

자연시간과 손실율을 고려한 데이터 트래픽 분석

임석구

요 약

차세대 고속 통신망을 구축하기 위해서는 해결해야 할 많은 문제들이 있는데, 이 중에서 기본적으로 고려해야 할 사항은 바로 망에 흐르는 트래픽의 특성 분석이다. 현재 제공되는 많은 인터넷 서비스들의 동작 특성은 자기 유사성(Self-similar)이라는 기조에 고려되던 트래픽 특성과는 완전히 다른 장기간 의존성의 성질을 가진다는 것이 증명되었다 이러한 장기간 의존성 성질을 표현하기 위한 모델로는 자기유사 모델이 있는데, 이것은 단기간 의존성을 표현하는 기존의 모델인 포아송 모델과는 상반되는 개념이다. 따라서 차세대 통신망의 설계 및 디멘셔닝을 위해서는 무엇보다도 데이터 트래픽의 주요 특성인 버스트성(Burstiness)과 자기유사성이 반영된 트래픽 모델이 요구된다. 여기서 자기유사성은 허스트 파라미터(Hurst Parameter)로 특성화 될 수 있다. 본 논문에서는 데이터 트래픽의 자기유사성 및 큐잉지연을 고려한 유효대역폭 산출식을 유도하여 시뮬레이션 결과와 비교 분석하였다.

An Analysis of Data Traffic Considering the Delay and Cell Loss Probability

Seog -Ku Lim

ABSTRACT

There are many problems that must solve to construct next generation high-speed communication network. Among these, item that must consider basically is characteristics analysis of traffic that flows to network. Traffic characteristics of many Internet services that is offered present have shown that network traffic exhibits at a wide range of scales - self-similarity. Self-similarity is expressed by long term dependency, this is contradictory concept with Poisson model that have relativity short term dependency. Therefore, first of all, for design and dimensioning of next generation communication network, traffic model that are reflected burstiness and self-similarity is required. Here self-similarity can be characterized by Hurst parameter. In this paper, the calculation equation is derived considering queueing delay and self-similarity of data traffic, and compared with simulation results.

Key words : self-similar, Hurst parameter, cell loss rate, cell delay probability

1. 서 론

세계 각국에서는 21세기 지식 정보사회의 도래를 대비하여 국가 정보 하부구조를 구축하는데 인터넷(Internet)을 근간으로 하고자 하는 노력이 활발히 추진되고 있다. 그러나 인터넷 사용자가 급속하게 증가하면서 인터넷은 느린 속도와 제한적인 멀티미디어 서비스, 보안 문제와 같은 사용자의 불만에 직면하게 되었고, 이를 해결하려는 노력이 전세계적으로 이루어지고 있

며, 이러한 노력은 차세대 인터넷(NGI: Next Generation Internet)*이라는 용어로 가장 잘 나타낼 수 있다. 차세대 고속 통신망을 구축하는데 있어서 가장 중요한 문제는 대역폭의 확대, 망의 확장성(Scalability), 그리고 다양한 서비스를 제공할 수 있는 능력으로 집약할 수 있다.

차세대 고속 통신망을 구축하기 위한 시도는 두 가지 방향으로 진행되고 있다고 분석된다. 첫째는 고속의 패킷 처리 능력을 가지는 기가비트 라우터(Gigabit Router)를 통하여 해결하자는 방향이다.

* 제일저자(First Author) : 임석구, 주소 : 충남 천안시 안서동 115번지 천안대학교 정보통신학부, 전화 : 041) 620-9478,

Fax : 041) 620-9507, E-mail : sklim@infocom.cheonan.ac.kr

접수일 : 2004년 2월 14일, 완료일 : 2004년 2월 27일

* 정회원, 천안대학교 정보통신학부

다른 방법은 데이터링크 계층의 스위칭 기술을 이용하여 패킷을 고속으로 전달하고자 하는 것이다.

그러나 앞에서 기술한 이러한 문제 해결에 앞서 기본적으로 고려해야 할 사항은 바로 인터넷 망에 흐르는 트래픽의 특성 분석이다. 인터넷에서 지원하고자 하는 서비스가 음성 뿐만이 아닌 고속 인터넷 서비스와 동영상 서비스와 같이 광대역을 필요로 하는 서비스이기 때문에 기존의 Markovian 모델에 기초한 음성 위주의 트래픽 모델링으로는 실제 서비스 트래픽을 모사할 수 없으며, 이러한 모델링을 통한 시스템 설계와 성능 예측은 부정확한 시스템 파라미터의 설정과 잘못된 성능 분석 결과를 초래할 수 있다.

현재 인터넷에서 제공되고 있는 다양한 멀티미디어 서비스의 트래픽 특성 중 가장 중요한 특성중 하나인 자기유사성(Self-similarity)은 기존에 고려되던 트래픽 특성과는 완전히 다른 장기간 의존성(Long Range Dependency)를 가진다. 이것은 음성중심의 트래픽에서 나타나는 단기간 의존성(Short Range Dependency)을 표현하는 기존 모델인 포아송 모델과는 상반되는 개념으로서 효율적인 네트워크 디멘저닝 절차와 트래픽 관리 기능 등에 상당히 중요한 영향을 미칠 것이며, 현재 전 세계적으로 이러한 트래픽 모델에 관한 연구가 활발히 진행 중이다[1][2][3][4][5].

또한 앞으로는 사용자의 우선순위(Priority)를 고려한 고속의 데이터 속도 보장요구 및 데이터 서비스중 화상 통신 등과 같은 지연에 민감한 서비스가 요구되고 있다. 이와 같이 지연에 대해 관대했던 기존의 데이터 특성과는 달리 차세대 시스템에서는 패킷 데이터 서비스에 대한 기존의 셀 손실 확률(CLR: Cell Loss Rate) 뿐만 아니라 지연(Delay) 및 지연확률(Delay Probability)이 중요한 GoS(Grade of Service) 파라미터로 대두되고 있다.

서론에 이어 2절에서는 자기 유사성을 갖는 데이터 트래픽의 큐잉 성능을 모델링한 해석적 모델인 Norros와 Z. FAN의 Effective Bandwidth Formula에 대해 지연을 고려한 수식을 산출하여 지연이 고려된 셀 손실확률과 셀 지연확률(Cell Delay Probability)관점에 대해 살펴보고 3절에서는 이러한 해석적 모델의 타당성을 검토하기 위해 시뮬레이션 결과와 비교 분석하였으며, 마지막으로 4절에서는 결론을 맺는다.

2. 유효 대역폭과 셀 손실 확률

기존의 음성 트래픽과는 달리 자기유사성을 가지는 데이터 트래픽의 경우 시스템이나 네트워크 내에서 수많은 Aggregation과정을 거치더라도 분산이 거의 줄어들지 않기 때문에 단순히 평균 트래픽의 양으로 시스템을 디멘저닝하기에는 우리가 따른다. 따라서 데이터 트래픽의 특징인 버스트함(Burstiness)과 자기유사성이 반영된 유효 대역폭(Effective Bandwidth)으로 네트워

크 트래픽을 수용하는데 필요한 전송용량 및 큐 모델의 서버용량 등을 평가할 필요가 있다.

본 논문에서는 자기유사 트래픽의 큐잉 성능인 셀 손실확률과 지연확률을 분석하기 위한 해석적 모델로써 Norros의 Effective Bandwidth Formula와 Bahadur-Rao Theorem을 적용해 Norros 수식을 보완한 Z. FAN의 Formula를 고려한다.

2.1 유효 대역폭

자기유사 트래픽 소스에 대한 Norros의 유효 대역폭(C)은 식(1)과 같이 정의된다[6].

$$C = m + [x(H)\sqrt{-2\ln\epsilon}]^{1/H} \cdot (ma)^{1/(2H)} B^{-(1-H)/H} \quad (1)$$

여기서, m은 트래픽 스트림의 평균 비트율(bits/sec)이고, a는 트래픽 스트림의 분산계수(σ^2/m) [bit-sec]이다. H는 트래픽 스트림의 자기 유사성 정도를 나타내는 Hurst 파라미터($0.5 \leq H < 1$)이다. B는 버퍼 크기(bits)이고 $x(H)$ 는 $x(H) = H^H(1-H)^{1-H}$ 이며, ϵ : Target Cell Loss Rate 이다. 예를 들어, $H=0.95$ 이고 $m=38.8\text{Kbps}$, $a=1\text{ kbits-sec}$ 을 갖는 FTP 파일 전송의 경우 요구되는 유효대역폭은 48.03Kbps이 된다(CLR : 10^{-6} , B : 16K cell)

식(1)에서 유효 대역폭(C)은 자기유사성을 나타내는 Hurst 상수(H)와 트래픽의 버스트함을 나타내는 분산계수(a)가 커짐에 따라 증가하고, 또한 버퍼크기(B)와 Target CLR(ϵ)이 커질수록 유효 대역폭(C)은 낮아짐을 알 수 있다.

2.2 셀 손실 확률

본 절에서는 자기유사 트래픽의 큐잉 성능중 하나인 셀 손실 확률(CLR)에 대한 Norros와 Z. FAN의 두 가지 셀 손실 확률 식에 대해 설명한다.

(1) Norros's Formula

버퍼의 크기가 B이고, 자기유사 트래픽을 처리하는데 요구되는 서버의 서비스율이 C인 Single-Queue Single-Server 시스템에서의 셀 손실 확률은 식(1)로부터 유도할 수 있으며, 식(2)와 같이 근사화 된다 [6][8].

$$\epsilon = \Pr\{Q > B\} = \exp\left[-\frac{(C-m)^{2H}}{2x(H)^2ma} B^{2-2H}\right] \quad (2)$$

식(2)를 이론적으로 유도하는데 있어서 기반이 되는 두 가지 가정사항은 매우 많은 i.i.d.한 트래픽이 통합되어졌을 때 Gaussian 특성을 따른다는 것과, 큐 길이의 분포의 꼬리부분이 Weibull 분포를 따른다는 가정이

다. 따라서, 이 식은 적은 버퍼 크기에 대해서 오차가 많은 단점이 있다.

(2) Z. Fan's Formula

Z. Fan은 Norros's Formula의 단점인 적은 버퍼 크기에 대한 오차를 줄이기 위해 Bahadur-Rao Theorem을 사용하여 근사적인 셀 손실 확률을 제시하였으며, 이를 정리하면 식(3)과 같다[9].

$$\Pr[Q > B] = \frac{1}{\sqrt{2\pi\sigma_{t^*}^2 \theta_{t^*}}} \cdot \exp\left[-\frac{(C-m)^{2H}}{2x(H)^2 ma} B^{2-2H}\right] \quad (3)$$

여기서,

$$\sigma_{t^*}^2 = ma(t^*)^{2H}, \quad \theta_{t^*} = \frac{B + Ct^* - mt^*}{ma(t^*)^{2H}},$$

$$t^* = \frac{HB}{(1-H)(C-m)} \text{이다.}$$

2.3. 셀 지연 확률

서론에 밝힌바와 같이 고속 패킷 데이터 서비스에서 요구되는 다양한 지연시간을 충족시키기 위해 시스템 설계 시 셀 손실 확률뿐만 아니라 지연시간을 충족시킬 수 있도록 해야 한다. 따라서 본 절에서는 앞 절에서 소개한 셀 손실 확률식[식(2),(3)]으로부터 셀 지연 확률을 유도하고자 한다.

셀 지연 확률을 유도하기 위해 버퍼의 크기는 셀 지연 시간(d)을 만족할 수 있는 만큼 충분히 크다고 가정하면, 버퍼크기 (B)와 셀 지연 시간(d)은 다음과 같은 관계를 만족해야 한다.

$$B \geq dC \quad (4)$$

식(4)은 셀 버퍼설계 시 지연시간(d)을 만족하기 위해서는 최소한 dC보다 크게 설계되어야 한다는 것을 의미한다. 이를 바탕으로 하여 지연이 고려된 유효 대역폭을 다음과 같이 산출할 수 있다.

(1) 셀 지연이 고려된 Norros's Formula

지연시간이 d이고, 자기유사 트래픽을 처리하는 데 요구되는 서버의 서비스율이 C인 Single-Queue Single-Serve 시스템에서의 셀 지연 확률은 식(2)와 식(4)의 관계로부터 유도할 수 있으며, 식(5)와 같이 근사화 된다.

$$\Pr[D > d] = \exp\left[-\frac{(C-m)^{2H}}{2x(H)^2 ma} (dC)^{2-2H}\right] \quad (5)$$

(2) 셀 지연이 고려된 Z. Fan's Formula

Z. Fan의 셀 지연 확률도 식(5)와 동일한 방식으로 유도하면 식(6)과 같이 근사화 된다.

$$\Pr[D > d] = \frac{1}{\sqrt{2\pi\sigma_{t^*}^2 \theta_{t^*}}} \cdot \exp\left[-\frac{(C-m)^{2H}}{2x(H)^2 ma} (dC)^{2-2H}\right] \quad (6)$$

여기서,

$$\sigma_{t^*}^2 = ma(t^*)^{2H}, \quad \theta_{t^*} = \frac{dC + Ct^* - mt^*}{ma(t^*)^{2H}},$$

$$t^* = \frac{HdC}{(1-H)(C-m)} \text{이다.}$$

3. 시뮬레이션

3.1 시뮬레이션 모델

기존의 논문들에서 자기유사 트래픽을 발생시키는 많은 방법들이 제안되었다[7],[10],[11],[12]. 본 논문에서는 시뮬레이션을 위해 참고문헌 [7],[10]에서 제안한 Fractional Gaussian Noise의 Fast Approximation 발생 방법을 사용하는 fft_fgn 발생기를 이용한다.

(그림 1)은 시뮬레이션 모델을 나타내며, Single-Server Single-Queue로 구성된다. 하나의 트래픽 소스는 하나의 출력링크에 의해 서비스되는 버퍼에 입력된다. 시뮬레이션에서 설정되어야 할 파라미터로는 네트워크 구성 측면에서는 버퍼 크기(B)와 출력 링크 용량(C)이고 트래픽 소스의 측면에서는 평균 비트율(m), 분산계수(a), Hurst 파라미터(H)이다. 각각의 파라미터들은 시뮬레이션의 입력 파라미터로 사용되며, 셀 손실 확률(CLR : Cell Loss Rate) 및 셀 지연 확률(Cell Delay Probability)은 시뮬레이션 출력으로 얻을 수 있는 결과가 된다.

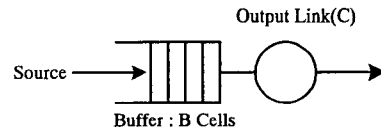


그림 1. 시뮬레이션 모델

3.2 시뮬레이션 결과

시뮬레이션 수행은 결과의 정확성을 위해 각 항목마다 총 8번씩 실시하였으며, 각각의 수행마다 Seed값을 다르게 하여 평균을 구하였다.

(1) 셀 손실 확률 분석

(그림 2)는 분산계수(a)의 변화에 따른 셀 손실 확률을 나타낸다. 여기서 시스템에 입력되는 평균 트래픽은 m=525(cells/sec)이며, 버퍼의 크기는 B=800 (cells), H=0.80, 출력링크의 용량은 C=700(cells/sec)임을 가정하였으며, 분산계수(Coefficient Variance) a는 10에서 100까지 변화시켰다. 해석적인 방법을 시뮬레이션 결과를 비교하여 보면, 분산계수는 셀 손실 확률에 매

우 민감함을 알 수 있으며, Norros Formula의 경우 앞에서 언급하였듯이 버퍼 크기가 낮은 경우에는 시뮬레이션 결과와 많은 오차가 발생하였음을 알 수 있고, Bahadur-Rao Theorem을 사용한 Z. Fan식의 경우는 시뮬레이션 결과와 거의 일치함을 알 수 있다.

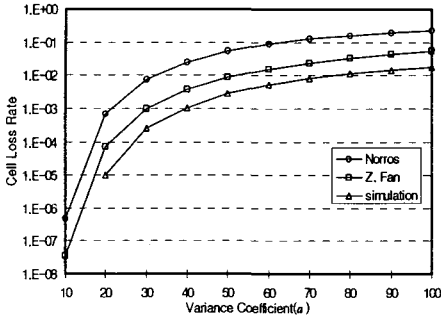


그림 2. 분산계수의 변화에 따른 셀 손실 확률

(그림 3)은 버퍼 크기(B)의 변화에 따른 셀 손실 확률을 나타내었는데 $H=0.83$, $m=525(\text{cells/sec})$, $a=23.9(\text{cell-sec})$, $C=700$ 으로 하였으며, 버퍼크기 B는 300에서 1000까지 변화시켰다. 이 결과는 그림 2의 결과 유사하게 Norros Formula의 경우 낮은 버퍼사이즈에서 시뮬레이션 결과와 많은 오차가 발생하였음을 알 수 있고, Bahadur-Rao Theorem을 사용한 Z. FAN식의 경우는 시뮬레이션 결과와 거의 일치함을 알 수 있다.

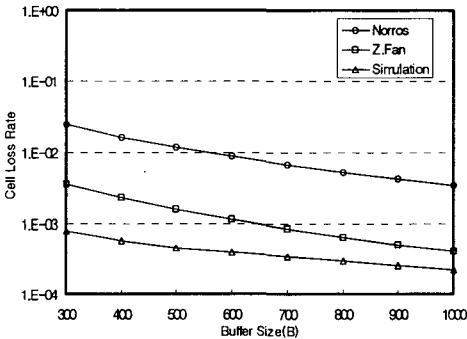


그림 3. 버퍼 크기의 변화에 따른 셀 손실 확률

(그림 4)는 Hurst 파라미터를 0.65 ~ 0.95까지 변화시키면서 셀 손실 확률의 변화를 나타내었는데, $B=800(\text{cells})$, $C=650(\text{cells/sec})$, $m=525(\text{cells/sec})$, $a=23.9(\text{cell-sec})$, 으로 설정하였다. 여기에서도 해석적 결과와 시뮬레이션 결과는 기본적으로 그 추이는 유사하고 이론식이 시뮬레이션 결과보다 큰 Upper-bound를 제공하므로 해석적 모델로 셀 손실 확률이나 유효 대역폭을 예측하여도 문제가 없는 것으로 분석할 수 있다.

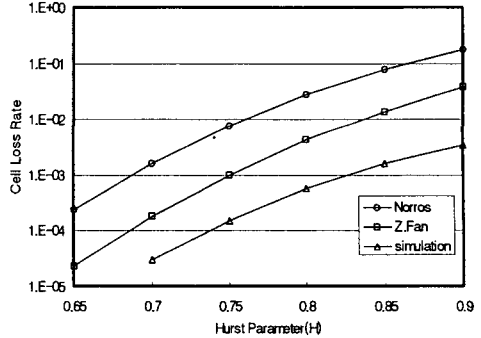


그림 4. Hurst 파라미터의 변화에 따른 셀 손실 확률

(2) 셀 지연 확률 분석

(그림 5)는 지연시간의 변화에 따른 셀 지연 확률의 변화를 나타낸다. $m=525(\text{cells/sec})$, $a=23.9(\text{cell-sec})$, $H=0.83$, $C=600$ 으로 하였다. 그림 5에 의하면, Norros Formula의 경우 앞의 셀 손실 확률의 결과와 같이 시뮬레이션 결과와 많은 오차가 발생하였음을 알 수 있고, Z. FAN의 식이 시뮬레이션 결과와 거의 일치함을 알 수 있다.

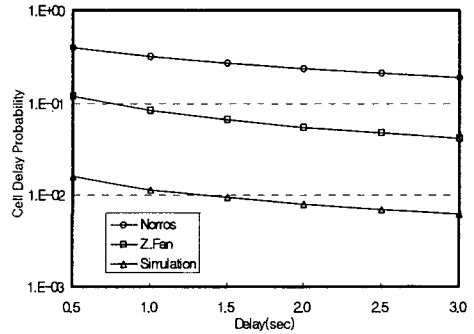


그림 5. 지연시간의 변화에 따른 셀 지연확률

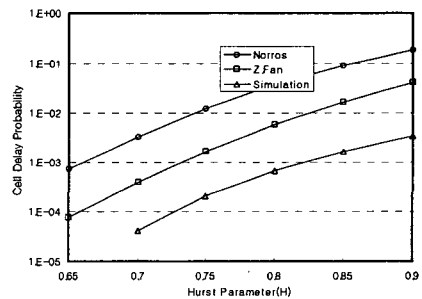


그림 6. Hurst Parameter의 변화에 따른 셀 손실 확률

(그림 6)은 Hurst 파라미터를 0.65 ~ 0.95까지 변화시키는 경우 셀 지연 확률의 변화를 나타낸다. 여기서 $d=1(\text{sec})$, $C=650(\text{cells/sec})$, $m=525(\text{cells/sec})$, $a=23.9(\text{cell-sec})$ 으로 설정하였다. 여기에서도 해석적 결과와 시뮬레이션 결과는 기본적으로 그 추이는 유사하

고 기존의 Norros와 Z. FAN의 해석적 모델이 시뮬레이션 결과보다 큰 Upper-bound를 제공함을 알 수 있다. 시스템 설계시 추후의 확장성이나 마진을 고려한다면, Z.FAN의 유효대역폭 산출식을 이용하여 셀 손실 확률이나 지연확률 및 유효 대역폭을 산출하여 시스템의 망 설계 및 디멘전닝에 적용하여도 큰 문제가 없을 것으로 분석된다.

4. 결론

이동통신시스템이 기존 음성 위주의 2G 시스템에서 고속 데이터 서비스 중심의 3G 시스템으로 Migration함에 따라 데이터 트래픽에 대한 특성 분석이 중요한 문제로 대두되고 있다. 데이터 트래픽은 음성 트래픽과는 달리 Burstiness와 Self-similarity를 가지고 있기 때문에 시스템 설계 및 디멘전닝 시에 기존의 음성과는 다른 방법이 요구되는 것이다.

본 논문에서는 이와 같은 고속 데이터 통신의 망 설계 및 시스템 디멘전닝 시 데이터 트래픽의 버스트성과 자기유사성을 고려함과 아울러 지연이 고려된 유효 대역폭을 산출식을 유도하였다.

시뮬레이션 결과에 의하면 두 가지 해석적 모델에 의한 결과는 기본적으로 시뮬레이션 결과와 추이가 비슷하고, 시뮬레이션 결과보다 큰 Upper-bound의 셀 손실 확률과 지연 확률을 제공함을 알 수 있었다. 또한, 두 해석적 모델을 비교해보면 Z. Fan's Formula가 Norros's Formula보다 시뮬레이션 결과에 좀 더 근접함을 알 수 있었다. 따라서, 3G 이동통신시스템의 망 설계 및 디멘전닝 시 추후의 확장성 및 마진을 고려한다면 복잡한 시뮬레이션에 의한 방법보다 Z. Fan's Formula에 의한 해석적 방법을 이용하는 것이 지연 시간을 만족하는 채널용량 및 중계선 용량, 그리고 시스템 내 다양한 디바이스의 Buffer 크기 등을 보다 쉽게 산출하는 Solution으로 활용될 수 있을 것이다.

시스템의 버퍼크기를 설계 시 지연시간과 시스템의 유효대역폭을 만족하도록 설계해야 할 것이며, 시스템의 유효대역폭 결정시 지연시간이 1초, 2초를 초과할 지연확률을 5%이하로 고려하고 있다.

추후로는 유무선 통합망에서 모델링한 가입자 트래픽 모델을 근간으로 하여 다양한 환경에서 가입자당 유효대역폭을 분석 및 시뮬레이션 해야 할 것이며, 이를 바탕으로 3G 고속 데이터 시스템인 1xEV-DO/DV의 망 설계 및 디멘전닝을 수행할 예정이다.

참고 문헌

[1] W. Willinger, "Self-similarity in wide-area network traffic," IEEE LEOS '97, vol.2, pp.462-463, 1997.
 [2] M. Garrett, W. Willinger, "Analysis Modeling and

Generation of Self-similar VBR Traffic," Proc. of SIGCOMM '94, August 1994.
 [3] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the Self-similar Nature of Ethernet Traffic," Proc. of SIGCOMM '93, Sept. 1993.
 [4] M. Crovella and A. Bestavros, "Self-similarity in World-wide Web Traffic: Evidence and Possible Causes," Proc. of ACM Sigmetrics Conference on Measurements and Modeling of Computer Systems, May 1996.
 [5] V. Paxson, S. Floyd, "Wide area traffic: the failure of Poisson modeling," IEEE/ACM Transactions on Networking, vol. 3, no.3, pp.226-244, 1995.
 [6] A. Patel and C. Williamson, "Effective Bandwidth of Self-Similar Traffic Sources: Theoretical and Simulation Results," Proceedings of the IASTED Conference on Applied Modeling and Simulation, Banff, AB, July 1997
 [7] Schuler, C. "fft_fgn". Research Institute for Open Communication systems, GMD FOKUS, Hardenbergplatz 2, D-10623 Berlin, Germany.
 [8] Leonard Kleinrock, "Queueing System," Volume 1: Theory, Wiley, 1975.
 [9] Z. Fan, P. Mars, "Accurate Approximation of Cell Loss Probability for Self-similar Traffic in ATM Networks," Electronics Letters, Vol. 32, No. 19, pp. 1719-1751, Sep. 1996.
 [10] V. Paxson, "Fast Approximation of Self-Similar Traffic," Technical Report LBL-36750, Lawrence Berkeley Laboratory and EECS Division, University of California, Berkeley, April 1995.
 [11] P. Pruthi, " Heavy-tailed ON/OFF Source Behavior and Self-Similar Traffic," ICC'95, 1995
 [12] J.S. Li, A. Wolisz, R.P.Zelein, "Fast Simulation of Selfsimilar Traffic," IEEE/ACM, pp1829-1833, 1998.



임 석 구

1983년 2월 : 한국항공대학교
 항공전자공학과(공학사)
 1987년 2월 : 서울대학교
 공과대학 전자공학과(공학석사)
 1999년 2월 : 한국항공대학교
 항공전자공학과(공학박사)
 1987년 1월 ~ 1992년 2월 : LG 정보통신
 1992년 2월 ~ 1994년 2월 : 한국전자통신연구원
 1994년 3월 ~ 2001년 2월 : 주성대학
 2001년 3월 ~ 현재 : 천안대학교 정보통신학부
 조교수

<관심분야> 트래픽 엔지니어링, 이동통신시스템
 성능분석, 광 네트워크