

# Burstiness 성질을 이용한 트래픽 측정방법

류승협, 이금모, 장요한, 이재용  
연세대학교 전기전자공학과  
e-mail:dooum@nasla.yonsei.ac.kr

## Traffic Measurement Using Burstiness Property

Seunghyup Ryoo, Geummo Lee, Yohan Chang, Jaiyong Lee  
Dept of Electrical & Electronic Engineering, Yonsei University

### 요 약

인터넷 트래픽은 폭발적으로 증가하고 있으며, 이에 따라 QoS를 보장하기 위한 여러 연구들이 진행되고 있다. 인터넷 트래픽에 대한 연구가 진행됨에 따라 자기 유사성을 가지며, 기존의 poisson을 이용한 모델링은 실제 트래픽과 차이가 있음이 밝혀졌다. 또한, 인터넷 트래픽은 burstiness 성질을 상당히 가지고 있으며, 이는 관측 시간간격에 크게 영향을 받지 않는 것으로 나타난다. 본 논문에서는 이러한 트래픽의 burstiness 성질을 이용한 트래픽 측정 방법을 제시하였다.

### 1. 서론

인터넷은 현재까지 그 수요가 폭발적으로 증가하였으며, 이에 따라 인터넷 환경에서의 QoS(Quality of Service)를 보장하려는 여러 연구가 진행되었다. 인터넷 트래픽의 모델링과 더불어 QoS 보장을 위하여 트래픽에 대한 측정과 모니터링에 대한 연구가 진행되었다.

트래픽을 측정하는 방법은 크게 2가지로 나누어 볼 수 있는데 Active 방식과 Passive 방식이 그것이다. Active 방식은 네트워크에 테스트 패킷을 네트워크에서 측정하려는 구간에 주입하여 실제로 이것이 어떠한 영향을 받는지를 측정하는 방법이다. Passive 방식은 측정하려는 구간에 probe를 설치하고 측정된 데이터를 monitoring station에서 분석하여 결과를 얻는 방법이다.

본 논문에서는 이러한 트래픽 측정을 위하여 인터넷 트래픽이 가지는 burstiness 성질을 이용한 방법에 대하여 논하였다. 최근 연구결과에 따르면 인

터넷 트래픽은 자기유사성을 가지고 있으며, 이로 인해 어떠한 시간 구간을 가지고 관측을 하더라도 burstiness 성질을 가지게 된다고 알려져 있다[1-4]. 이를 이용하여 측정 대상 트래픽의 상대적인 burstiness 구간을 찾아내고 이 구간에서 샘플링을 통한 트래픽 측정방법을 제시하였다.

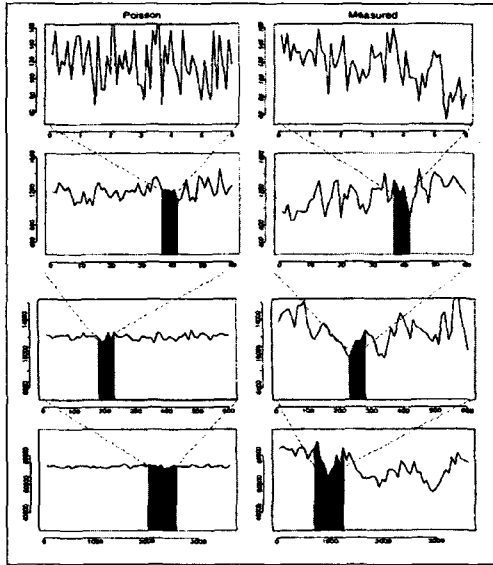
### 2. 인터넷 트래픽의 특성

예전부터 네트워크 트래픽에 대한 여러 모델링 방법이 연구되어 왔으며, 그중에 가장 널리 이용되었던 모델은 poisson 트래픽 모델이었다. 하지만 최근의 연구는 이러한 poisson 트래픽 모델에 대한 한계를 제시하였는데 이는 poisson 트래픽으로 나타낼 수 없는 실제 트래픽의 자기 유사성에 관한 부분이다.

Poisson 트래픽과 실제 트래픽을 측정하여 본 결과 두 가지의 트래픽은 차이를 나타내는 것을 볼 수 있다. 다음의 <그림 1>은 시간 간격에 따른 트래픽의 양을 나타낸 측정 결과이다[1]. Poisson 트래픽은 시간 간격이 커짐에 따라 burstiness 성질이 확연하게 줄어들지만, 인터넷 트래픽의 경우는 시간 간격

본 연구는 한국과학재단 목적기초연구  
(R01-2002-000-00531-0) 지원으로 수행되었음.

이 커지는 것에 상관없이 어느 정도의 burstiness 성질을 보여주고 있으며, 자기 유사성을 가지는 것을 볼 수 있다.



<그림 1> 인터넷 트래픽의 자기 유사성

인터넷 트래픽은 자기 유사성외에도 다음과 같은 특징을 가진다. 첫 번째로, LRD(Long Range Dependence)라는 성질을 가진다. LRD 특성을 가지는 트래픽은 autocovariance가 hyperbolic하게 감소하는 현상을 보여준다. 두 번째는, heavy-tailed 분포를 가진다는 특징이다. 이는 트래픽이 burstiness 성질을 가질 때의 특징으로 알려져 있다.

트래픽의 burstiness는 패킷이 평균값보다 많이 들어오거나 혹은 적게 들어오는 현상을 보이는 시간적인 구간이 상대적으로 길고 자주 나타나는 것으로 볼 수 있다. 특히, 어떤 시간구간에서 패킷이 평균값보다 크게 들어오게 되면 이는 네트워크에 영향을 줄 수 있다.

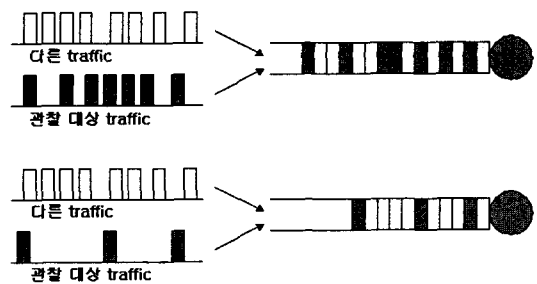
트래픽의 burstiness가 심한 경우에는 네트워크에 어떤 부분에서 혼잡을 유발할 수 있으며, 이로 인하여 다수의 패킷이 한꺼번에 loss될 수도 있다. 이러한 loss를 방지하기 위해서는 평균값보다 상당히 큰 버퍼 공간을 필요로 하는 단점도 유발한다. 앞의 <그림 1>을 다시 살펴보면, 인터넷 트래픽의 burstiness는 poisson 모델과는 달리 버퍼링에 의한 smoothing 효과가 작음을 알 수 있다.

### 3. 트래픽 burstiness와 샘플링

트래픽을 측정하는데에는 대부분의 경우 샘플링을 이용하여 트래픽의 일부를 샘플링하여 하게 된다. Passive 방식은 측정할 데이터량이 많을수록 이를 처리하는데 많은 부담을 가져오며, Active 방식도 네트워크에 삽입한 테스트 패킷이 많을수록 이의 처리와 네트워크에 대한 부담이 가중된다. 따라서, 트래픽의 전체에 대하여 측정하였을 때보다 정확도가 떨어지더라도 트래픽의 일부분에 대한 측정을 하게 된다.

샘플링 방법은 기본적으로 time-based 방식과 count-based 방식이 있다. Time-based 방식은 일정한 시간 마다 패킷을 샘플링 하는 방법으로 구현하기가 간단하지만, 트래픽의 burstiness와 같은 특성을 정확히 반영하지 못할 수 있다. Count-based 방식은 일정 개수의 패킷 마다 샘플링을 하는 방법으로 트래픽의 burstiness 특성을 어느 정도 반영한다.

본 논문에서 사용하고자 하는 트래픽 측정방법은 Active 방식에 기반을 두고 있으며, 측정대상이 되는 트래픽의 burst한 부분에 테스트 패킷을 삽입하는 방법이다. 이러한 방법은 시간상으로 상당히 인접한 패킷들은 상대적으로 인접하지 않는 패킷들보다 네트워크에서 더욱 유사한 환경을 거쳐 갈 확률이 높다는 가정을 전제로 한다. 즉, 삽입된 테스트 패킷의 수가 적어도 어느 정도의 정확도를 유지할 수 있게 된다. 다음의 <그림 2>는 네트워크에서 하나의 큐를 살펴본 것이다.



<그림 2> 하나의 큐에서 인접한 패킷들과 상대적으로 인접하지 않은 패킷들

큐에 도착하는 패킷들이 시간적으로 인접한다면, 그렇지 않은 패킷들과 비교할 때 큐에서 차지하는 위치 역시 상대적으로 인접할 확률이 높다고 볼 수 있다. 이는 다른 트래픽에 대한 영향을 상대적으로 덜 받을 가능성이 높음을 의미한다. 즉, 인접한 패킷

들은 상대적으로 그렇지 않은 패킷들보다 큐에서의 delay와 같은 특성의 차이가 적을 확률이 높음을 예측할 수 있으며, 이는 burstiness 부분에서의 테스트 패킷의 특성은 그와 인접한 여러 패킷과 유사한 환경을 거칠 확률이 높으므로, 이들을 대표하는 목적으로 사용될 수 있음을 의미한다.

**4. Burstiness를 이용한 트래픽 측정방법**

본 논문에서는 다루고 있는 트래픽 측정방법은 [5]에서 제안한 Embedded Active 모니터링 방식을 기반으로 하여 3장에서 제안한 샘플링 방법을 적용한 방식이다. 이를 위하여 상대적으로 burst한 부분을 찾아내는 트래픽을 측정이 필요하다. 트래픽의 burstiness한 부분은 다음과 같은 과정을 통하여 트래픽 측정 구간에서 트래픽이 보내지는 부분에서 탐지되며, 이 부분에서 측정을 위한 테스트 패킷이 발송된다.

**Step 1 :** 측정을 시작한 시간  $t_s$ 를 기록하고 이후로 측정 대상이 되는 트래픽의 n번째 패킷이 도착할 때의 시간  $t_{old}$ 을 기록한다. 이를 가지고 n개 패킷이 도착한 시간 간격을 구하여 구간 평균 트래픽률  $R_a = \frac{n}{t_{old} - t_s}$ 을 계산한다. n값은 관리자에 의해 경험적으로 지정될 수 있다.

**Step 2 :** 다시 패킷을 카운트 하여 n번째 패킷이 도착하면 패킷을 샘플링하고 그때의 시간  $t_{new}$ 를 기록하고, 현재까지 도착한 패킷의 개수  $n_{total}$ 을 계산한다.  $n_{total}$ 은 n의 정수배가 된다. 그리고, 측정이 시작된 시간  $t_s$ 와 이전의 측정 시간인  $t_{old}$ 를 이용하여 평균 트래픽률  $R_a$ 와 구간 평균 트래픽률  $R_i$ 를 다음과 같이 계산한다.

$$R_a = \frac{n_{total}}{t_{new} - t_s}, R_i = \frac{n}{t_{new} - t_{old}}$$

**Step 3 :** 평균 트래픽률과 현재 구간의 트래픽률을 비교하여 다음의 조건을 만족하면 테스트 패킷을 발송한다. 조건에 만족하지 않으면 step 2에서 샘플링한 패킷을 버린다.

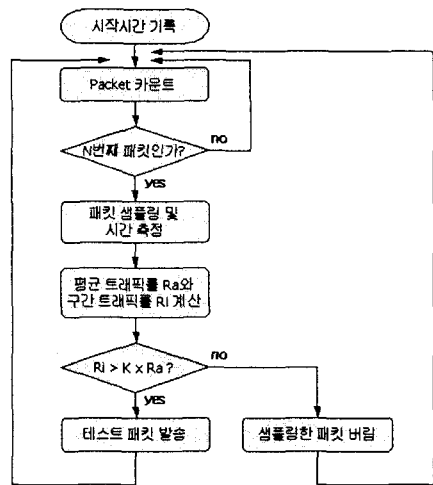
$$R_i > kR_a$$

이때의 k값은 현재 구간의 트래픽률이 평균 트래픽률에 대하여 어느 정도 높을 때 active packet이 발송될 것인가에 대한 기준이 된다. 이는 관리자에 의

해 경험적으로 선택될 수 있다.

테스트 패킷은 step 2에서 샘플링한 패킷의 헤더를 가지며, 그때의 시간  $t_{new}$ 와 이전 측정과의 시간 간격  $t_{new} - t_{old}$ , 평균 트래픽률, 테스트 패킷이 속하는 구간의 트래픽률을 payload로 한다.

**Step 4 :**  $t_{old}$ 에  $t_{new}$ 값을 넣어 갱신하고 step 2부터의 과정을 측정이 끝날 때까지 반복한다. 과정을 반복함에 따라 평균 트래픽률은 계속 갱신되게 된다. 이에 따라 평균 트래픽률에 기초한 burstiness를 판단하고 측정을 위한 테스트 패킷을 발송한다. 다음의 <그림 3>은 앞의 과정을 나타낸 것이다.



<그림 3> 제안한 트래픽 측정 방법의 순서도

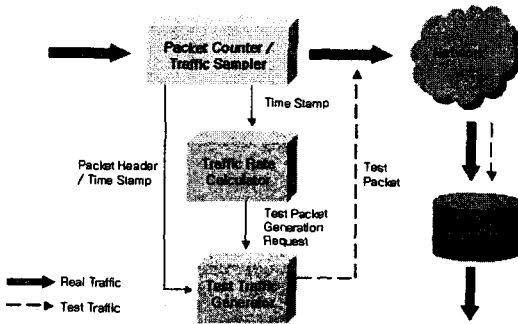
앞서 설명한 과정을 거쳐 네트워크로 보내진 테스트 패킷은 트래픽 측정 구간을 통과하여 트래픽이 도착하는 다른 end부분에서 받게 된다. 이 부분에서는 실제 트래픽에 섞인 테스트 패킷을 골라내고 이것이 담고 있는 timestamp정보를 통하여 delay나 delay jitter등의 정보를 얻어내어 database에 저장한다.

본 논문에서 제안한 트래픽 측정 방법은 패킷 단위의 근사적인 트래픽률을 계산한다. 발송된 테스트 패킷은 측정 대상이 되는 네트워크를 통과하면서 각 라우터와 같은 네트워크 구성원에게 또다른 정보를 제공할 수 있다. 예를 들면, 라우터의 경우 테스트 패킷의 payload부분에 실린 평균 트래픽률과 구간별 트래픽률을 이용하여, 현재의 트래픽이 평균에 비해 어느정도 높은지를 알려줌으로써, 버퍼에서의 drop

확률을 조정하거나, 라우팅 경로를 변경하는 등의 필요한 조치를 취할 수 있다.

제안한 방식의 단점으로는 정확하게 샘플의 수를 결정하기는 상대적으로 어렵다는 것이다. 하지만,  $n$  과  $k$ 값의 설정에 따라 네트워크에 보내지는 테스트 패킷의 수를 대략적으로 조절할 수 있다.

앞서 제안한 트래픽 측정을 위한 트래픽 측정기의 구성을 <그림 4>에 보였다. 트래픽 측정기는 Packet Counter/Traffic Sampler, Traffic Rate Calculator 및 Test Traffic Generator의 3부분으로 구성된다. Packet Counter/Traffic Sampler는 측정 대상이 되는 트래픽의 패킷을 카운트하여  $n$ 개마다의 패킷을 샘플링하고, 이때의 시간을 측정하여 Traffic Rate Calculator에게 보낸다. Traffic Rate Calculator에서는 측정이 시작된 후부터 현재까지의 평균 트래픽률과 현재 구간에서의 평균 트래픽률을 계산하고 비교하여 조건을 만족하면 테스트 패킷을 Test Traffic Generator에 요청한다. 요청을 받은 Test Traffic Generator는 Packet Counter/Traffic Sampler로부터 샘플한 패킷의 헤더와 시간 정보를



<그림 4> 트래픽 측정기의 구성

받아서 테스트 패킷을 만들고 이를 네트워크에 전송한다. 테스트 패킷이 네트워크를 지나서 Monitoring Station에 도착하면 실제 트래픽과 분리되며 delay 와 같은 정보를 얻을 수 있게 된다.

### 5. 결론 및 향후 연구 과제

본 논문에서는 인터넷 트래픽의 특성을 살펴보고 이를 고려한 트래픽 측정방법을 다루었다. 이전의 모델과는 달리 인터넷 트래픽은 관찰하는 시간 간격을 달리 하여도 burstiness 특성이 계속 나타난다. 이러한 인터넷 트래픽의 성질을 이용하여 트래픽의

burstiness 부분에서 테스트 패킷을 삽입함으로써 정확도를 높이려는 시도를 하였으며, 이를 수행하기 위한 트래픽 측정 방법을 제시하였다. 이는 burstiness 부분의 패킷들은 네트워크에서 유사한 환경을 거쳐갈 확률이 높다는 가정에 기초한다.

제안한 트래픽 측정 방법은 기존의 time-based 방식이나 count-based 방식처럼 테스트 패킷의 수를 조절하기에 까다롭고, 측정의 cost를 더 필요로 하지만, 적은 샘플로 비교적 정확한 결과를 나타낼 수 있을 것으로 예상된다. 따라서, active한 측정방법을 사용하는 경우에는 네트워크에 주는 부담을 줄일 수 있을 것이라 생각된다.

향후 연구 과제는 다양한 트래픽의 종류에 따른 효율적인  $n$ 값과,  $k$ 값을 실험적인 방법을 통하여 밝히는 것이다. 또한, 기존의 샘플링 방법보다 상대적으로 정확도와 샘플의 양의 관계를 실험을 통하여 살펴볼 것이다.

### 참고문헌

- [1] Paxson, V. and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," IEEE/ACM Transactions on Networking, 1995
- [2] Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson, On the Self-similar Nature of Ethernet Traffic (Extended Version), IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 2, NO. 1. 1994
- [3] Kihong Park, Walter Willinger, Self-Similar Network Traffic and Performance Evaluation
- [4] 박재성, IP 네트워크에서 인터넷 트래픽의 자기 유사성 및 QoS 제공을 위한 큐 관리 기법, Ph.D Thesis, Yonsei University (2001)
- [5] 강상우, 이금모, 김동근, 이재용, 임베디드 모니터링 기술과 모바일 대행자를 이용한 망 종단간 QoS 파라미터 모니터링, JCCI 2003
- [6] Yuming Jiang et al, "Challenges and approaches in providing QoS monitoring", International Journal of Network Management, 2000
- [7] Thomas Lindh, "A New Approach to Performance Monitoring in IP Networks- combining active and passive methods", PAM2002
- [8] William Stallings, High-Speed Networks and Internets, 2nd Edition, Prentice-Hall