIMA(2, *d*, 1)이 선택되었으며, Out Bps에서는 FARIMA(1, *d*, 1)이 선택되었다. 장기기억특성과 이분산성을 동시에 고려한 FARIMA-GARCH 모형은 In Bps에서는 FARIMA(2, *d*, 1)-GARCH(1, 1)이 선택되었으며, Out Bps에서는 FARIMA(1, *d*, 1)-GARCH(1, 1)가 선택되었다. 선택된 모형에 대한 모수 추정값은 Table 3.2에서 Table 3.5와 같다.

추정된 모형의 모수를 이용하여 모형의 성능을 판단하기 위해 6일치 자료 1728개의 자료에 대하여 1~12 step ahead 예측을 하였다. 예측에 대한 모형의 성능은 식 (3.2)와 같은 MAPE(Mean Absolute Percentage Error)를 이용하여 비교하였다. 보통 RMSE를 사용하지만 여기서는 예측의 상대적 크

**Table 3.5.** Parameter estimates of FARIMA-GARCH

In Bps: FARIMA(2, d, 1)-GARCH(1, 1)					In Bps : FARIMA(1, d, 1)-GARCH(1, 1)				
	Estimate	Standard error	t-value	p-value		Estimate	Standard error	t-value	p-value
$d$	0.1555				$d$	0.1634			
$\phi_1$	0.2496	0.02258	11.058	<.0001	$\phi_1$	0.2572	0.0496	5.1821	<.0001
$\phi_2$	-0.0417	0.0169	-2.4655	0.0136	$\theta_1$	-0.7435	0.0337	-22.0725	<.0001
$\theta_1$	-0.7184	0.0185	-38.824	<.0001	$\alpha_0$	0.0003	0.0002	1.8605	0.0628
$\alpha_0$	0.0005	0.0001	4.126	<.0001	$\alpha_1$	0.0409	0.014	2.9288	0.0034
$\alpha_1$	0.0537	0.0078	6.853	<.0001	$\beta_1$	0.9322	0.0236	39.5313	<.0001
$\beta_1$	0.9095	0.0147	61.4977	<.0001					

기를 비교하기 위해서 MAPE를 사용하였다.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100\% \quad (3.2)$$

여기서  $n$ 은 예측을 위해 사용된 표본수이며,  $Y_t$ 는  $t$ 시점에서의 실제값,  $\hat{Y}_t$ 는  $t$ 시점에 대한 예측값이다. 각 모형에 대해 1부터 12시차까지의 Multi-step 예측을 사용하여 5분 후의 예측부터 1시간 후의 예측을 실시하였다. 식 (3.2)를 이용하여 산출한 MAPE는 In Bps에 경우는 Table 3.6과 같고, Out Bps에 경우는 Table 3.7과 같다.

Table 3.6에 외부에서 교내로 들어오는 BPS에 대한 결과를 살펴 보면 ARMA-GARCH 모형과 FARIMA-GARCH 모형이 타 모형에 비하여 작은 MAPE를 보여주어 우수한 예측 결과를 보여준다. 특히 5분 후의 예측인 1 ahead Forecasting인 5분 후 예측부터 1시간 후에 대한 예측인 12 시차 미래예측 모두에서 타 모형에 비하여 작은 값으로 더 좋은 예측을 보여주고 있다. 이분산성과 장기기억특성을 고려한 FARIMA-GARCH 모형이 타 모형에 비하여 더 나은 효과를 보여주고 있다. FARIMA 모형에 경우는 이분산을 고려한 모형들 보다 예측력이 낮게 나타나는데 이에 따라 단기 트래픽 예측시 이분산을 고려하여야 할 것으로 보여진다. 교내 서버에서 외부로 나가는 트래픽인 Out Bps에 경우 장기기억특성을 고려한 FARIMA와 FARIMA-GARCH 모형이 전반적으로 타 모형들에 비하여 우수한 성향을 보여주고 있다. 이는 In Bps에 비하여 Out Bps가 이분산성에 의한 영향보다 장기기억특성의 영향을 더 받는 것으로 보여진다.

#### 4. 결론

본 연구는 인터넷 트래픽 양에 대한 실제 자료에 대하여 장기기억 특성과 이분산성을 고려한 모형을 통한 성능비교를 실시하였다. 현재 실제로 측정된 자료는 거의 없는 실정이나 본 연구를 위하여 최근 3년간의 교내의 신축 트래픽 자료를 수집하였다. 본 연구를 위하여 먼저 교내 서버로 들어온 트래픽인 In Bps와 교내서버에서 외부로 나간 트래픽인 Out Bps에 대하여 분석 및 예측을 실시하였다. In Bps에 경우 장기기억특성과 이분산성을 동시에 고려한 FARIMA-GARCH 모형이 가장 좋은 예측 정확도를 보이며 이분산성을 고려한 모형들이 우수한 예측 정확도를 보였다. 교내에서 외부로 나간 트래픽인 Out Bps에서도 FARIMA-GARCH 모형이 우수한 예측력을 보였으며, 이분산성을 고려한 AR-GARCH와 ARMA-GARCH 모형보다는 장기기억 특성이 고려된 FARIMA 모형이 우수한 결과를 보여주고 있다. 이에 따라 트래픽 예측시 들어오는 트래픽 양과 외부로 나가는 트래픽 양에 대하여 특성에 따라서 다른 모형의 적용이 필요함이 보여진다. 특히 학교와 같이 공공기관과 같은 곳으로 들어오는 트래픽의 경우는 장기기억 특성과 이분산성을 고려한 모형에 적용이 필요할 것으로 보인다. 본 연구에서

Table 3.6. MAPE for In Bps(%)

	5분후	10분후	15분후	20분후	25분후	30분후	35분후	40분후	45분후	50분후	55분후	60분후
AR-GARCH	9.4686	12.0174	13.5153	14.7666	15.3394	16.5716	17.7462	18.9556	19.9803	20.0208	20.8360	21.0712
ARMA-GARCH	9.4529	11.9127	13.5298	14.6071	15.2982	16.5597	17.5737	18.6503	19.2630	19.8713	20.7184	20.9829
FARIMA	9.5243	12.4653	13.9931	15.3170	15.9440	16.9121	18.1768	19.3992	20.0650	20.6646	21.3436	21.2485
FARIMA-GARCH	9.4484	11.9278	13.5134	14.6761	15.3867	16.6261	17.5890	18.6808	19.3850	19.5411	20.0275	21.0284
AR-GARCH	8.9806	11.2693	12.3183	13.4480	14.1528	15.2228	16.2268	16.9971	18.3823	18.9495	19.7020	20.3335
ARMA-GARCH	8.9500	11.1117	12.2248	13.3279	14.1060	15.1862	16.1374	16.9447	17.6785	18.3797	19.4283	20.2095
FARIMA	9.2051	11.8403	12.8732	14.0170	14.6285	15.5856	16.6976	17.4234	18.1820	18.7080	19.5321	20.3975
FARIMA-GARCH	8.9395	11.1039	12.1899	13.3295	14.0626	15.0431	15.9554	16.7529	17.4654	18.1173	19.0571	19.9516
AR-GARCH	8.9147	11.0012	12.0614	12.8924	13.5308	14.2784	15.0095	15.7001	16.3001	16.7963	17.4993	18.1133
ARMA-GARCH	8.8133	10.7485	11.7698	12.6129	13.3533	14.0945	14.8281	15.5264	16.1573	16.7445	17.5192	18.1710
FARIMA	9.2152	11.6379	12.6884	13.5061	14.1514	14.8263	15.5033	16.2657	16.8299	17.3079	17.9160	18.4713
FARIMA-GARCH	8.8282	10.7252	11.6983	12.5623	13.2575	13.9068	14.6364	15.3256	15.9141	16.4840	17.1573	17.8439

Table 3.7. MAPE for Out Bps(%)

	5분후	10분후	15분후	20분후	25분후	30분후	35분후	40분후	45분후	50분후	55분후	60분후
AR-GARCH	12.2516	15.9923	18.1383	20.0627	22.2902	23.0451	24.0403	25.6917	26.6525	28.7626	30.1008	31.1824
ARMA-GARCH	11.4111	14.9802	17.3509	19.1134	20.8991	22.0531	23.0786	24.8176	26.2900	28.0555	29.7381	31.2802
FARIMA	11.3766	14.9114	17.1790	18.8329	20.5564	21.5741	22.7510	24.3091	25.6927	27.4579	28.7840	30.0124
FARIMA-GARCH	11.3828	14.9202	17.1738	18.2128	20.6605	21.1454	21.6754	23.0203	24.1796	26.6186	28.3986	29.5715
AR-GARCH	11.7125	15.8728	18.1104	19.6708	21.6739	23.0380	24.5265	26.2406	27.4869	28.9701	30.1396	31.1146
ARMA-GARCH	11.0738	14.9055	17.2279	18.7483	20.5245	22.1682	23.6863	25.2652	26.7013	28.0889	29.2433	30.4232
FARIMA	11.0840	14.8894	17.1583	18.6358	20.3399	21.8452	23.3931	24.8839	26.3093	27.6872	28.7310	29.7292
FARIMA-GARCH	11.4919	14.7006	17.1484	18.5628	20.1720	21.3152	23.2986	24.2383	26.2819	27.0389	28.2855	29.3229
AR-GARCH	11.7670	15.8288	17.9536	19.7472	21.6396	22.9989	24.3830	25.4635	26.7323	28.1214	29.2778	30.4111
ARMA-GARCH	11.1951	14.8843	17.1823	18.8550	20.5293	21.9782	23.3402	24.5576	25.7923	27.1974	28.2691	29.4998
FARIMA	11.1920	14.8747	17.1480	18.8131	20.3629	21.8232	23.2162	24.3972	25.6389	27.0215	28.0443	29.1678
FARIMA-GARCH	11.0829	14.8630	17.0916	17.8972	19.5200	21.6918	23.0002	24.0154	25.2147	26.6141	27.8303	28.9636

는 예측에 대하여 1시차 미래 예측부터 12시차 예측까지 실시하였으며, 모형에 따라 시차별 예측 정확도가 다르므로 관리 상황에 따라 적합한 예측 방법을 선택하여야 할 것으로 할 것이다. 향후 좀 더 광범위한 지역, 예를 들어, 대기업, 지자체, 단위 등 큰 지역에서 발생하는 트래픽의 분석 및 예측을 통하여 통신망의 안정적인 관리 및 수요 예측을 이룰 수 있다고 사료된다.

## References

- Basu, A., Mukherjee, A. and Klivansky, S. (1996). Time series models for internet traffic, In Proceedings IEEE INFOCOM 96, *Fifteenth Annual Conference of the IEEE Computer Societies*, **4**, 24–28.
- Baillie, Richard T. Chung, Ching-Fan. Tieslau, Margie A. (1996). Analysing inflation by the fractionally integrated FARIMA-GARCH model, *Journal of applied econometrics*, **11**, 23–40.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics*, **31**, 307–327.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. Inflation, *Econometrica*, **50**, 987–1008.
- Jiang, J. and Papavassiliou, S. (2006). Enhancing network traffic prediction and anomaly detection via statistical network traffic separation and combination strategies, *Computer communications*, **29(10)**, 1627–1638.
- Kim, S. (2007). Time series models for performance evaluation of network traffic forecasting, *The Korea Journal of Applied Statistics*, **20**, 219–227.
- Kim, S. (2011). Forecasting internet traffic by using seasonal GARCH models, *Journal of Communications and Network*, **13**, 621–624.
- Liu, J., Shu, Y., Zhang, L. and Xue, F. (1999). Traffic modeling based on FARIMA models, *IEEE Can. Conf. on Electrical and Computer Engineering*, 162–167.
- Shu, Y., Yu, M., Yang, O., Liu, J. and Feng, H. (2005). Wireless traffic modeling and prediction using seasonal ARIMA models, *IEICE-transactions on Communications*, **10**, 3992–3999.
- Tikunov, D. and Nishimura, T. (2007). Traffic Prediction for Mobile Network using Holt-Winter's Exponential Smoothing, *Telecommunications and Computer Networks, 15th International Conference*, 1–5.



# 장기기억 특성과 이분산성을 고려한 인터넷 트래픽 예측을 위한 시계열 모형 연구

손흥구<sup>a</sup> · 김삼용<sup>a,1</sup>

<sup>a</sup>중앙대학교 응용통계학과

(2013년 11월 26일 접수, 2013년 12월 24일 수정, 2013년 12월 24일 채택)

---

## 요약

본 논문은, 장기기억 특성과 이분산성을 고려한 인터넷 트래픽 예측 모형을 제안하고자 한다. 트래픽 과부하를 대비하기 위해서, 트래픽 용량은 트래픽의 예측치와 트래픽의 변동 크기에 따라 트래픽의 최대용량을 설정하여야 한다. 이를 위하여 교내 트래픽 자료 중 교내로 들어오는 트래픽과 교외로 나가는 트래픽에 이분산성과 장기기억 모형의 유용성을 확인하였다. 이에 대하여 AR-GARCH 모형, ARMA-GARCH 모형과 장기기억모형인 Fractional ARIMA와 장기기억과 이분산성을 고려한 Fractional ARMA-GARCH 모형을 적용하여 모형의 예측성능을 비교하였다.

주요용어: Fractional ARIMA,, Fractional ARIMA-GARCH, Internet Traffic, Forecasting, Bps.

---

---

이 논문은 2013년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2013-008277).

<sup>1</sup>교신저자: (156-756) 서울 동작구 흑석동 221, 중앙대학교 응용통계학과, 교수. E-mail: sahm@cau.ac.kr