데이터 트래픽에서의 Self-Similar 특성

김창호*, 황인수*, 최삼길*, 김동일*, 이동철**, 박기식**

*동의대학교, **한국전자통신연구원

Self-Similarity Characteristic in Data traffic

Chang-ho Kim*, In-Soo Hwang*, Sam-gil Choi*, Dong-ill Kim*

Dong-chul Lee**, Ki-shik Park**

*Dongeui University, **ETRI*

E-mail : kimch@hyomin.dongeui.ac.kr

요 약

기존의 큐잉 분석은 네트워크 용량설계 및 성능 예측을 할 때 상당히 유용하지만, 대부분의 실제 경우에서 큐잉 분석으로부터 예측된 결과는 실제 관측된 성능과는 상당히 다르다. 특히, 최근 몇 년 동안 몇몇 환경에서 트래픽 패턴이 Poisson 패턴보다는 차라리 Self-similar 패턴에 더 가깝다는 것이 입증되어 왔다. 본 논문에서는 이러한 Self-similar 확률과정의 정의 및 트래픽의 특성, 그리고 최근 논문들에서 보고된 사례에 대해서 고찰해 봄으로써, 실제의 데이터 트래픽 특성에 대해 더 나은 이해를 제공하고자 한다.

ABSTRACT

The classical queuing analysis has been tremendously useful in doing capacity planning and performance prediction. However, in many real-world cases. it has found that the predicted results form a queuing analysis differ substantially from the actual observed performance. Specially, in recent years, a number of studies have demonstrated that for some environments, the traffic pattern is self-similar rather than Poisson. In this paper, we study these self-similar traffic characteristics and the definition of self-similar stochastic processes. Then, we consider the examples of self-similar data traffic, which is reported from recent measurement studies. Finally, we wish you that it makes out about the characteristics of actual data traffic more easily.

I. 서 론

Self-similarity는 널리 알려진 서로 다른 이론(fractal과 chaos)과 관련된 개념이다. 본 논문에서는 self-similarity에 대한 기본개념을 설명하고, 그 다음 self-similarity의 존재 및 핵심적인 특성을 포함해, self-similar 데이터 트래픽에 대해 살펴본다. 또한 Poisson 트래픽과 비교되는 이러한 형태의 트래픽의 성능관계에 대해 살펴보고, self-similar 트래픽을 모델링하고 키 파라미터를 분석한다.

II. Self-similarity

self-similarity에 대한 이해를 돕기 위해 인위적인 예를 들어보도록 하겠다[1]. 1-Mbps frame relay 라인과 4000 bit로 고정된 길이의 프레임을 모니터링 한다고 가정한다. 그리고 각 프레임이 전송되는데 걸리는 시간이 4㎳라고 한다. 또한, 다음과 같은 도착시간(각 프레임의 첫 bit가 도착하는 시간)이 수신기에서 기록되어졌다고 가정한다.

0	8	24	32	72	80	96	104
216	224	240	248	288	296	312	320
648	656	672	680	720	728	744	752
864	872	888	896	936	944	960	968

여기서 어떠한 패턴이나 통계적인 특성을 식별하기는 어렵다. 그러나, 이 트래픽은 데이터 트래

픽에서 예상했던 것처럼 버스트(burst)하게 보인다. 몇 개의 도착시간들은 서로 군집되어 있고 어떤 부분들은 약간의 갭이 있다. 이중에 5개 프레임 시간(20ms)이하의 갭들을 모아서(aggregate) 어떠한 프레임의 그룹이 되는 한 클러스터로 간주한다고 가정하고, 각 클러스터의 시작 시간을 기록하면 다음과 같다.

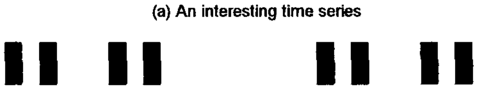
0 72 216 288 648 720 864 936

여전히 어떠한 패턴을 관찰하기는 어렵다. 따라서 집합(aggregation)의 정도를 더 크게 해본다. 즉, 10개의 프레임 시간(40ms)이하의 갭을 모아서 더 큰 클러스터를 정의한다. 그러면 다음과 같은 도착시간이 나온다.

0 216 648 864

이 경우에 그 갭들은 216, 432, 216이 된다. 그 패턴은 작은 갭에 의해 두 개의 클러스터가 반복되고, 큰 갭에 의해 나뉘는 두 개의 클러스터가 된다. 이전의 8개의 클러스터로 된 집합을 살펴보면, 이러한 패턴이 반복된다는 것을 알 수 있다. 처음 4개의 도착 시간들은 '도착, 짧은 갭, 도착, 긴 갭, 도착, 짧은 갭, 도착'의 패턴을 형성하고, 마지막 4개의 도착 시간들도 똑같이 반복된다. 원래의 32개의 도착 데이터 집합으로 돌아가서 살펴보면, 이러한 동일한 패턴이 8번 반복되는 것을 알 수 있다. 그러므로 그 패턴은 원래의 데이터와 같은 모양을 가지는 서로 다른 레벨의 집합이 반복된다는 것을 알 수 있다. 이것은 타임 시퀀스가 해상도(resolution)의 정도에 상관없이 동일한 패턴으로 나타나는 것이다. 이것이 self-similarity의 핵심이다.


(a) An interesting time series


(b) Aggregated time seriese with level 1(20ms)


(c) Aggregated time seriese with level 2(40ms)

그림 1. Self-similar 시계열(time series).

self-similarity는 단지 최근에 들어서야 데이터 통신 트래픽 분석에 적용된 중요한 개념이다. 또한 self-similarity가 도처에 산재해 있다는 것이 밝혀졌다[2]. self-similarity는 차원(dimension)상의 서로 다른 확대비율이나 서로 다른 스케일에서 보았을 때 동일하게 보이거나 동일하게 행동하는 자기 유사한 현상이다.

앞에서 예로든 패턴은 그림으로 보는 것이 좀 더 이해하기 쉽다. [그림 1.a]는 시간에 따른 프레임의 연속 도착을 그린 것이다. 각각의 수직 라인들은 수신기가 첫 bit에서 마지막 비트까지 전체의 프레임을 흡수하는데 걸리는 시간인 4ms에 비례하는 폭을 가진 한 개의 프레임을 나타낸다. [그림 1.b]는 4개의 큰 클러스터로 모아진 데이터를 보여준다. 이 그림에서 '도착, 짧은 갭, 도착, 긴 갭, 도착, 짧은 갭, 도착'의 패턴이 데이터의 서로 다른 해상도에서 나타난다는 것을 보기가 쉽다.

앞의 이러한 인위적인 예는 chaos, fractal에 관한 유명한 구조인 칸토르 집합(Cantor set)에서 유래한 것이다. self-similar한 특징들은 실제의 현상에 대해서 무한하게 유지되는 것은 아니다. 그렇지만 아주 큰 범위의 스케일을 통해 대부분의 현상들은 self-similarity를 드러낼 것이다.

비록 이 예가 단순하며 인위적인 예제이긴 하지만, 이러한 고찰을 토대로 self-similar 데이터 트래픽에 대한 어떠한 통찰력을 얻을 수 있다. 아마도, 네트워크 성능의 관점에서 가장 현저한 특징은 집단화의 지속성(persistence of clustering)일 것이다. Poisson 트래픽에서, 집단화는 단기간 내에 발생하지만, 장기간이 지나면 smooth out된다.

III. Self-similar Data Traffic

앞에서 언급한 self-similarity의 형태는 정확한 self-similarity라고 할 수 있다. 주어진 패턴이 정확하게 서로 다른 스케일에서 반복된다. 이러한 정확한 self-similarity는 결정론적인 시계열에 대해 구성되어진다. 그러나 데이터 트래픽은 확률적인 과정으로 가장 잘 고찰되어지고, 통계적인 방법으로만 self-similarity를 언급할 수 있다[1].

일반적으로 결정적이고 주기적인 신호는 시간이동에 관하여 불변인 특징이 있다. 즉 그 신호는 시간상으로 여러 주기가 이동하더라도 동일한 신호이다. 이에 비해, 정상확률과정에 대해서는 그 과정의 통계는 시간이동에 불변이다. 또한 평균은 시간에 독립적이고 자기상관 함수는 단지 시간의 차이에만 의존한다[4].

self-similar 확률과정은 기존의 논문들에서 여러 가지 방법으로 정의되어져 왔다. 본 논문에서는 먼저 연속시간 확률과정에 대해 살펴본 후, 데이터 트래픽과 관련된 이산시간 확률과정에 대해

서 살펴본다[1].

3.1 연속시간 정의

Self-similar 확률과정의 일반적인 정의는 다음과 같이 연속시간 변수의 직접 스케일링에 기초한다. 어떠한 실수 $a>0$에 대해, 확률과정 $a^{-H}x(at)$가 $x(t)$와 통계적으로 동일한 특성을 가진다면, 확률과정 $x(t)$는 파라미터 $H(0.5 \leq H \leq 1)$를 가지고 통계적으로 self-similar하다. 이러한 관계는 다음의 3가지 조건으로 표현된다[3].

$$1.\ E[x(t)] = E[x(at)] \qquad \text{Mean}$$

$$2.\ Var[x(t)] = \frac{Var[x(at)]}{a^{2H}} \qquad \text{Variance}$$

$$3.\ R_x(t,s) = \frac{R_x(at,as)}{a^{2H}} \qquad \text{Autocorrelation}$$

Hurst 또는 self-similarity 파라미터 H는 self-similarity의 핵심척도이다. 다시 말하면, H는 통계적인 현상의 지속성(persistence)에 대한 척도이고 확률과정의 장기간 종속에 대한 척도이다. H=0.5의 값은 self-similarity의 부재를 나타내고, H가 1에 가까울수록, 지속성의 정도 또는 장기간의 종속의 정도는 더욱 커진다. 즉, H=0.5에 대하여 과거와 미래의 증가에 대한 상관성이 없어지고, H>0.5에 대해여 지속성의 두드러진 특징을 가진다.

3.2 이산시간 정의

정상 시계열(stationary time series) x에 대해, m-aggregated 시계열 $x^{(m)}=\{x_k^m, k=0,1,2,...\}$를 인접한 m크기의 블록을 겹침 없이(nonoverlapping) 원래의 시계열을 합계함으로써 정의한다. 이것은 다음과 같이 표현된다.

$$x_k^{(m)} = \frac{1}{m}\sum_{i=km-(m-1)}^{km} x_i \qquad (1)$$

예를 들어, x(3)는 다음과 같이 정의된다.

$$x_k^{(3)} = \frac{x_{3k-2}+x_{3k-1}+x_{3k}}{3} \qquad (2)$$

aggregated 시계열을 관찰하는 한 방법은 타임스케일을 압축하는 기법과 같다. 즉, $x^{(1)}$는 이러한 시계열의 최대 해상도이고, $x^{(3)}$는 비율 3으로 축소한 것이다. 만약 이것에 확률과정의 통계(mean, variance, correlation, etc)가 동일한 압축 사본을 간직하고 있다면, self-similar 과정으로 다룰 수 있다.

$x^{(m)}$의 에르고딕 과정에 대해, 시간평균은 조화평균과 동일하고, 시간평균의 분산은 조화평균과 동일하다. 만약 시간평균의 분산이 m이 매우 커짐에 따라 zero로 수렴하게 되면, 이것은 self-similar 과정이 아니다. 만약 확률과정 x가 모든 m=1,2,...에 대해서 다음과 같다면 파라미터 $\beta(0<\beta<1)$에 대하여 정확하게(exactly) self-similar하다고 한다.

$$Var(x_{(m)}) = \frac{Var(x)}{m^{\beta}} \qquad \text{Variance}$$

$$R_{x^{(m)}}(k) = R_x(k) \qquad \text{Autocorrelation}$$

파라미터 β는 앞에 정의했던 Hurst 파라미터 $(H=1-(\beta/2))$와 관련이 있다. 정상과정 및 에르고딕 과정에서는 $\beta=1$이고 시간 평균의 분산은 1/m비율로 감소하게 되지만, self-similar 과정에서는 시간평균의 분산은 더욱 천천히($1/m^{\beta}$) 감소하게된다.

확률과정 x가 충분히 큰 모든 k에 대해 다음과 같다면 근사적으로(asymptotically) self-similar하다고 한다.

$$Var(x_{(m)}) = \frac{Var(x)}{m^{\beta}} \qquad \text{Variance}$$

$$R_{x^{(m)}}(k) \rightarrow R_x(k),\ \ as\ m \rightarrow \infty \qquad \text{Autocorrelation}$$

따라서, 이러한 self-similarity의 정의에 의해서, aggregated 과정의 자기상관은 원래의 확률과정과 동일한 형태를 가진다. 이것은 변이성 또는 버스트한 정도가 서로 다른 타임 스케일에서 동일하게 나타난다는 것을 의미한다.

3.3 장기간 의존성(Long-range Dependence)

self-similar 과정의 중요한 특징들 중의 하나는 장기간 의존성으로 나타난다. 일반적으로, 단기의존과정은 자기공분산이 최소한 지수적으로 급격히 감소한다는 조건을 만족시킨다.

$$C(k) \sim a^{|k|} \quad ,|k| \rightarrow \infty,\ 0<a<1 \qquad (3)$$

일반적으로 기존의 논문에서 고려된 데이터 트래픽 모델들의 형태는 단기간 의존 과정만을 고려했다.

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x},\ |x|<1 \qquad (4)$$

여기서 $\sum_k C(k)$가 유한하다는 것을 알 수 있다. 이에 반해, 장기간 의존과정은 hyperbolical하게 감소하는 자기공분산을 가진다.

$$C(k) \sim |a|^{-\beta} \quad ,|k| \rightarrow \infty,\ 0<\beta<1 \qquad (5)$$

여기서 β는 앞에서 정의된 파라미터이고

H=1-(β/2)이기 때문에 Hurst 파라미터와 연관이 있다. 이 경우 $\sum_k C(k)=\infty$이 됨을 알 수 있다. 그림 2는 자기공분산의 단기간 의존성과 장기간 의존성을 도시한 것이다. 장기간 의존성은 self-similar 과정들에서의 지속적 현상을 반영한다. 즉, 모든 타임스케일에서의 군집 및 버스트한 특성의 존재를 나타낸다.

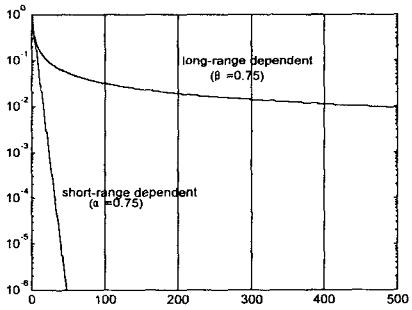


그림 2. 장기간의존성과 단기간 의존성의 비교

3.4 Heavy-tailed Distributions

앞의 aggregated 시계열, 장기 의존성에 관한 앞의 공식은 동등한 정의이다. heavy-tailed 분포는 다소 다른(더욱 포괄적인)특징을 갖고 있지만, 본질적으로, 이러한 분포로 self-similar 확률과정을 정의하는 것이 가능하다. heavy-tailed 분포 접근법의 매력 중 하나는 시뮬레이션 모델들을 다루기 쉽게 한다는 것이다.

heavy-tailed 분포는 패킷의 도착(interarrival) 시간 및 버스트 길이와 같은 트래픽 과정을 설명하는 확률의 특성을 나타내기 위해 사용되어 질 수 있다. 랜덤변수 X의 분포가 다음과 같다면 heavy-tailed하다고 한다.

$$1-F(x)=\Pr[X>x]\sim\frac{1}{x^{\alpha}} \quad as\ x\rightarrow\infty,\ 0<\alpha \qquad (7)$$

일반적으로 heavy-tailed 분포를 가진 랜덤변수는 높은 분산(심지어 무한대의 분산)을 갖는다. 가장 단순한 heavy-tailed 분포는 파라미터 k와 α(k, α>0)를 가진 Pareto 분포이다. 밀도함수 와 분포함수는 다음과 같다.

$$f(x)=\ F(x)=0 \quad (x\le k) \qquad (8)$$

$$f(x)=\frac{\alpha}{k}\left(\frac{k}{x}\right)^{\alpha+1},\ F(x)=1-\left(\frac{k}{x}\right) \quad (x>k,\ \alpha>0) \qquad (9)$$

따라서 기대값은 (10)식과 같이 나타난다.

$$E[x]=\frac{\alpha}{\alpha-1}k \quad (\alpha>1) \qquad (10)$$

파라미터 k는 랜덤변수가 취할 수 있는 최소의 값을 지정한다. 파라미터 α는 랜덤변수의 기대값 및 분산을 결정한다. 만약 $\alpha\le 2$이면, 분포함수는 무한 분산을 가지고, $\alpha\le 1$이라면, 무한한 기대값과 분산을 가지게 된다. 그림 3은 log-linear 스케일의 Pareto 및 지수밀도 함수를 비교한 것이다. 이 그림에서 지수밀도 함수는 거의 직선으로 나타나고, Pareto 분포의 tail은 지수함수보다 매우 더 서서히 감소한다. 그러므로 'heavy tail' 분포 함수라 한다.

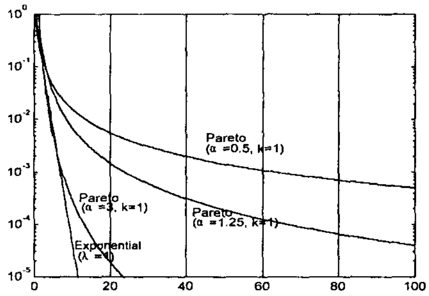


그림 3. Pareto 및 지수 확률 밀도 함수

IV. Self-similar 데이터 트래픽의 예

최근 몇 년 동안 여러 연구에서 실제의 네트워킹환경에서 데이터의 트래픽 패턴이 self-similar 과정에 의해 아주 잘 모델링된다는 것이 증명되었다. 이 장에서는 몇몇 논문에서 발표되었던 두 가지의 self-similar한 데이터 트래픽의 예를 소개한다.

4.1 Ethernet Traffic

Ethernet traffic에서는 기존의 Poisson 트래픽 가정을 사용하는 straightforward 큐잉분석이 모든 네트워크 트래픽을 모델링하는데 적합하지 않음이 밝혀졌고, Ethernet traffic에 대한 새로운 모델링과 분석의 접근법이 대두었다[5].

그림 4는 이러한 Ethernet 트래픽 특성과 Poisson 모델링의 문제점을 단적으로 보여준다. 그림 4.a는 실제의 Ethernet 트래픽을 여러 가지 해상도의 단계에 따라 나타낸 것이다. 여기에서 흥미로운 것은 모든 그림들의 분포적인 관점에서 볼 때, Ethernet 트래픽은 큰 스케일(hours, minutes)이나 작은 스케일(seconds, milliseconds)에서도 그 버스트함이 잘 나타나있다. 즉, 모든 타임스케일에서도 버스트함을 나타낸다.

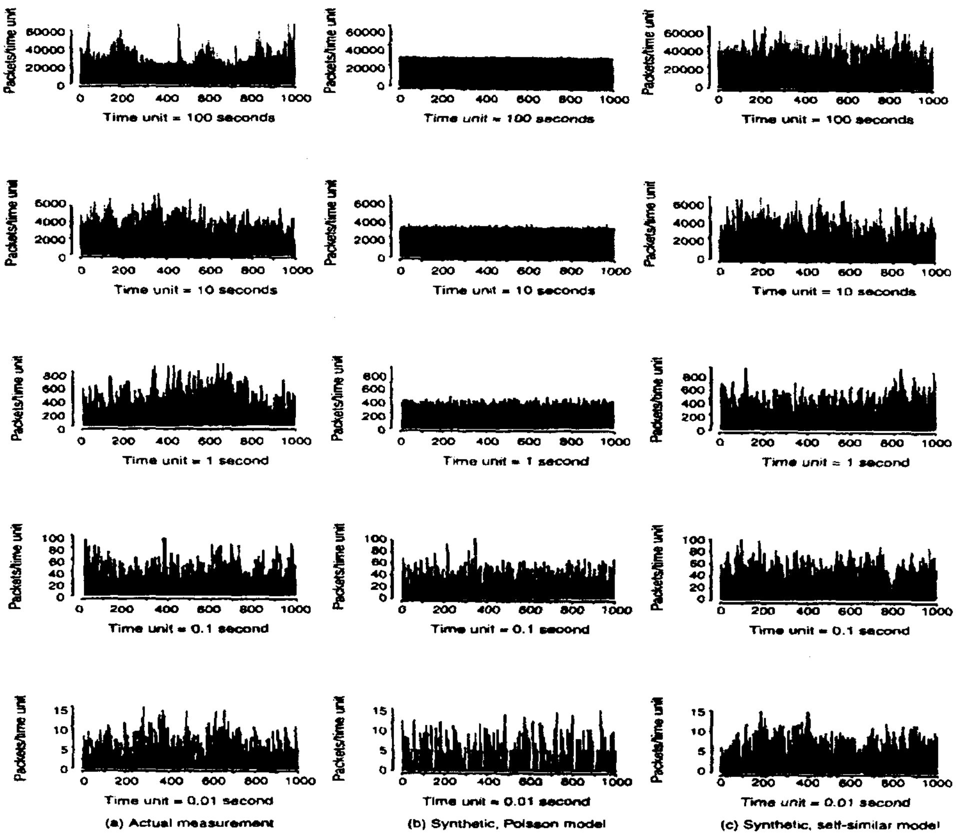


그림 4. 실제 Ethernet 트래픽과 인위적인 트래픽의 비교[6]

이와 대조적으로, 그림 4.b는 Ethernet 그림과 동일한 방식으로 생성되었지만, 합성된 트래픽 데이터를 사용해서 그려진 것이고, 이 데이터는 실재의 데이터와 거의 비슷한 평균패킷 사이즈와 도착율을 구해 Poisson 모델을 사용해서 생성되어진 것이다. 고 해상도(time unit=0.1 seconds)에서 트래픽은 버스트함을 잘 나타내고 있지만, 해상도가 점점 낮아짐에 따라 smooth out해진다. 이것은 Poisson 모델에서 정의된 것처럼 정상 및 에르고딕 확률과정으로부터 기인한 것이다.

그림 4.c 는 실재 Ethernet 트래픽의 Hurst 파라미터를 추정한 값인 H=0.9를 사용한 self-similar 트래픽 모델로 생성된 것이다[6]. 이 그림은 실제의 Ethernet 트래픽의 그것과 일반적으로 동일한 특성을 보여준다. 즉, 모든 타임 스케일에서 Ethernet 트래픽의 버스트함을 아주 잘 보여주고 있다.

4.2 World Wide Web 트래픽

50만개 이상의 Web 도큐먼트들에 대한 요청이 수반된 Web트래픽 연구의 예로써, 웹 브라우저에 의해 생성된 트래픽 패턴이 self-similar 하는 것이 증명되었다. 각각의 브라우저를 ON/OFF 소스로 모델링하였고, 그 데이터가 Pareto 분포와 매우 잘 들어맞고, 다양한 측정을 통해 그것이 1.16~1.5범위의 α값을 가진 Pareto 분포라는 것이 발견되었다[7].

Web 트래픽의 이러한 self-similar 행동은 Web 트래픽이 전송될 때 파일의 랜덤한 선택이 반영됐다고 볼 수 있다. 즉, 사용자들이 Web을 사용할 때 전송되는 파일의 크기에 상관없이 다음링크들에 의해 파일들을 선택하기 때문에 기인한 것으로 볼 수 있다. 즉, Web상에 분포된 파일들의 크기에 의해 전송되는 트래픽 패턴 결정되기 때문이다. 실제로 Internet을 통한 Web 파일들의 전송은 heavy-tailed 크기 분포를 가진다는 것이 발견되었다.

V. Self-similar 트래픽의 성능관계

이러한 데이터 트래픽의 self-similar한 특성이 실제의 성능에 어떠한 영향을 미치는가에 대한 문제가 제기되어왔고, 많은 연구가 행해져 왔다. 그중 가장 중요한 문제는 self-similarity가 성능에 막대한 영향력을 미친다는 것이다.

특히, Ethernet에 대한 중요한 발견은 Ethernet 상의 부하가 점점 더 높아질수록, 계산된 Hurst 파라미터 H가 더욱 높아지거나, 이에 상응하는 self-similarity의 정도가 더 높아진다는 것이었다. 이러한 결과는 성능의 쟁점에서 가장 관련 있는 높은 부하상태에서 정확하기 때문에 극히 중요한 것이다. 또한 Ethernet 분석에 대한 중요한 결과는 성능을 예측하기 위한 전통적인 큐잉모델들이 부적당하다는 것이었다[5]. 예를 들어, 데이터 트래픽에 관련된 일반적인 가설은 많은 수의 독립적인 트래픽 스트림들을 멀티플렉싱하는 것이 Poisson과정으로 귀착된다는 것이다. 기존의 이러한 가정과 결과적인 큐잉분석이 초창기의 ATM switch 제조업자들이 매우 적은 버퍼(10~100cells)를 가진 1세대 스위치를 생산하도록 했고, 이러한 스위치들이 현장에 배치되고 실재의 트래픽을 수용했을 때, 기대한 범위를 훨씬 넘어선 cell손실들이 발생하게 되었고 그 스위치들을 다시 설계하도록 하는 결과를 초래했다. 이를 예로 알 수 있듯이, 입력이 self-similar하다면, 증가되는 지연과 증가되는 버퍼 사이즈의 요구조건은 self-similar 스트림의 어떠한 멀티플렉싱으로 나타날 것이라는 결론이 나온다[5]. 이것은 ATM, frame relay와 100BASE -T와 같은 스위치들과 WAN 라우터들, Ethernet과 같은 공유매체 LANs 그리고 통계학적인 멀티플렉서들에게도 적용된다.

VI. 결 론

본 논문에서는 self-similar의 정의와 self-similarity를 나타내는 특성파라미터를 고찰 보았고 데이터 트래픽에서의 실제 예를 통해 기존의 Poisson 큐잉이론을 통한 분석이 self-similar 특성을 가진 트래픽 분석에는 적합하지 않음을 설명했다. 그러나 전통적인 큐잉분석이 이제 부적절해졌다는 의미는 아니다. 즉 self-similarity가 모든 데이터 트래픽에서 적용될 수 있는 것은 아니다. 어떤 때는 적절하고 어떤 때는 적절하지 않을 수 있다[8]. 이러한 문제는 아직 활발히 진행중인 연구과제 이기도 하다.

기존의 큐잉 이론이나 Poisson 가정의 경우 단기간 의존성만을 고려한 것으로 self-similarity의 장기간 의존성을 나타내기 위해서는 새로운 파라미터 및 분포함수가 필요하다는 것을 알았다. 또한 Ethernet 상의 부하가 증가할수록 Hurst 파라미터 H는 더욱 높아지고 이에 상응하는 self-similarity의 정도 또한 높아지기 때문에 파라미터 H가 self-similarity를 나타내는 중요한 파라미터 임을 알 수 있었다. 그리고 이러한 H 파라미터는 정확한 추정하는 방법을 결정하기 위해 다양한 접근법이 사용되어져 왔고 앞으로 수행되어야 할 과제 중에 하나이다.

본 논문에서 고찰한 Self-similar 트래픽의 특성을 토대로 하여, 추후 연구과제로서는 실제의 트래픽을 측정한 결과로 직접 Hurst 파라미터를 추정해 보고, 그 결과를 토대로 NS(Network Simulator)를 사용해 시뮬레이션한 결과를 토대로 Self-similar 트래픽 모델의 적용성(applicability) 및 타당성 대해서 검증해 볼 것이다.

참고문헌

[1] William Stallings, High-Speed Networks, Prentice Hall,p125-145,p181-207,1997

[2] Schroeder, M.,Fractals, Chaos, Power Laws :*Minutes from an Infinite Paradise.* Freeman, 1991.

[3] Wornell, G. Signal Processing with Fractals: A Wavelet-based Approach. Upper Saddle River, NJ,Prentice Hall,1996

[4] Peyton Z. Peebles,JR. Probability, Random Variables, and Random Signal Principles, McGraw Hill, p134-198,1993

[5] Leland,W.,Taqqu,M.,Willinger,W.,Wilson,D. On the Self-similar Nature of Ethernet Traffic(Extended Version), IEEE/ACM Transaction on Networking, Feb,1994

[6] Willinger,W.,Wilson,D.,Taqqu,M. Self-similar Traffic Modeling for High-speed Networks, ConneXions,Nov,1994

[7] Corvella,M.,Bestavros,A. Self-similarity in World-Wide Web Traffic: Evidence and Possible Causes, Proceedings, ACM Sigmetrics Conference on Measurement and Modeling of Computer Systems, May, 1996

[8] Grossulauser,M. Bolot,J. On the Relevance of Long-range Dependence in Network Traffic, Proceedings, SIGCOMM'96,Aug,1996