

SMS 트래픽의 Self-similarity

하준, 신우철, 박진경, 최천원
단국대학교 대학원 전자컴퓨터공학과
전화 : 02-799-1095

Self-similarity of SMS Traffic

Jun Ha, Woo Cheol Shin, Jin Kyung Park, Cheon Won Choi
Department of Electronics and Computer Engineering, Graduate School
Dankook University
cchoi@dku.edu

Abstract

As the wireless mobile telecommunication system has been developed with astonishment, its offering service has also widely been expanded including various data service. Currently, the wireless mobile telecommunication network presents voice service that covers for the most part of the whole service areas. For this reason, the availability of the switching capacity in the mobile switching center(MSC) is manipulated by the required volume of voice service. However, considering the increase of data service, it is desirable for the current switching method to be modified for more efficiency. In this Paper, we analyze the data traffic caused by providing data service in the wireless mobile telecommunication network. For this, we are firstly going to review the result of the analysis in the feature of the data traffic. Secondly, based on the review, we are also going to perform analyzing the other feature of the data traffic normally generated in the wireless mobile telecommunication network.

We expect that this paper would be utilized as an elementary source for the feature of the SMS data traffic and it will be an honour for ourselves to work on it.

I. 서론

무선이동통신 기술이 발전함에 따라 무선이동통신망에서 제공되는 서비스가 다양해지고 있다. 이로 인해 음성 서비스뿐만 아니라 SMS, WWW, E-Mail 등 데이터 서비스로 그 사용이 확대되고 있다. 또한 데이터 서비스를 이용하는 사용자의 수와 그 사용량도 증가하는 추세이다. 제 3세대 이동통신망이 보급됨에 따라 서비스의 종류는 더욱 다양해지고, 데이터 서비스의 사용량은 더 증가될 것으로 예상된다. 현재의 이동통신망에서는 음성 서비스가 총 서비스 사용량의 대부분을 차지하고 있고, 그로 인해 교환국(Mobile Switching Center: MSC)에서의 교환 용량의 산출은 음성 서비스의 사용량에 의해 조율되고 있다. 그러나 앞으로 제공되리라 예상되는 다양한 데이터 서비스의 사용량 증가를 본다면, 현재의 음성 위주의 교환 용량 산출 방식은 수정되어야 하고, 이를 위하여 데이터 서비스의 제공으로 인해 발생하는 트래픽의 성질에 대한 분석이 선행되어야 한다 [1][2].

본 논문에서는 무선이동통신망에서 발생하는 데이터 트래픽의 self-similarity를 판별하기 위하여, 우선 self-similarity의 정의를 검토하고, 이어서 트래픽이 self-similarity를 갖는지 여부를 판별하기 위한 검정 방법을 고찰한다. 그리고, 현재 활발하게 사용되고 있는 데이터 서비스 중 하나인 단문 서비스(Short Message Service: SMS)를 대상으로 첫째, MSC에서 발생하는 SMS의 호 수를 측정한다. 둘째, 측정 자료를 바탕으로 하루에 발생하는 SMS 호 수의 평균, 분산 등의 통계량을 추정한다. 셋째, 측정 자료를 이용하여 SMS 트래픽의 self-similarity를 검정한다.

본 논문은 다음과 같이 구성된다. 2절에서는

self-similarity의 정의와 검정 방법에 대하여 기술한다. 3절에서는 측정 자료를 토대로 SMS 트래픽의 self-similarity를 검정한다.

II. Self-similarity의 정의와 검정

2.1 Self-similarity의 정의

Self-similarity는 Mandelbrot에 의해 제안된 개념이다 [3]. 프로세스가 self-similarity를 갖고 있다는 것은 시간 혹은 공간의 관점에서 이 프로세스를 관찰할 때 크기의 유사성이 보존됨을 의미한다.

Covariance-stationary 시퀀스 $X = \{X_n, n=1,2,\dots\}$ 의 자기 상관 함수(autocorrelation function)를 r 이라고 하자. 즉, $k \in \{1,2,\dots\}$ 에 대해

$$r(k) = E(X_n X_{n+k}). \quad (1)$$

(상수 μ 와 σ 에 대해 모든 $n \in \{1,2,\dots\}$ 에서 시퀀스 X 의 평균 함수(mean function)와 분산 함수(variance function)는 $E(X_n) = \mu$ 와 $E[(X_n - E(X_n))^2] = \sigma^2$ 와 같이 상수 함수이다.) 또한 자기 상관 함수 r 은 $k \rightarrow \infty$ 에 따라 다음의 성질을 갖는다고 가정하자.

$$r(k) \rightarrow c k^{-\beta}. \quad (2)$$

식 (2)에서 $\beta \in (0,1)$ 이고 c 는 상수이다. $n \in \{1,2,\dots\}$ 과 $m \in \{1,2,\dots\}$ 에 대해 확률 변수 $Y_n^{(m)} = \frac{1}{m}(X_{m(n-1)+1} + \dots + X_{mn})$ 이라고 정의하자. 또한 시퀀스 $Y^{(m)} = \{Y_n, n=1,2,\dots\}$ 의 자기 상관 함수를 $r^{(m)}(k)$ 라고 하자. (시퀀스 $Y^{(m)}$ 은 covariance-stationary 시퀀스이다.) 만약 모든 $m \in \{1,2,\dots\}$ 과 $k \in \{1,2,\dots\}$ 에 대해 $r^{(m)} = r(k)$ 라면 시퀀스 X 는 self-similarity를 갖는다고 정의한다 [4]. 한편 모든 $k \in \{1,2,\dots\}$ 에 대해 자기 상관 함수 $r^{(m)}(k)$ 가 $m \rightarrow \infty$ 에 따라 다음을 만족한다고 가정하자.

$$r^{(m)}(k) \rightarrow r(k). \quad (3)$$

그러면 시퀀스 X 는 점근적으로 self-similarity를 갖는다고 정의한다 [4]. 식 (2)의 자기 상관 함수는 쌍곡선형으로 감소하며, 지수형으로 감소하는 자기 상관 함수와 비교하여 매우 느리게 감소함을 알 수 있다.

또한 $\beta \in (0,1)$ 이므로 self-similarity를 갖는 시퀀스의 자기 상관 함수 값의 합은 무한대로 접근한다. 자기 상관 함수 값의 합이 무한대로 접근하는 사실은 self-similarity의 검정에 유용하게 사용된다.

Self-similarity의 분석에는 전통적으로 Hurst 파라미터가 사용된다. Hurst 파라미터 H 는 다음과 같이 정의된다 [4].

$$H = 1 - \frac{\beta}{2}. \quad (4)$$

식 (2)에서 $\beta \in (0,1)$ 이므로 Hurst 파라미터 H 는 다음을 만족한다.

$$H \in (\frac{1}{2}, 1). \quad (5)$$

또한 H 의 값이 1에 가까울수록 시퀀스의 self-similarity가 강하다고 할 수 있다 [4][6].

2.2 Self-similarity의 검정

만약 어떠한 프로세스가 self-similarity를 갖고 있다면 다음 세 가지 특성이 나타난다.

1. 프로세스의 표본으로부터 구한 분산의 추정 값이 표본 크기의 역수보다 매우 느리게 감소한다.
2. 자기 상관 함수가 지수 함수보다 훨씬 느리게 감소한다.
3. 스펙트럼 밀도 함수가 멱 법칙을 따른다.

상기 특성을 이용하여 임의의 프로세스가 self-similarity를 갖는지를 판별할 수 있다 [3].

Self-similarity를 검정하기 위한 방법으로 variance-time plot, re-scaled adjusted range plot(R/S plot), periodogram plot, 그리고 Whittle estimator 등이 발표되었다 [4]. 본 논문에서는 이러한 방법 중 variance-time plot 방법과 R/S plot 방법을 이용하여 SMS 트래픽의 self-similarity를 검정한다.

(1) Variance-time plot 방법

Variance-time plot 방법은 다음과 같다. $\overline{X}(n)$ 을 n 개의 표본으로부터 구한 평균의 추정 값이라고 하자. 그러면

$$\log(\text{Var}(\overline{X}(n))) = -\beta \log(n) + c \quad (6)$$

(여기에서 c 는 상수이다.) 식 (6)에서 $\log(n)$ 과 $\log(\text{Var}(\overline{X}(n)))$ 을 이용하여 β 의 추정 값을 구할 수

있다. 또한 식 (4)로부터 Hurst 파라미터 H 를 추정할 수 있다 [5][6].

(2) R/S plot 방법

R/S plot 방법은 다음과 같다. $\bar{X}(n)$ 과 $S(n)^2$ 을 n 개의 표본으로부터 구한 평균과 분산의 추정 값이라고 하자. 또한 $k \in \{1, 2, \dots\}$ 에 대해 $W_k = (X_1 + \dots + X_k) - k\bar{X}(n)$ 이라고 정의하자. 그러면 R/S 통계량을 다음과 같이 정의한다.

$$\frac{R(n)}{S(n)} = \frac{[\max(0, W_1, \dots, W_n) - \min(0, W_1, \dots, W_n)]}{S(n)} \quad (7)$$

일반적으로 임의의 프로세스는 $E\left(\frac{R(n)}{S(n)}\right) \sim cn^H$ 을 만족한다. (여기에서 c 는 상수이다.) 따라서 $n \rightarrow \infty$ 이면 [5][6]

$$\log\left(\frac{R(n)}{S(n)}\right) \approx H \log(n) + c. \quad (8)$$

식 (7)과 (8)로부터 $\log(n)$ 과 $\log\left(\frac{R(n)}{S(n)}\right)$ 을 이용하여 Hurst 파라미터 H 의 추정 값을 구할 수 있다.

본 논문에서는 SMS 트래픽에 대하여 variance-time plot 방법과 R/S plot 방법을 이용하여 파라미터 β 와 Hurst 파라미터 H 의 추정 값을 구하고 식 (5)에 주어진 $\beta \in (0, 1)$ 과 $H \in (\frac{1}{2}, 1)$ 의 조건을 충족하는 지를 확인하여 SMS 트래픽의 self-similarity를 검증한다.

III. SMS 트래픽의 Self-similarity

본 논문에서는 SMS 트래픽의 self-similarity를 검증하기 위하여 무선이동통신망에서 발생한 SMS 호 수의 측정 자료를 사용한다. SMS 호 수의 측정은 MSC에서 이루어졌으며 8 개월 동안 매월 보편성을 갖는 일정 수의 날을 선정하여 선정된 날마다 시간대 별로 발생하는 SMS 호 수를 측정하였다 [8].

그림 1은 SMS 트래픽의 variance-time plot을 보여 준다. 이 그림에서 표본의 크기는 80이고 중첩 표본 수는 8이다. 그림 1에서 variance-time plot 방법으로 구한 파라미터 β 의 추정 값은 0.07814이고 Hurst 파라미터 H 의 추정 값은 0.96093이다. 따라서 Hurst 파라미터 H 의 추정 값이 식 (4)의 조건을 만족하므로

SMS 트래픽은 self-similarity를 갖는다고 판정할 수 있다.

그림 2는 SMS 트래픽의 variance-time plot을 보여 준다. 이 그림에서 표본의 크기는 80이고 중첩 표본 수는 4이다. 그림 2에서 variance-time plot 방법으로 구한 파라미터 β 의 추정 값은 0.15243이고 Hurst 파라미터 H 의 추정 값은 0.92379이다. 따라서 Hurst 파라미터 H 의 추정 값이 식 (4)의 조건을 만족하므로 SMS 트래픽은 self-similarity를 갖는다고 판정할 수 있다.

그림 3은 SMS 트래픽의 R/S plot을 보여 준다. 이 그림에서 표본의 크기는 80이다. 그림 3에서 R/S plot 방법으로 구한 Hurst 파라미터 H 의 추정 값은 0.71585이다. 따라서 Hurst 파라미터 H 의 추정 값이 식 (4)의 조건을 만족하므로 SMS 트래픽은 self-similarity를 갖는다고 판정할 수 있다.

IV. 결론

무선이동통신망에서 MSC에서의 교환 용량 산출은 음성 트래픽 위주로 산정되고 있다. 하지만, 서비스의 종류가 다양해지고, 데이터 트래픽의 발생량 또한 증가함에 따라, 기존의 음성 위주의 MSC에서의 교환 용량 산정 방식은 수정되어야 할 것이다. 이를 위해서는 무선 이동통신망에서 발생하는 기존의 음성 트래픽과는 다른 성질을 가지고 있는 데이터 트래픽은 성질에 대한 분석이 반드시 필요하다.

본 논문에서는, 데이터 트래픽의 성질 중 하나인 self-similarity 성질을 검증하기 위하여, 현재 가장 많이 사용되고 있는 서비스 중 하나인 SMS 트래픽의 측정값을 이용하였다. SMS 트래픽의 평균과 분산의 추정값을 이용하여, variance time plot과 R/S plot 방법으로 Hurst 파라미터 값을 추정하고, 이를 이용하여, 무선이동통신망에서 측정된 SMS 트래픽 데이터가 self-similarity 성질을 가지고 있음을 알 수 있었다.

본 논문에서는 무선이동통신망에서 발생하는 데이터 트래픽의 self-similarity 성질을 검증하였고, 향후 다양한 데이터 서비스의 제공으로 인해 발생할 데이터 트래픽의 성질을 규명하는 기초 자료가 될 것으로 기대된다.

참고문헌

- [1] H. Holma and A. Toskala, *WCDMA for UMTS*. John Wiley & Sons, 2000.
- [2] D. Knisely, S. Kumar, S. Laha, and S. Nanda,

"Evolution of Wireless Data Services: IS-95 to cdma2000", IEEE Communications Magazine, Vol. 36, No. 10, pp 140-149, October 1998.

- [3] M. Listani and V. Eramo, "Architectural and Technological Issues for Future Optical Internet Networks", IEEE Communications Magazine, Vol. 38, No. 9, pp. 82-92, September 2000.
- [4] W. Leland, M. S. Taqqu, W. Willinger, "On the Self-similar Nature of Ethernet Traffic", IEEE/ACM Transactions on Networking, Vol. 2, No. 1, pp. 1-15, February 1994.
- [5] M. E. Crovella, A. Bestavros, "Self-similarity in World Wide Web traffic: evidence and possible causes", IEEE/ACM Transactions, Vol. 5, No. 6, pp. 835-846, December 1997.
- [6] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1982.
- [7] Kihong Park, Walter Willinger, *Self-Similar Network Traffic and Performance Evaluation*, John Wiley & Sons, 2000.
- [8] 하준, 이정수, 김민석, 하장용, 최천원, "SMS 트래픽의 측정, 분석, 모델", 대한전자공학회 하계종합학술대회 논문집, 24권 1호, pp. 125-128, 2001.

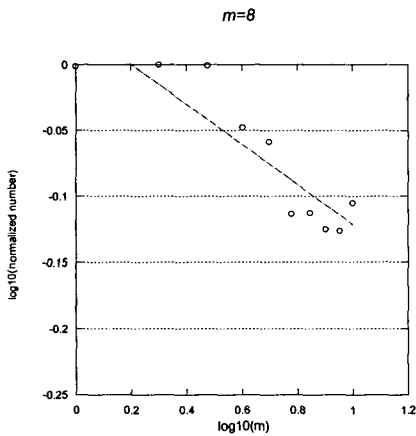


그림 1. SMS 트래픽의 variance-time plot

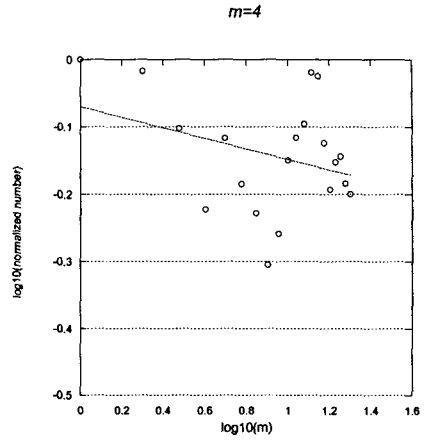


그림 2. SMS 트래픽의 variance-time plot

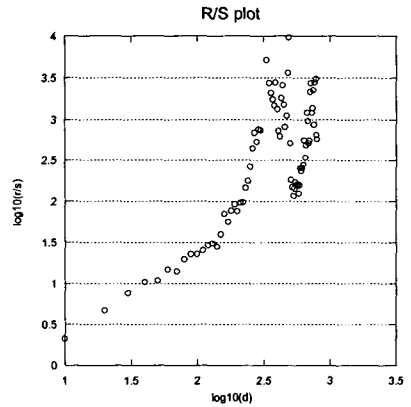


그림 3. SMS 트래픽의 R/S plot