

## SMS 트래픽의 측정, 분석, 모델

하준\* 이정수\*\* 김민석\*\* 하장용\*\* 최천원\*  
단국대학교\* SK 텔레콤\*\*

### Measurement, Analysis and Model of SMS Traffic

Jun Ha\* Jeoung Soo Lee\*\* Min Seok Kim\*\* Jang Yong Ha\*\* Cheon Won Choi\*  
Dankook University\* SK Telecom\*\*

#### 요약

본 논문에서는 SMS 트래픽을 측정하고 트래픽의 주요 통계량에 대해 측정 자료를 기반으로 추정값을 구한다. 이어서 측정자료에 대한 관찰에 기초하여 가입자 큐잉 시스템을 근간으로 하는 SMS 트래픽 모델을 구축한다. 또한 이러한 모델로부터 모의 실험을 통해 트래픽을 발생시키고 주요 통계량에 대한 추정값을 구해서 측정 자료를 바탕으로 한 추정값과 비교하여 개발된 SMS 트래픽 모델의 정확성을 검증한다.

#### 1. 서론

이동통신망이 제공하는 서비스는 음성 서비스뿐만 아니라 WWW 서비스 등 데이터 서비스로 확대되고 있다. 또한 데이터 서비스의 사용자 수 및 사용량도 증가하는 추세이다. 이러한 서비스의 다양화와 데이터 서비스 사용량의 증가 경향은 제 3 세대 이동통신망의 보급과 함께 심화될 것으로 예상된다 [3][7][9][10]. 현재 이동통신망에서는 전통적인 음성 서비스가 총 서비스 사용량의 대부분을 차지하며 따라서 서비스의 원활한 제공을 위한 교환국(Mobile Switching Center: MSC)의 교환 용량도 음성 서비스의 사용량에 조율되고 있다. 앞으로 예상되는 서비스 다양화와 데이터 서비스 사용량의 증가에 비추어 이러한 교환 용량의 산출 방식은 수정되어야 하고, 이를 위하여 데이터 서비스의 제공으로 발생하는 트래픽의 성질에 대한 분석 및 모델이 선행되어야 한다. 현재 GSM, IS-95 등 제 2 세대 이동통신망에서 활발한 사용이 있는 데이터 서비스 중 하나는 단문 서비스(Short Message Service: SMS)이다 [7][9][10]. 본 논문에서는 IS-95 기반의 제 2 세대 이동통신망에서 제공되는 본 연구는 SK Telecom의 지원으로 수행되었음.

는 SMS를 대상으로 첫째, MSC에서 발생하는 트래픽(구체적으로 시간별 발생하는 SMS의 호 수)을 측정하고 둘째, 측정 자료를 바탕으로 시간대별 발생하는 SMS 호 수의 평균, 분산, 99th percentile 등 통계량을 추정한다. 셋째, 각 서비스 사용자에게 대해 단일 서버 및 단일 버퍼의 가입자 큐잉 시스템을 구성하고, 측정 자료에 대한 관찰에 기초하여 이러한 큐잉 시스템의 호 도착 process를 nonhomogeneous Poisson process로 설정하며, MSC에서 SMS 호 발생 process의 모델로서 가입자 큐잉 시스템의 수락된 호 도착 process가 중첩되어 형성되는 process를 제시한다. 끝으로 이러한 호 발생 process의 모델이 갖는 정확성을 검증하기 위해 모의 실험을 통해 가입자 큐잉 시스템의 수락된 호 도착 process를 발생시키고 SMS 호 수에 관한 통계량의 추정값을 추출하여 측정 자료를 통해 구한 추정값과 비교한다.

본 논문은 다음과 같이 구성된다. 2 절에서는 SMS 트래픽의 측정 방법이 서술되고 3 절에서는 SMS 호 수와 관련된 통계량의 추정 방법이 기술된다. 4 절에서는 SMS 호 발생 process의 모델이 설정되고, 측정 자료와 SMS 트래픽 모델에 대한 모의 실험 결과로부터 구한 추정값의 비교는 5 절에서 수행된다.

#### 2. SMS 트래픽의 측정

발생하는 SMS 호 수의 측정은 MSC에서 이루어진다. 이러한 측정은 두 가지 방향으로 추진된다. 첫째, 월별 보편성을 갖는 일정 수의 날을 선정하여 각 날 기준 시간대에 발생하는 SMS 호 수를 측정한다. (본 논문에서 기준 시간대는 하루 중 음성 서비스 호가 가장 많이 발생하는 시간대로 설정한다.) 둘째, 보편성을 갖는 날 중 일

부를 택하여 시간대별 발생하는 SMS 호 수를 측정한다.

구해진 시간대별 SMS 호 발생 빈도가 도시되어 있다.

### 3. SMS 트래픽의 측정 자료 분석

월별 보편성을 갖는 날 기준 시간대에 측정된 SMS 호 수를 이용하여 매 월 임의의 날 기준 시간대에 발생하는 SMS 호 수의 평균, 분산, 99th percentile의 추정값을 구한다. 또한 하루 중 시간대별 발생하는 SMS 호 수의 측정 자료를 바탕으로 시간대별 SMS 호의 발생 빈도를 추정한다.

첫째, 평균 및 분산의 추정값은 moment 추정 방식으로 구해지며 99th sample percentile로 99th percentile을 추정한다 [2][6]. 한 달 중 보편성을 갖는 임의의 시간대에 발생한 SMS 호 수를  $Y$ 라고 하자. 또한  $Y$ 의 평균, 분산, 99th percentile을 각각  $\mu$ ,  $\sigma^2$ ,  $\gamma$ 라고 하자.  $y_1, \dots, y_n$ 을 한 달 중  $n$ 일간 측정된 일별 기준 시간대에 발생한 SMS 호 수라고 하자. 이러한 샘플로부터  $\mu$ 와  $\sigma^2$ 의 추정값  $\hat{\mu}$ ,  $\hat{\sigma}^2$ 를 다음과 같이 정한다 [2].

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\mu})^2 \quad (1)$$

또한 99th percentile의 추정값  $\hat{\gamma}$ 은 다음과 같다 [6].

$$\hat{\gamma} = y_{\lfloor 0.99n \rfloor} \quad (2)$$

(위 식에서  $\lfloor x \rfloor$ 는  $x$ 보다 크거나 같은 최소 정수이다.) 둘째, 시간대별 SMS 호의 발생 빈도는 다음과 같이 추정한다.  $h=1, \dots, 24$ 에 대해  $\zeta(h)$ 를 임의의 SMS 호가 하루 중 시간대  $(h-1, h]$ 에 발생할 확률이라고 하자. 한 달 중  $n$ 일간 시간대별 발생하는 SMS 호 수에 대한 측정이 수행되어  $\{y_i(h) : i=1, \dots, n, h=1, \dots, 24\}$ 의 샘플을 얻었다면,  $\zeta(h)$ 의 추정값  $\hat{\zeta}(h)$ 를 다음과 같이 정한다.

$$\hat{\zeta}(h^*) = \frac{1}{n} \sum_{i=1}^n \frac{y_i(h^*)}{\sum_{h=1}^{24} y_i(h)} \quad (3)$$

(식 (3)에서  $y_i(h)$ 는  $i$ 날 시간대  $(h-1, h]$ 에 발생한 SMS 호 수이다.) 그림 1에는 측정 자료를 이용하여

### 4. SMS 트래픽의 모델

#### 4.1 가입자 큐잉 시스템

각 가입자는 호가 도착하면 호 지속 시간동안 서비스가 제공되는 큐잉 시스템으로 간주될 수 있다. 가입자 큐잉 시스템에 호가 도착하면 일단 버퍼에 저장되고 서비스 규칙(service discipline)에 따라 버퍼에 저장된 호들의 서비스를 받는 순서가 결정된다. 예를 들어 FCFS(First Come First Served)의 서비스 규칙이 적용된다면 앞서 도착한 호의 서비스가 모두 종료된 후 새로 도착한 호의 서비스가 개시된다. 각 호의 서비스는 호 지속 시간동안 이루어지며 서비스가 종료되면 호는 가입자 큐잉 시스템을 떠난다.

일반적으로 임의의 시각에 한 가입자는 둘 이상의 호에 참여할 수 없다. 또한 호가 지속되는 동안에 새로운 호의 발생 시도가 있다면 이 발생 시도는 거부된다. 따라서 가입자 큐잉 시스템은 호의 서비스가 진행되는 사이에 도착하는 호는 버퍼에 저장되지 않고 큐잉 시스템으로의 입장이 거부되는 단일 서버 및 단일 버퍼(single-server single-buffer) 큐잉 시스템이 된다 [4][8].

#### 4.2 SMS 호 지속 시간

호 지속 시간의 모델로 지수 분포와 데이터 서비스 트래픽이 갖는 self-similarity 성질을 고려하여 Pareto 분포를 설정한다 [1][11][13]. 즉,  $S$ 를 호 지속 시간이라고 하면 다음과 같은 분포를 갖는다. 지수 분포 모델에서

$$P(S \leq t) = 1 - e^{-\frac{t}{E(S)}} \quad (4)$$

또한 Pareto 분포 모델에서

$$P(S \leq t) = 1 - \left( \frac{t}{\min\{S\}} \right)^{\left( 1 + \frac{\min\{S\}}{E(S)} \right)} \quad (5)$$

#### 4.3 SMS 호 도착 process

그림 1에 나타난 호 발생 빈도에 대한 관찰에 기초하여 가입자  $k$ 의 가입자 큐잉 시스템에서 호 도착 process를 intensity 함수  $\lambda^*(k, t)$ 의 nonhomogeneous Poisson process로 설정한다 [5][12]. 여기에서 intensity 함수  $\lambda^*(k, t)$ 는 다음과 같이 유추된다..

그림 2에 도시된 바와 같이 측정 자료를 이용하여 기준 시간대  $(h^*-1, h^*]$ 에 발생하는 가입자  $k$ 의 평균 호 수  $\eta(k, h^*)$ 를 정할 수 있다. 이를 이용하여 시간

대  $(h-1, h]$ 에서 발생하는 가입자  $k$ 의 평균 호 수  $\eta(k, h)$ 는 다음과 같이 계산된다.

$$\eta(k, h) = \eta(k, h^*) \frac{\zeta(h)}{\zeta(h^*)} \quad (6)$$

(식 (6)에서  $\zeta(h)$ 는 시간대  $(h-1, h]$ 에서 SMS 호의 발생 빈도이다.) 이 때  $\eta(k, h)$ 는 가입자 큐잉 시스템에 도착하는 호 중 수락된 평균 호 수이다.  $\eta^*(k, h)$ 를 시간대  $(h-1, h]$ 에 가입자  $k$  큐잉 시스템에 도착하는 총 평균 호 수라고 하자. 주어진 시간대에서 큐잉 시스템은 M/G/1/1 시스템이 되므로 다음의 관계식을 얻는다 [4][8].

$$\eta(k, h) = \eta^*(k, h)[1 - \eta^*(k, h)E(S)] \quad (7)$$

식 (7)의  $\eta^*(k, h)$ 를 이용하여 intensity 함수  $\lambda^*(k, t)$ 는 다음과 같이 표현된다.

$$\lambda^*(k, t) = \sum_{h=1}^{24} \eta^*(k, h) \cdot I_{(h-1 < t \bmod 24 \leq h)} \quad (8)$$

#### 4.4 호 발생 process

4.3 절에 기술된 호 도착 process와 4.2 절에 서술된 호 도착 시간으로 특징지워지는 단일 서버 및 단일 버퍼 가입자 큐잉 시스템에서 수락된 호 도착 process를 고려하자. 모든 가입자에 대해 이러한 호 도착 process의 중첩으로 구성되는 process를 MSC에서의 SMS 호 발생 process로 설정한다.

가입자 큐잉 시스템은 단일 서버 및 단일 버퍼 큐잉 시스템이므로 각 가입자는 임의의 시간에 최대 한 개의 호에 참여할 수 있다. 따라서 가입자 수가  $\nu$  라면 이동통신망에서 최대  $\nu$  개의 호가 동시에 진행될 수 있다. 그러나, 실제 이동통신망에서는 무선 링크의 용량 제약으로 이러한 최대  $\nu$  개의 호에 대한 동시 지원이 이루어지지 못한다. 그러나, 가입자 큐잉 시스템에서 이러한 시도는 반드시 성공하므로 호 발생률의 낙관적인 (optimistic) 추정값을 얻게 된다.

### 5. 모의 실험을 통한 SMS 트래픽의 발생과 비교 분석

본 절에서는 4 절에 서술된 SMS 트래픽 모델을 바탕으로 모의 실험을 통해 SMS 트래픽을 발생시키고, 기준 시간대에 발생하는 SMS 호 수의 샘플을 추출하여, 평균, 분산, 99th percentile 등의 추정값을 구한다. 이어서

측정자료를 근거로 얻은 주요 통계량의 추정값과 모의 실험을 통해 구한 추정값을 비교한다.

모의 실험에서 SMS 호 지속 시간은 지수 분포 혹은 Pareto 분포를 갖는다고 가정되고 이때 평균 호 지속 시간은 1 sec으로 설정된다. Pareto 분포 모델에서는 추가로 최소 호 지속 시간이 평균 호 지속 시간의 반으로 고정된다. 또한 모의 실험에서 8 개월간 매 달 10 일 동안 SMS 트래픽을 발생시키며 가입자 수는 측정 자료로부터 구해진 매 달 평균 가입자 수가 사용된다. 이때 모든 가입자의 동질성이 가정된다.

그림 2에는 모의 실험을 통해 얻은 기준 시간대에 발생하는 SMS 호 수의 평균의 추정값이 측정 자료를 기초로 구해진 추정값과 비교되어 있다. 그림 2에서 측정 자료로부터 얻은 평균의 추정값과 비교해서 모의 실험을 통해 구한 평균의 추정값은 유의할 수준의 차이를 보이지 않음을 알 수 있다.

그림 3에는 모의 실험을 통해 구한 기준 시간대에 발생하는 SMS 호 수의 99th percentile의 추정값과 측정 자료를 이용하여 얻은 추정값이 도시되어 있다. 이 그림에서 모의 실험을 통해 얻은 99th percentile의 추정값은 측정 자료에서 구한 추정값에 비해 다소 작음을 관찰할 수 있다. 이러한 현상은 SMS 트래픽 모델에 내재한 다음 두 가지 요인에 기인한다고 판단된다. 한 가지는 한 달 동안 가입자 수는 변동이 없다는 가정이며 다른 요인은 가입자가 동질성을 갖는다는 가정이다. 이러한 두 가지 가정이 위배될 경우 기준 시간대에 발생하는 SMS 호 수의 분산은 증가하며 따라서 모의 실험을 통해 얻은 99th percentile의 추정값이 측정 자료로부터 구한 추정값에 보다 근접하게 된다.

### 6. 결론

본 논문에서는 이동통신망에서 제공되는 SMS로 인해 발생하는 호 수를 측정하고 이러한 측정 자료를 바탕으로 기준 시간대에 발생하는 SMS 호 수의 평균, 분산, 99th percentile 등의 추정값을 구하였다. 이러한 기준 시간대에 발생하는 평균 SMS 호 수와 하루 중 시간대별 SMS 호 발생 빈도에 대한 관찰값을 바탕으로, nonhomogeneous Poisson process로 설정된 호 도착 process와 지수 분포 혹은 Pareto 분포로 대표되는 호 도착 시간으로 특징 지워지는 단일 서버 및 단일 버퍼 가입자 큐잉 시스템을 구성하여, 이러한 가입자 큐잉 시스템에서 수락된 호 도착 process를 이용한 SMS 트래픽 모델을 구성하였다.

본 논문에서 개발된 SMS 트래픽 모델에 대한 모의 실험을 통해 기준 시간대에 발생하는 평균 SMS 호 수라는 측면에서 본 모델은 상당히 높은 정확성을 나타냄이

확인되었다. 그러나 SMS 호 수의 99th percentile에서는 이러한 추정값이 다소 차이를 나타내었다. 이는 SMS 트래픽 모델에 내재한 가입자의 동질성 가정에 기인한다고 판단되며, 따라서 가입자의 균화를 시도하고 가입자 군별 추정 자료를 확보하여 가입자 이질성을 모델에 반영함으로써 본 SMS 트래픽 모델은 개선될 것이다. 본 논문에서 개발된 트래픽 모델은 SMS에 국한되지 않고 앞으로 다양화 및 활성화될 데이터 서비스의 트래픽 모델 개발에 기여할 것으로 예상된다.

**참고문헌**

- [1] R. Addie, M. Zukerman, and T. Neame, "Broadband traffic Modeling : Simple Solutions to Hard Problems," IEEE Communications Magazine, Vol. 36, No. 8, pp. 88-95, August 1998.
- [2] P. Bickel and K. Doksum, *Mathematical Statistics : Basic Ideas and Selected Topics*. Holden-Day, 1977.
- [3] P. Chaudhury, W. Mohr, and S. Onoe, "The 3GPP Proposal for IMT-2000", IEEE Communications Magazine, Vol. 37, No. 12, pp. 72-81, December 1998.
- [4] J. Cohen, *The Single Server Queue*. North-Holland, 1982.
- [5] R. Durrett, *Essentials of Stochastic Processes*. Springer-Verlag, 1999.
- [6] T. Ferguson, *A Course in Large Sample Theory*. Chapman & Hall, 1996.
- [7] V. Garg, *IS-95 CDMA and cdma2000*. Prentice Hall, 1999.
- [8] D. Gross and C. Harris, *Fundamentals of Queueing Theory*. 3rd edition, John Wiley & Sons, 1998.
- [9] H. Holma and A. Toskala, *WCDMA for UMTS*. John Wiley & Sons, 2000.
- [10] 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.
- [11] W. Leland, "On the Self-similar Nature of Ethernet Traffic," IEEE/ACM Transactions on Networking, Vol. 2, No. 1, pp. 1-15, February 1994.
- [12] S. Ross, *Stochastic Processes*. 2nd edition, John Wiley & Sons, 1996.
- [13] TIA TR45.5, "The cdma2000 ITU-R RTT Candidate Submission," Draft document TR45/98.03.03.22 R1, June 1998.

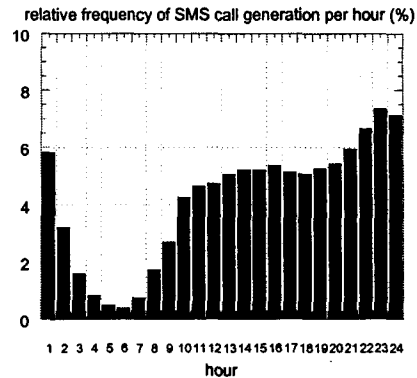


그림 1 시간대별 SMS 호 발생 빈도의 추정값

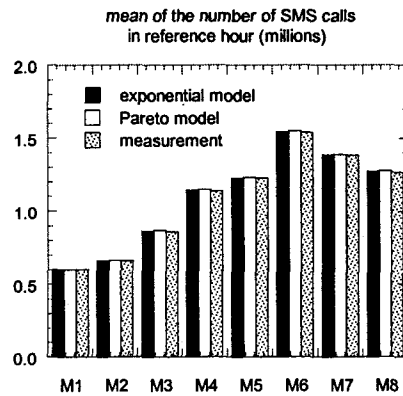


그림 2 기준 시간대에 발생하는 SMS 호 수의 평균의 추정값

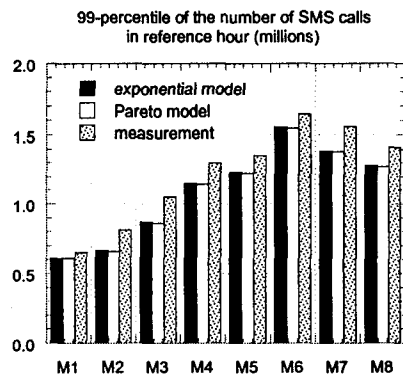


그림 3 기준 시간대에 발생하는 SMS 호 수의 99th percentile의 추정값