

무선 랜 트래픽의 분석과 모델링

정희원 대쉬도즈 얀힌*, 이 성 진**, 원 유 집***

Modeling and Analysis of Wireless Lan Traffic

Dashdorj Yamkhin*, Seongjin Lee**, Youjip Won*** *Regular Members*

요 약

본 논문에서는 802.11 무선 랜 네트워크 트래픽의 실제 관측 자료에 대한 연구 결과를 보인다. 패킷 트레이스는 대학 캠퍼스의 무선 랜 시설에서 얻은 자료로서 총합된 트래픽(aggregate traffic), 업스트림 트래픽(upstream traffic), 다운스트림 트래픽(downstream traffic), 그리고 TCP 패킷으로만 구성된 통합된 트래픽으로 이 4개의 트래픽 데이터를 수집하였다. 수집한 데이터에서 byte count 프로세스와 packet count 프로세스로 구성된 트래픽의 시계열과 시계열의 주변분포, 그리고 패킷 크기 분포에 대한 분석을 한다. 4개의 모든 데이터의 byte count 프로세스와 packet count 프로세스에서 장기 의존성 성질이 나타났다. 사용자가 인터넷으로 접속하는 업스트림 트래픽의 평균 패킷 크기는 151.7 byte였는데 다른 데이터의 평균 패킷 크기는 모두 260 byte 이상이었다. 최대 크기를 갖는 페이로드(payload)는 업스트림에서 3%, 그리고 다운스트림에서 10%로 나타났다. 이런 분명한 패킷 크기 분포의 차이에도 불구하고 모든 4개의 데이터에서는 허스트(Hurst) 값이 모두 유사하게 나왔다. 허스트 값만으로는 트래픽의 확률적 특성을 충분히 설명할 수가 없다. 트래픽의 특성을 fractional-ARIMA(FARIMA) 그리고 fractional Gaussian noise(FGN)으로 모델링을 한다. FGN은 연산을 하는데 있어서는 더 효율적이었고, FARIMA는 트래픽 특성을 정확하게 모델링하는데 더 좋은 결과를 얻었다.

Key Words : Modeling, Analysis, Self-similarity, Long-range Dependenc, Fractional-ARIMA, Fractional Gaussian Noise

ABSTRACT

In this work, we present the results of our empirical study on 802.11 wireless LAN network traffic. We collect the packet trace from existing campus wireless LAN infra-structure. We analyzed four different data sets: aggregate traffic, upstream traffic, downstream traffic, tcp only packet trace from aggregate traffic. We analyze the time series aspect of underlying traffic (byte count process and packet count process), marginal distribution of time series, and packet size distribution. We found that in all four data sets there exist long-range dependent property in byte count and packet count process. Inter-arrival distribution is well fitted with Pareto distribution. Upstream traffic, i.e. from the user to Internet, exhibits significant difference in its packet size distribution from the rests. Average packet size of upstream traffic is 151.7 byte while average packet size of the rest of the data sets are all greater than 260 bytes. Packets with full data payloads constitutes 3% and 10% in upstream traffic and the downstream traffic, respectively. Despite the significant difference in packet size distribution, all four data sets have similar Hurst values. The Hurst alone does not properly explain the stochastic characteristics of the underlying traffic. We model the underlying traffic using fractional-ARIMA (FARIMA) and fractional Gaussian Noise (FGN). While the fractional Gaussian Noise based method is computationally more efficient, FARIMA exhibits superior performance in accurately modeling the underlying traffic.

* 본 연구는 MOST/KOSEF 의 SRC/ERC 프로그램 (Grant No. R11-2000-073-00000) 의 지원을 받아 수행되었습니다.

* 한양대학교 전자컴퓨터통신공학과 분산멀티미디어 연구실(dashdorj | james | yjwon @ece.hanyang.ac.kr)

논문번호 : KICS2008-04-189, 접수일자 : 2008년 2월 23일, 최종논문접수일자: 2008년 7월 27일

I. 서 론

지난 몇십년간 네트워크와 통신 기술은 인터넷 트래픽에서 급성장하는 분야가 되었다. 통합된 광대역 네트워크는 데이터, 목소리, 그림과 영상 등의 트래픽 종류를 다룰 수 있는 능력이 요구되고 있다. 또한, 이런 서비스에서 생성된 트래픽은 서로 통계적 특성이 다르지만 이 네트워크는 각 세션마다 특정 수준의 처리 속도를 유지하도록 요구되고 있다. 엘르 들어 음성 트래픽은 수십 Kbps의 대역폭을 보장을 해줘야 하지만 지연에는 매우 민감하다. 그렇지만 파일 전송과 같은 고속 데이터 트래픽을 요구하는 랜 또는 광지역 정보 통신망은 수백 Mbps를 보장하여야 하며 패킷 손실에 매우 민감하다. 복잡한 네트워크 트래픽의 경우 정교한 모델링과 분석이 필요하며 공학적으로 매우 자연스러운 접근 방법이다. 과거의 모델링 도구나 기술은 이론적으로나 경험적으로 제한적으로 네트워크의 특성과 행동을 분석하고 이해할 수가 있었다. 하지만 통신 트래픽에서 발견된 free scale 성질은 과거의 기술들에 비해서 데이터의 특성을 더 정확하게 근사화 할 수 있는 방법들을 이끌어 냈다. 결과적으로 자기 유사성 프로세스는 장기 의존성을 나타내는 수문학^[2], 지구 물리학^[6], 생물학^[3], 이동통신 네트워크^[14] 그리고 경제학^[5]과 같은 다양한 과학 분야에서 성공적인 데이터 모델링에 사용되었다.

본 연구의 목적은 패킷 수준의 트래픽의 특성을 포괄적으로 이해하는 것과 시계열의 패킷 크기 분포와 주변 분포와 같은 통계적 특성을 이해하는 것이다. 뿐만 아니라 유선 네트워크에서 만연한 트래픽의 Long Range Dependent 성질이 유선에서도 적용이 됨을 보인다. 본 연구에서 사용된 트래픽 데이터는 실제 운용되고 있는 802.11 네트워크에서 추출된 것이다. 통합된 트래픽에 관한 업스트림과 다운스트림 모두에서 장기 기억 특성을 갖고 있었다. 하지만 업스트림과 다운스트림의 패킷 크기 분포는 서로 다른 양상을 나타냈다. 평균 패킷 크기의 경우 다운스트림 트래픽이 298 바이트의 크기를 갖고 있던 반면 업스트림의 경우 151.7 바이트였다. 이것은 다운스트림의 경우 대부분의 패킷들이 최대 페이로드 크기를 갖고 있었기 때문이다. 또한 본 연구에서는 자기유사성 성질을 갖는 트래픽을 분석하고 이해하기 위하여 부분 자기회귀누적이동평균(fractalional - ARIMA, FARIMA)와 부분 가우스 잡음(fractional

Gaussian Noise - FGN) 모델로 네트워크 트래픽의 장기 기억 성질의 분석적 모델을 제시한다. FARIMA 프로세스가 FGN에 비해 짧은 구간 분석을 하거나 장기 의존성 성질을 분석하는데 더 좋은 특성을 나타냈다.

II. 관련 연구

무선 환경에서의 이동성 패턴의 관점에서의 인간의 행동을 파악하는 것에 초점을 맞춘 논문이 몇개 있다. 이 연구들은 대학교의 캠퍼스, 강당이나 기업 망과 같은 다양한 무선 환경들을 반영한다. Balachandran et al.^[11]은 공공지역에서의 무선 네트워크의 사용자 행동과 네트워크의 성능에 대한 분석을 했다. 무선 사용자에 대한 매개변수 모델을 생성하여 분석과 모의시뮬을 위한 연구 그리고 어플리케이션의 워크로드 분석 결과를 위해 사용하였고 더 나아가 무선 네트워크를 설치하는데 필요한 사용자량 계획이나 잠재적인 네트워크의 최적화를 분석을 위해 사용하였다. Gurtov et al.^[8]은 무선 링크와 트랜스포트 프로토콜 간의 상호작용의 관점에서 무선 링크가 실제와 비등하고 보편적인 모델링을 할 수 있는 트랜스포트 프로토콜에서의 간단한 기법을 개발했다. Park et al.^[18]은 파일 크기에서 두꺼운 꼬리(heavy-tailed)의 정도가 트래픽의 자기 유사성 정도를 직접적으로 나타낸다는 것을 보였다. 허스트 변수 H와 파일 크기 분포에서 지수법칙을 이용해 자기 유사성을 확인해보면 여러 종류의 네트워크 상황에서도 x 와 y 축에 대해서 선형관계를 갖고 있다. 이 관계는 프로토콜 계층에 중요한 영향을 가져다준다. 이 기법은 자기 유사성을 갖는 네트워크 트래픽이 다양한 문헌에서 중요하게 다루고 있는지에 대해 쉽게 이해할 수 있도록 해준다. Mah et al.^[12]은 HTTP에 의해 생성된 네트워크 트래픽의 경험적 모델을 개발했다. 서버나 클라이언트의 로그 기록에 의존하는 대신 HTTP 대화기록을 조사한다. 이 트래픽 분석을 통해서 통계를 얻고 HTTP 파일의 크기, Web 페이지 당 파일의 갯수, 사용자의 검색 방식과 같은 수치들에 대한 분포를 얻어 낼 수 있다. Paxson et al.^[19]은 자기 유사성을 갖는 표본 경로를 부분 가우시안 잡음을 이용하여 가공하였다. 이들이 개발한 방법은 현존하는 다른 방법들보다 빠르며 실제 자기유사성 표본 경로와도 매우 유사한 값들을 얻어 냈다.

Grossglauser et al.^[7]는 최근의 모델링 기법들은 두가지 변수의 영향력에 대해서 고려하지 못하고 있다. 첫 번째는 성능평가와 예측 문제에 있어서 유한한 시간 척도가 가지는 영향이고 그다음은 프로세스의 주변 분포와 같은 일차 통계자료의 영향이다. 이들이 사용한 방식은 유체 유동을 모델링하는 기법으로 유체 비율의 상관 함수는 허스트 변수에 의한 점근적으로 2차 자기 유사성을 보이고 절사 시간지연이 끝날 때 영값을 갖는다.

Tudjarov et al.^[7]는 TCP와 UDP와 같은 다른 종류의 프로토콜들을 분석했고, 상관계수와 공분산 그리고 자기 유사성 정도에 관해서 통계적 분석을 했다. 그 실험의 결과로 트래픽에 대해서 허스트 변수의 값이 0.7-0.75 사이인 것을 확인했다. 최대우도법을 사용하여 추출된 시계열을 사용하여 Pareto와 지수 분포와 같은 기존의 분포들과의 적합도를 검증하였다. 이 때 Pareto 분포는 자기 유사성 프로세스이고 지수분포는 그렇지 않다. 트리 구조의 모델을 지렛대 삼아 multiscale 대기행렬 분석을 한 Vinay et al.^[21]는 이것을 통해 간단하게 대기행렬 확률의 꼬리부분을 근사화할 수 있게 했다. 이 분석은 Multifractal Wavelet 모델에 적용할 수 있을 뿐만 아니라 일반적인 트리구조의 모델과 부분 가우시안 잡음에도 적용이 가능하다. 이들이 한 실험의 결과를 통해서 분명하게 알 수 있는 것은 다양한 시간 해상도에 대한 주변 분포가 대기행렬에 영향이 있고, 가우시안에 대한 가정이 자기 의존성의 성질을 갖고 있다고 하더라도 대기행렬 확률의 꼬리부분을 예측하는데 우수한 성능을 나타낸다는 것을 이끌어 낸다.

본 논문은 802.11을 사용하는 무선 네트워크의 워크로드에 대한 것이고, 이 자료를 통해서 통계적 특성과 허스트 변수를 정하고 무선 기술 분야에 FGN과 FARIMA의 자기 유사성 방법을 통해 모델링하는 것을 목적으로 한다. 이것과 비슷한 목적을 갖고 연구한 논문들이 있는데 이 논문들에서는 무선 네트워크의 자기 유사성을 이끌어 내는 것을 간단하게 설명하고 있다^{[6][17]}. 본 연구는 자기상관함수나 도착 시간 분포에서 높은 자기 유사성 성질을 갖고 있다고 알려진 프랙탈 관점에서 분석을 한다. III장에서는 자기 유사성의 수학적 개념을 설명을 하고 IV장에서는 측정 환경을 소개한다. V장에서는 분석 결과와 트래픽 모델을 설명을 하고 마지막으로 VI장에서 논문의 결론을 맺는다.

III. Synopsis: Long Range Dependent Property

장기 기억 프로세스 또는 자기 유사성 프로세스는 WAN (Wide area network)^[11], LAN (Local Area Network)^[14], variable rate video^{[10][26]} 등과 같은 네트워크 트래픽의 다양한 특성을 설명하기 위해서 사용되었다. 어떤 프로세스가 갖는 특성이 자기 유사성 성질의 여부는 어느 정도는 주관적인 문제이다. 그렇지만 자기 유사성이라는 개념은 실제 현상에 대한 영향력 있는 설명을 할 수 있다고 알려져 있다.

계속 설명하기에 앞서 자기 유사성에 대한 수학적 정의에 대해 먼저 살펴본다. 기초적인 자기 유사성 프로세스의 정의는 다음과 같다. 연속시간 확률 프로세스 X_t 는 어떤 양의 상수 c 에 대해 ct rescaled 프로세스 허스트 변수 $H \sim ((0 < H < 1))$ 을 갖는 강한 자기 유사성을 성질을 갖고 있다.

연속 시간 확률 프로세스 X_t 는 강한 자기 유사성 성질을 갖는 경우는 어떤 양의 stretching 요인 c 와 시간 척도 ct 를 갖는 rescaled 프로세스 $c^{-H}X_{ct}$ 가 원래의 프로세스 X_t 와 동일하고 자기 유사성 변수 H 가 $0 < H < 1$ 의 값을 갖을 때 이다^[23]. 이것은 시계열의 임의의 수열 $t_1, t_2, t_3, \dots, t_n$ 과 $c > 0$ 에 대해서 $c^{-H}X_{ct_1}, c^{-H}X_{ct_2}, \dots, c^{-H}X_{ct_n}$ 는 $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ 와 동일한 분포를 갖고 있다는 말이다^[23]. 이산 시간의 경우, $\{X_k\} = \{X_k; k=0, 1, 2, \dots\}$ 를 평균이 μ 이고 분산이 σ^2 인 정상 프로세스라고 하고 $\{X_k^{(m)}\} = \{X_1^{(m)}, X_2^{(m)}, \dots\}$, ($m=1, 2, 3, \dots$)는 집단 평균이라 하자. 이때 집단 평균은 다음과 같이 표현된다.

$$X_k^{(m)} = (X_{km-m+1} + \dots + X_{km})/m,$$

($k \geq 1$). 자기 상관 함수가 $k \rightarrow \infty$, $0 < \beta < 1$ 일 때 $\rho_k \rightarrow k^{-\beta}$ 를 갖는 X_t 프로세스는 $\rho_k^{(m)} = \rho_k$ 가 $m=1, 2, 3, \dots$ 의 값을 가질 때 정확하게 $H=1-(\beta/2)$ 를 갖는 자기 유사성을 띤다고 한다. 다시 말해, X_k 프로세스와 $\{X_k^{(m)}\}$, $m \geq 1$ 인 평균을 한 프로세스는 똑같은 상관구조를 갖고 있다. $\rho_k^{(m)} \rightarrow \rho_k$ 가 $m \rightarrow \infty$ 로 다가갈 때 X_k 프로세스는 점근적으로 $H=1-(\beta/2)$ 인 자기 유사성을 갖는다^[25].

어떤 프로세스를 자기 유사성을 갖는 프로세스로 만드는 기법들은 여러 가지가 있다. 그 중에 가장

많이 사용되고 있는 방법은 Fractal Gaussian Noise 방법을 사용하는 것이다^[23]. $\{Y_k\} = \{X_{k-X_{k-1}}\}$, $k > 0$ 인 증분 프로세스는 X_k 가 fractional Brownian motion (FBM) random 프로세스일 때 FGN 프로세스라고 한다. FBM 은 분석이 가능한 기법이기 때문에 주로 자기 유사성을 갖는 프로세스를 모델링하는데 많이 사용된다^[16].

자기 유사성 프로세스는 몇가지 성질을 갖고 있다 [24][11]. (i)표본 평균의 분산이 표본 크기의 역수보다 더 천천히 감소하는 Slowly decaying variance 성질을 갖는다. 다시 말해서 c 가 상수이고 $0 < \beta < 1$ 일 때 $\tau \rightarrow \infty$ 로 다가가면 $Var[X_k^m] \rightarrow c\tau^{-\beta}$ 이다^[22]. 이 프로세스의 경우 허스트 변수는 $H = 1 - (\beta/2)$ 를 나타낸다^[26]. (ii)자기 유사성 프로세스의 공분산 구조는 $\gamma_x(t, s) = (\sigma^2/2)\{t^{2H} - |t-s|^{2H} + s^{2H}\}$ 를 갖는다. (iii) 자기 상관 함수 ρ_k 는 무한 합의 성질을 갖고 있다. 다시 말해서 $\sum_{k=0}^{\infty} \rho_k = \infty$ 의 값을 갖기 때문에 총합을 구할 수가 없다. 자기상관계수가 지수승보다 더 급하게 쌍곡선이 감소하는 속도로 감소한다.

자기 유사성을 나타내는 H 값을 얻는 다양한 방법들 중 rescaled adjusted R/S Statistics^{[23][28]}과 variance time plots^[11] 그리고 wavelet analysis^{[26][27]} 등이 있다. 본 연구에서는 첫 번째 것인 R/S plot을 사용하여 트래픽의 H 변수를 측정한다. 수들의 집합 $\{X_1, X_2, \dots, X_n\}$ 이 있다고 할 때 허스트 변수 H는 rescaled adjusted range 방명인 $R(n)/S(n)$ 으로 구할 수 있다. 이 때 $R(n) = \max \{ \sum_{i=1}^k (X_i - \mu), 1 \leq k \leq n \}$ $- \min \{ \sum_{i=1}^k (X_i - \mu), 1 \leq k \leq n \}$ 이다. R/S statistics의 점근적 기울기는 H 변수를 나타낸다^{[23][16]}. 더 자세한 분석은 본 논문의 범위를 넘어가므로 여기까지만 소개한다.

IV. Measurement Setup

그림 1에서는 네트워크의 구성과 연결도를 나타내고 있다. 넓은 지역을 포함하고 비용절감을 하기 위해서는 유선 네트워크를 구성하는 것보다는 무선 네트워크 기술을 이용하여 구성하는 것이 더 효율적이다. 특히 몽고와 같이 한 도시에서도 사람들이 떠엄떠엄 군집을 이루어 모여 사는 경우라면 무선

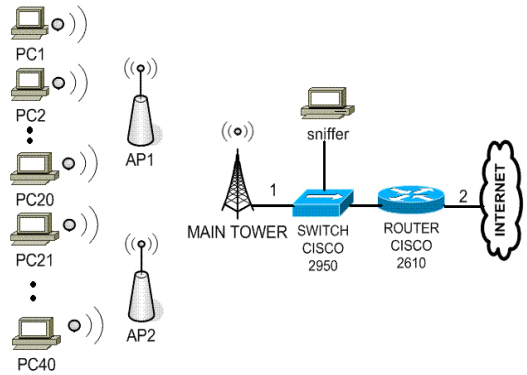


그림 1. Wireless Network Configuration

네트워크를 사용하는 것이 더 좋다. 무선 네트워크에서는 하나의 6개 섹터 안테나 시스템을 사용하는 데 하나의 섹터 안테나는 약 60°의 범위를 담당하고 인접한 안테나의 범위와 약간 겹친다. 40개의 무선 사용자는 서비스 제공자의 2개의 Access Point에 연결이 되어 있다. 모든 연결의 라우팅과 조정 그리고 QoS와 같은 용량의 관리는 라우터가 담당한다. 각 무선 사용자의 전송 용량 범위는 64에서 512 kbps를 갖고 있다.

스니퍼 프로그램을 통해서 패킷 트레이스를 수집했고 이 프로그램 또는 장치는 네트워크에 전송되는 정보를 잡아내는 역할을 한다. 이러한 장비가 이렇게 동작할 수 있는 이유는 이더넷이 처음 정의 되었을 때 공유를 목적으로 하였기 때문에 자신에게 온 메시지가 아니라도 그 메시지를 컴퓨터가 받을 수 있게 만들어져 있기 때문이다. 랜으로 연결된 하나의 컴퓨터는 두개의 주소를 갖게 된다. 하나는 MAC 주소로서 각 노드를 고유하게 구분할 수 있는 것이며 이것은 네트워크 카드에 저장되어 있다. MAC 주소를 사용하여 이더넷 프로토콜이 프레임을 만들어 하나의 장치에서 다른 장치로 데이터를 전송한다. 그리고 다른 하나는 애플리케이션들이 사용하는 IP 주소이다. 데이터 링크 계층에서는 이더넷 헤더와 목적지 장치의 MAC 주소를 IP 주소 대신에 사용한다. 네트워크 계층에서는 IP 네트워크 주소를 데이터 링크 계층이 필요로 하는 MAC 주소로 대응하는 역할을 담당한다. 이것이 하는 일은 주로 ARP 캐쉬로 알려진 포로 저장되어 있는 목적지 장치의 MAC 주소를 먼저 참조한다. 만약 해당 IP 주소로 대응되는 값이 입력이 되어 있지 않다면 Address Resolution 프로토콜은 request 패킷을 broadcast하여 목적지 주소를 가진 장

치가 소스 장치에게 자신의 MAC 주소를 응답하도록 한다. 응답을 받고 나면 그 MAC 주소는 소스 장치의 ARP 캐쉬에 추가 된다. 그리고 이 MAC 주소는 소스 장치에 의해서 목적지 장치와 통신을 하기 위해서 사용이 된다. 트래픽을 수집하기 위해서 사용된 컴퓨터의 운영체제는 윈도우 2000 프로페셔널을 사용하였고 이 컴퓨터에서 10^{-6} 초의 정밀도를 갖는 스니퍼를 사용하여 수집을 하였다. 스니퍼는 네트워크에 연결이 되어 그림 1에서 1과 2로 표시 되어 있는 곳을 지나는 트래픽을 연속적으로 저장한다. 지점 1에서 여러 무선 사용자가 생성하는 트래픽이 한데 모여서 인터넷과 연결이 되는 지점이다. 지점 2에서는 가장 늦게 인터넷과 연결되는 트래픽들이 모여 있다. 모든 패킷들은 TCPDUMP 형식으로 저장이 되었고 12.7 백만개의 패킷들을 수집하였다. 이들 중 70 %는 TCP 데이터그램으로 이루어져 있다.

그림 1에서 나타난 인터넷 접속도는 유선 접속 사용자가 AP를 통하여 네트워크에 정상 접속하는 것을 나타내고 있는데, 이와 같은 구성은 무선 네트워크 트래픽에서 자기 유사성 또는 장기 의존성 성질을 분석하기에 적합한 기본 구성을 제공한다.

V. Traffic Statistics

5.1 Primitive Statistics

패킷 트래픽은 2005년 3월 18일 10시부터 17시까지 수집하였다. 표 1에 수집한 4종류의 데이터 집합을 보인다. 표 2에서는 트래픽의 통계를 나타낸다. 패킷 크기의 평균은 eth.dat, eth.src, eth.dst와 tcp.dat에 대해 각각 267.4, ~ 151.7, ~ 298.5 그리고 270.7 byte 이다. eth.src로 수집된 자료는 다른 세 개의 파일들과 다른 특성을 갖고 있다. 이 트래픽에서 평균 패킷 크기는 151.7 바이트로 다른 3개의 트래픽들의 평균인 267 바이트 또는 그 이상인 것

표 1. File Description of Wireless Network

n/n	Data set	Description of Files
1	eth.dat	Aggregate Traffic Captured at point 1
2	eth.src	Upstream Traffic Captured at point 2
3	eth.dst	Downstream Traffic Captured at Point 2
4	tcp.dat	TCP Traffic Captured at Point 2

표 2. Traffic of Data Sets (10:00-17:00, Mar-18-2005)

Data set	Number of Packets	Packet Size		Inter Arrival time (msec)	
		μ	σ^2	μ	σ^2
eth.dat	12,715,077	267.4	204760	2	1 e-006
eth.src	4,554,667	151.7	80103	.5	6.7 e-005
eth.dst	5,586,555	298.5	239430	4.5	7.1 e-005
tcp.dat	8,468,547	270.7	279430	3	4.5 e-005

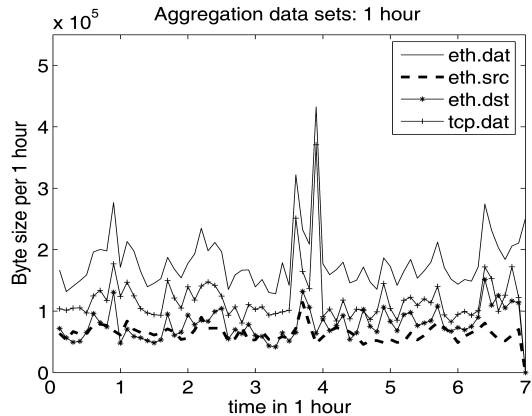


그림 2. Statistical Analysis for Data Sets

에 비해 작다. 패킷 크기의 분산 역시 eth.src이 다른 트래픽들에 비해 작다. 다른 트래픽 트래이스의 33 %의 분산 값을 갖고 있다. 평균 도착간격 시간은 eth.dat, eth.src, eth.dst 와 tcp.dat에 대해 각각 2 msec, 5.5 msec, 4.5 msec 그리고 3 msec 이었다. 그림 2는 1시간 주기로 총합을 낸 트래픽 트래이스에 대한 결과를 나타낸다.

5.2 패킷 크기의 분포

트래픽의 패킷 크기 분포에 대한 분석을 한다. 전술한 것과 같이 업스트림 트래픽 eth.src의 패킷 크기 분포는 다른 데이터들에 비해 그 크기가 작다. 그림 3은 패킷 series의 히스토그램을 나타낸다. eth.dat, eth.dst, 그리고 tcp.dat 의 10 % 이상의 패킷들은 1500 바이트의 페이로드를 갖고 있었다. 반면에 eth.src에서는 3%만이 1500 바이트의 크기를 갖는 패킷들이었다. 그림 3(a), 3(b), 3(c) 그리고 3(d)는 eth.dat, eth.src, eth.dst, 그리고 tcp.dat의 무선 네트워크의 트래픽의 패킷 크기 분포를 나타내고 있다. 무선 트래픽 데이터에 수집된 패킷 크기들은 60에서 1514 바이트 사이로 분포가 되어 있다. 히스토그램에서 알수 있는 것은 eth.dat, eth.src,

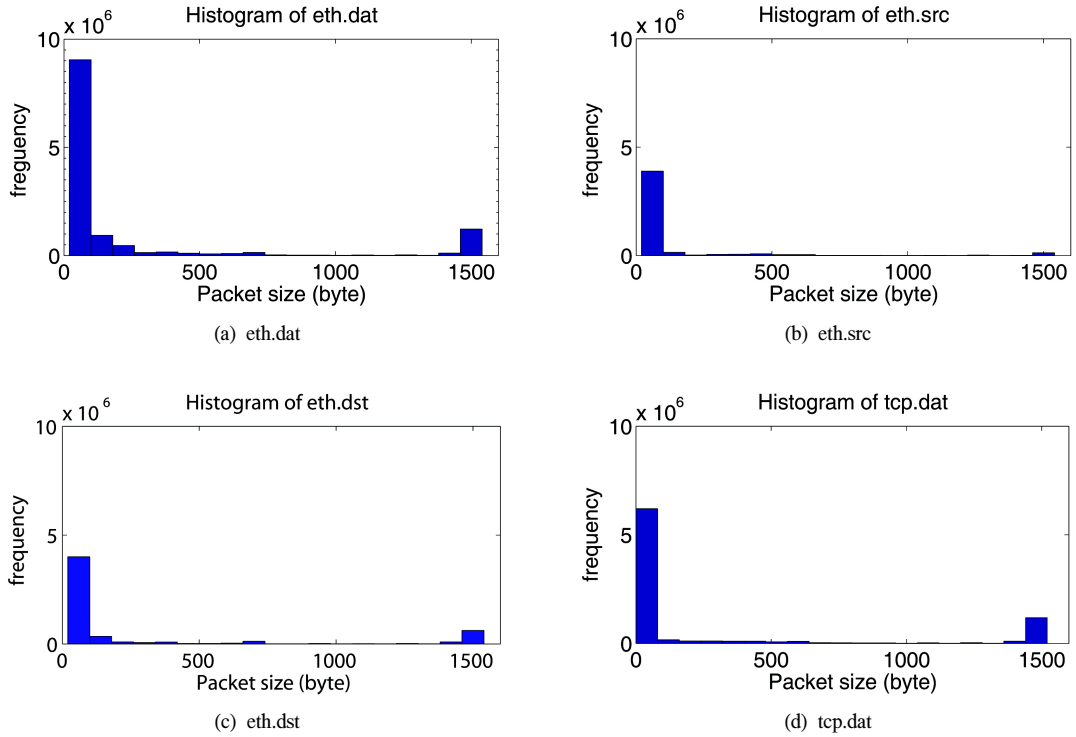


그림 3. Histogram in Data Sets

표 3. Packet Size Statistics of Wireless Network

Rank	eth.dat		eth.src		eth.dst		tcp.dat	
	Size	Average	Size	Average	Size	Average	Size	Average
1	60	24.6 %	60	33.0 %	62	30.0 %	40	66.6 %
2	62	16.6 %	74	25.2 %	60	17.3 %	1500	12.1 %
3	74	13.0 %	62	8.6 %	1514	10.4 %	66	2.0 %
4	1514	9.0 %	66	5.7 %	74	8.2 %	1488	1.4 %
5	66	3.0 %	80	2.7 %	106	3.6 %	5	0.6 %
6	64	28.0 %	1514	2.6 %	78	3.2 %	46	0.6 %
7	214	2.3 %	102	2.0 %	64	2.6 %	192	0.5 %
8	150	2.0 %	91	1.7 %	80	2.3 %	1420	0. %
9	106	1.7 %	64	1.4 %	709	2.0 %	628	0.4 %
10	94	1.6 %	87	1.2 %	66	1.5 %	54	0.4 %
11	78	1.5 %	82	0.7 %	84	1.1 %	1400	0.3 %
12	80	1.4 %	99	0.7 %	1434	0.9 %	1216	0.3 %

eth.dst 그리고 tcp.dat 의 분포인 그림 4(a), 4(b), 4(c) 와 4(d)를 보면 무거운 꼬리 분포를 갖고 있다는 것이다. 표 3에서는 일반적인 패킷 크기의 비율들을 나타내고 있다. 패킷 크기가 80 바이트 이하인 것들은 주로 제어 메시지들을 포함하고 있는 경우가 많다. 약 85 % 이상의 패킷들이 100 바이트

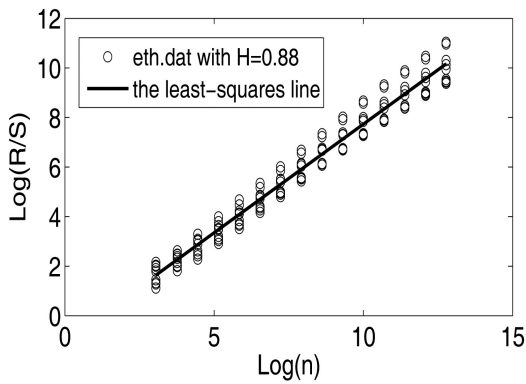
이하의 크기들을 갖고 있는 것을 확인할 수가 있다. 이 현상은 모든 트레이스에서 공통적으로 확인되었다. 패킷의 갯수를 따져보면 최대 크기의 페이로드를 갖는 패킷들이 다운 스트림에서 약 10 %였고 업스트림에서는 2.6 % 였다.

표 4. Description of Model Parameters

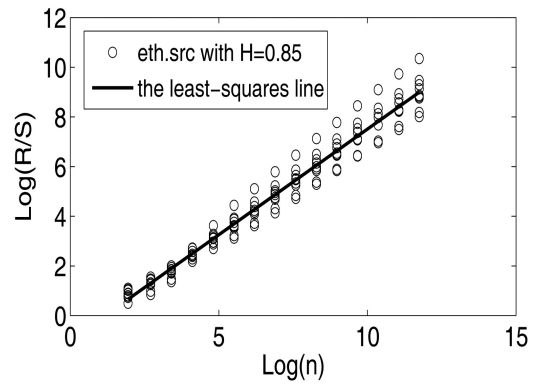
	H Parameter							
	eth.dat		eth.src		eth.dst		tcp.dat	
	10 msec	100 msec	10 msec	100 msec	10 msec	100 msec	10 msec	100 msec
Byte Count	0.82	0.78	.074	0.85	0.81	0.75	0.82	0.79
Packet Count	0.84	0.79	0.80	0.79	0.83	0.80	0.84	0.82

표 5. Description of Model Parameters

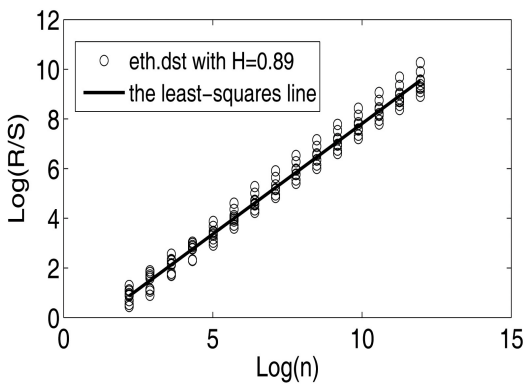
Data set	μ (byte)	σ^2 (byte)	H Parameter		
			Trace Data	FARIMA	FGN
eth.dat	267.4	204760	0.88	0.85	0.82
eth.src	151.7	80103	0.85	0.84	0.83
eth.dst	298.5	239430	0.89	0.84	0.83
tcp.dat	270.7	279430	0.89	0.83	0.83



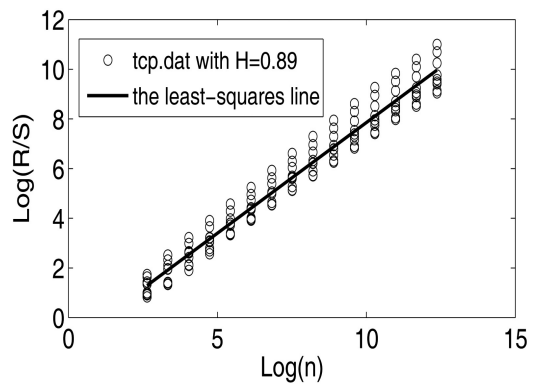
(a) eth.dat: H=0.88



(b) eth.src: H=0.85



(c) eth.dst: H=0.89



(d) tcp.dat: H=0.89

그림 4. Degree of Self-Similar in Datasets: R/S Plots

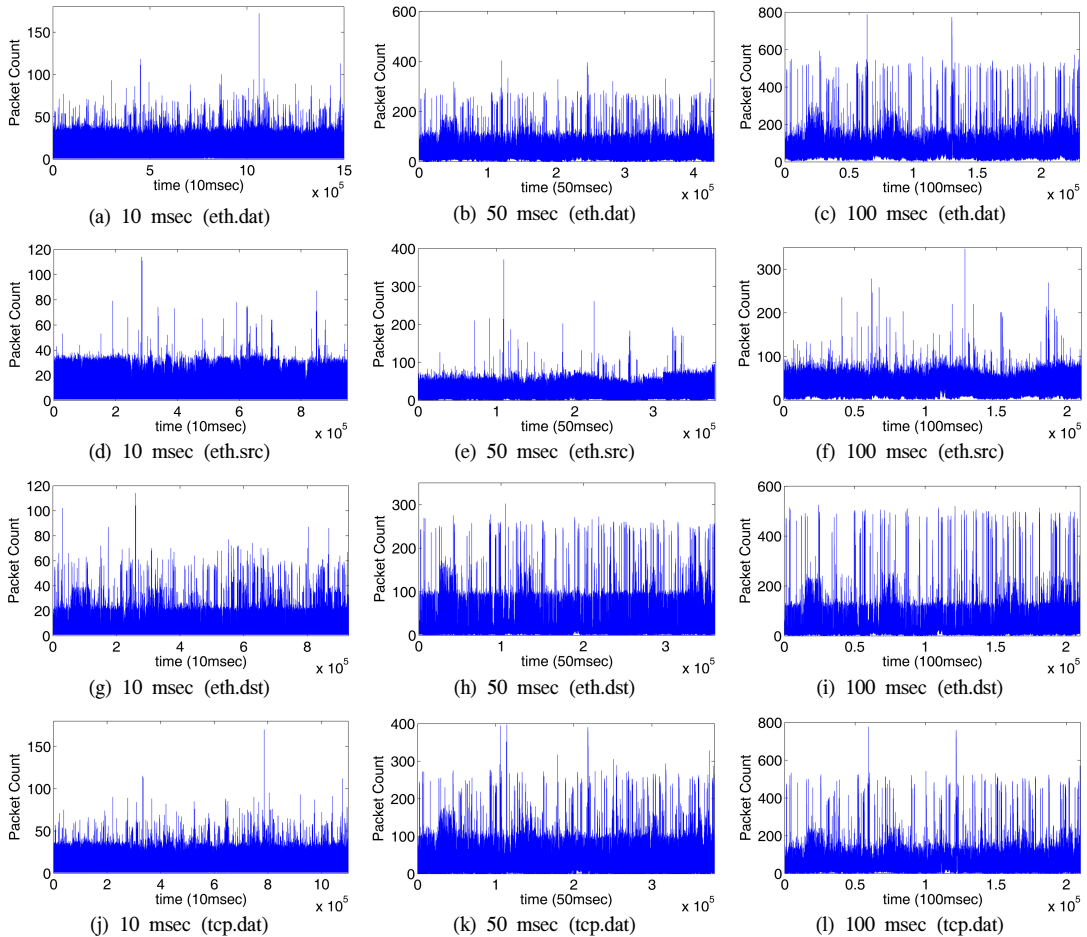


그림 5. Byte Count for eth.dat, eth.src, eth.dst, and tcp.dat

VI. Long Range Dependency

6.1 Empirical Observation

먼저 다른 시간 척도를 갖는 트래픽의 burstiness를 시각적으로 분석한다. 패킷 갯수와 바이트의 크기 프로세서를 10 msec, 50 msec, 그리고 100 msec의 간격을 갖도록 하여 합한 후 도표로 burstiness를 표현한다. 그림 5 와 6에서는 간격을 넓히더라도 burstiness는 눈으로 확인할 수 있을 정도로 완만하게 변하지는 않는다는 것을 확인할 수 있다. 표 4에는 허스트 변수의 값이 나와 있고, 표 5에는 FGN과 FARIMA로 구현된 임의의 데이터의 값들의 통계 값들이 나와 있다.

6.2 Sample Autocorrelation

그림 9에서는 트래픽의 표본 자기 상관계수를 나타내고 있다. ACF의 값은 지연이 커짐에 따라 매

우 느린 속도로 감소하고 하고 있는 것을 확인할 수가 있다. 계층 3의 네트워크 트래픽이 또는 계층 2의 이더넷 트래픽이 자기 유사성을 띄도록 만드는 다양한 요인들이 있을 수 있다. 그 중에는 파일 크기 분포가 무거운 꼬리 분포의 특성을 갖고 있는 경우와 TCP의 혼잡제어 알고리즘²³⁾이 원인일수가 있으며, 사용자가 컴퓨터를 사용하는 on-off 분포라든가, 또는 이 모든 것들의 조합등의 이유로 그렇게 될 수가 있다. 이러한 특성은 자기 유사성을 표현하는 H 값을 예측하는데 있어서 몇 가지 방법들을 제안한다. 전술한 R/S plot을 사용하여 허스트 변수를 예측한다. 총합된 프로세스의 R/S statistics는 매우 큰 n에 대하여 log-log plot은 선형적으로 증가하는 특성이 있다. 이 R/S 표본의 선형 회귀 분포는 허스트 변수 H의 예측 값이 된다. R/S statistics는 그림 4(a), 4(b), 4(c)와 4(d)에 나타나 있다.

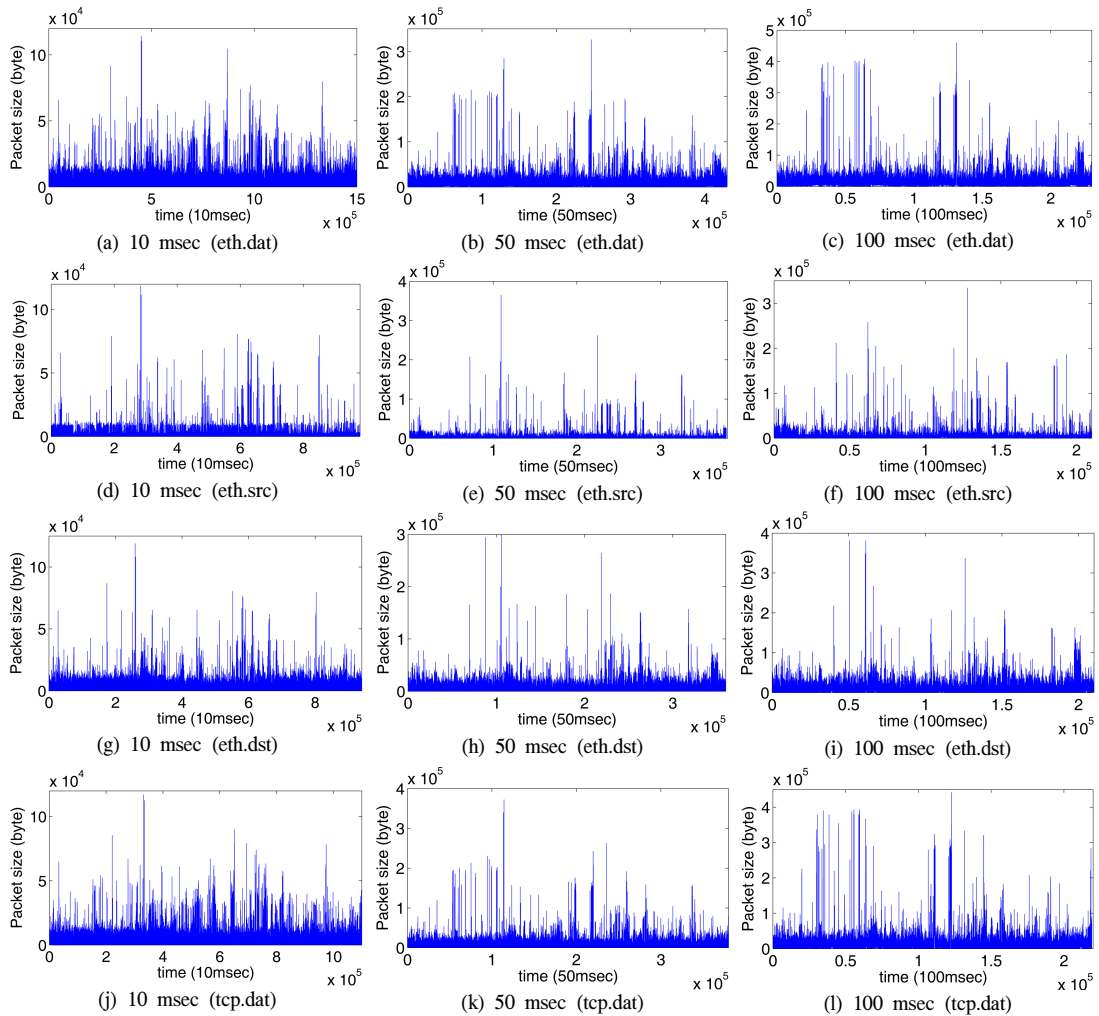


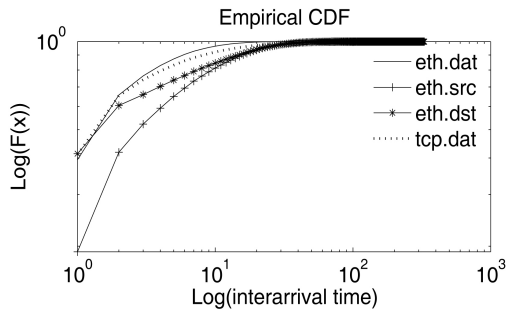
그림 6. Packet Count for eth.dat, eth.src, eth.dst, and tcp.dat

VII. Modeling of Wireless LAN Traffic

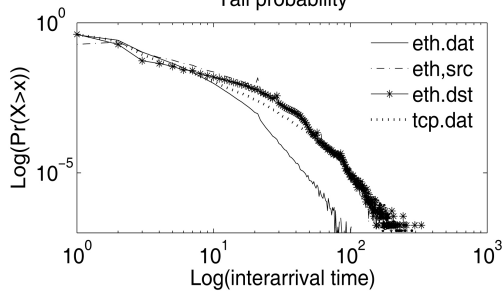
7.1 Background

트래픽의 장기 기억의 특성을 모델링하기 위해서 여러 종류의 분석적 기법들이 소개 되었다. 본 논문에서는 Fractional Gaussian Noise (FGN)과 Fractional ARIMA (FARIMA) 기법을 사용하여 트래픽의 특성을 모델링하려 한다. FGN 알고리즘은 FGN의 파워 스펙트럼과 같도록 표본 경로를 가공하는 것에 근본을 두고 있다. 이렇게 생성된 표본 경로는 자기 유사성 성질을 갖는 네트워크 트래픽의 트레이스를 모의실험하는 것에 사용이 될 수가 있다. 이 알고리즘은 FGN의 파워 스펙트럼을 빠르게 근사화하는 기법이고 네트워크 도착 프로세스에서 장기 의

존성 또는 히스트의 변수의 값이 얼마나 강하게 나타나는지 빠르게 예측해 볼 수 있는 응용 기법이기도 하다¹⁹⁾. 또한, 인공된 FGN을 생성하기 위해서 FFT 알고리즘 하용하였고 이것을 통해 트래픽 데이터를 생성해냈다(자세한 알고리즘은 다음 절에 소개하도록 한다). 최근 실제 트래픽의 측정을 통해 단기 의존성과 장기 의존성이 트래픽 트레이스에 같이 나타나고 있는 것이 발표 되었다¹⁶⁾. 그렇기 때문에 모델링을 했을 때 장기 의존성과 단기 의존성을 동시에 나타낼 수 있는 것이야 한다. FARIMA의 경우 이 두가지 모두를 표현하는데 좀더 나은 성능을 보인다고 평가한다. 본 논문에서는 실제 트래픽 트레이스가 FARIMA 모델과 적합하도록 만드는 순서와 주어진 변수들을 사용하여 FARIMA 프로세스를 생성하는 방법에 대



(a) Cumulative Distribution Function for Inter-Arrival Time Tail probability

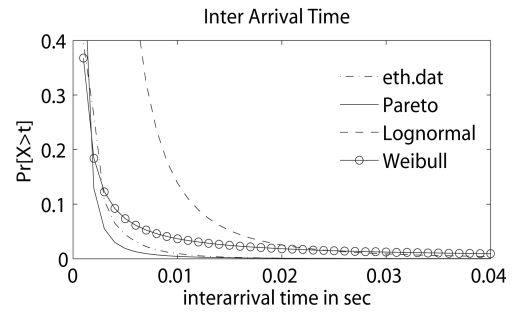


(b) Tail Probability function for Inter-Arrival Time

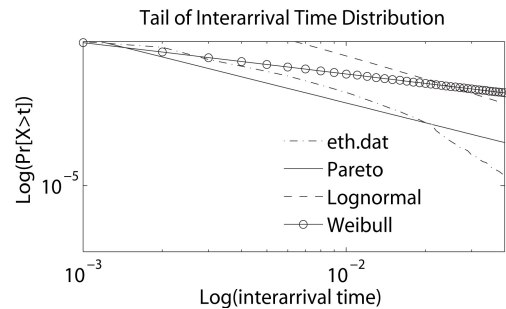
그림 7. Inter-Arrival Time Statistics

해서 설명한다^[8].

먼저 FARIMA를 구하는 Hosking의 알고리즘을 사용하는 Gaussian Motion noise 생성기를 통해서 패킷 프로세스를 재생해 냈다. 무선 랜 트래픽의 자료에서 허스트 변수를 예측했고 이 값으로 무선 랜 트래픽을 가공하였다. 재생된 트래픽에서 FGN 데이터 수열과 FARIMA 수열의 자기상관계수를 계산해 냈을 때 상관도가 매우 높음을 보였고 이것은 그림 9에서 보는 것과 같이 실제 무선 랜 트래픽 데이터와 유사한 것을 보였고 장기 의존성 역시 있는 것으로 보인다. 그림 4(a), 4(b), 4(c) 그리고 4(d)는 R/S plot을 나타낸다. R/S plot의 선형 회귀선 주변으로 보이는 점들이 갖는 군집성은 H=0.88의 장기 의존성을 갖고 있는 것을 나타낸다. 그림 8에서는 Pareto, Weibull, 그리고 Lognormal 분포들을 갖고 도착 간격 시간에 대한 시물레이션한 것을 나타낸다. 그림 9에서는 FGN과 FARIMA의 시물레이션 데이터로 진행한 장기 의존성 실험 결과를 나타내는데 장기 의존성과 자기 유사성 성질을 보이고 있다. 두꺼운 꼬리 분포를 갖는 도착 간격 시간에 대한 시물레이션의 결과를 갖고 장기 의존성 검증 실험을 했을 때 Shilunhin et.al^{[16][17]}이 발표한 내용과 비슷한 H 값을 갖는 결과로 장기 의존성이 있음을 보이고 있다.



(a) Inter-Arrival Time Fitting; eth.dat Vs. Pareto, Weibull, Lognormal Distribution ($\alpha=1.24$)



(b) Tail Probability Function for Inter-Arrival Time

그림 8. Probability of Inter-Arrival Time

7.2 Self-Similar Process from Fractional Gaussian Noise

자기 유사성 프로세스로서 많이 사용되는 두개의 모델은 FGN과 FARIMA 이다^[19]. 이 두 모델 외에 자기 유사성 프로세스를 가공하기 위한 다른 통계적 모델들을 $M/G/\infty$ ^[20], Random Midpoint Displacement^[16], Fractional Brownian Motion^[23], 그리고 Wavelet model^[26]이 있다. 본 연구에서는 FGN과 FARIMA를 이용하여 실제 측정된 트래픽 데이터를 모델링하고, 두 모델이 자기 유사성 프로세스를 얼마나 효과적으로 설명하는지 분석한다. FGN은 Fractional Brown Motion (FBM)의 증분 프로세스이다. 증분을 δ 로 표기하고 $B_H(t)$ 를 Fractional Brownian Motion이라고 할 때 FGN은 $X_H(t) = 1/\delta [B_H(t+\delta) - B_H(t)]$ 으로 표현된다. $X_H(t)$ 은 $N(0, \sigma^H)$ 의 정규 분포를 따른다. 정규 공분산 함수는 다음과 같이 나타낼 수 있다.

$r(\tau) = (|\tau + 1|^{2H} - 2|\tau|^{2H} + |\tau - 1|^{2H})/2$ 총합된 프로세스 $X_H^m(t)$ 는 H가 0과 1 사이의 어떤 값을 대입하더라도 똑같은 분포를 갖고 있다. FGN은 분산과 허스트 값으로 완전히 표현이 된다. 네트워크

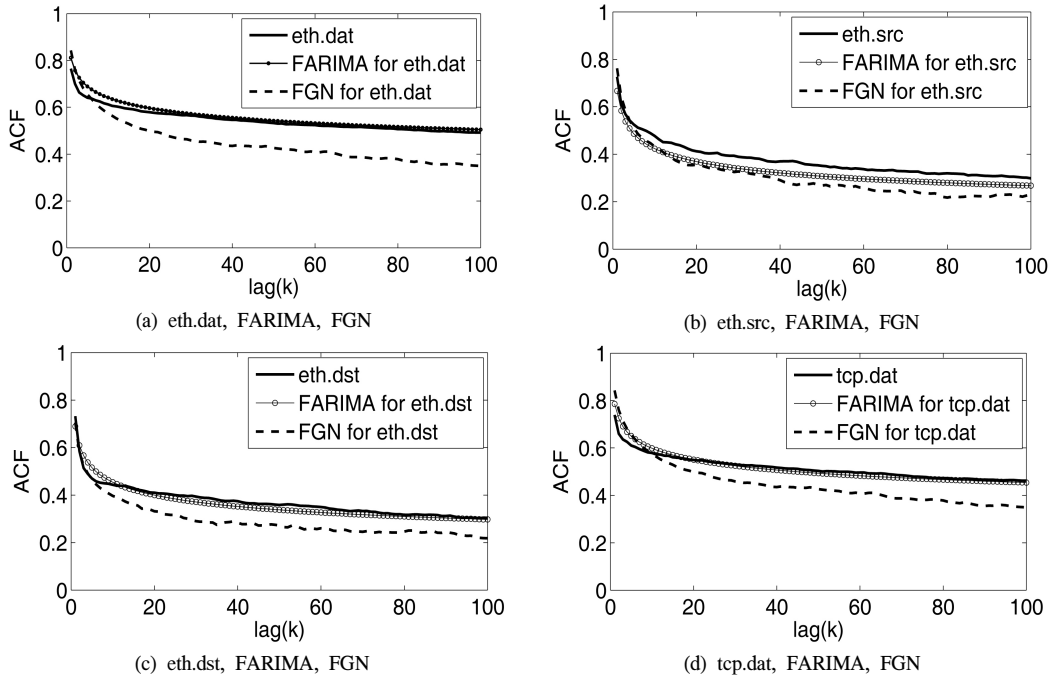


그림 9. Autocorrelations of Sample Data Sets and ACF of Synthetic Traffic Generated by FGN and FARIMA

크 분야에서 FGN 모델을 적용할 수 있는 이유는 트래픽을 ON 인 동안에는 두꺼운 꼬리 분포를 갖는 독립적인 ON/OFF 소스들의 중첩으로 이해할 수 있기 때문이다^[24]. 이러한 경우 평균 도착 시간을 제하고 중심 극한 정리에 따라 필요한 정규화 과정을 거치면 총합된 ON/OFF 소스들은 Gaussian FBM으로 수렴을 하게 된다. 그러므로 자기 유사성을 갖는 중분 프로세스 형태의 트래픽의 경우 분산과 H값이 주어진다면 FGN+mean으로 계산이 될 수 있다. 하지만, FGN 모델의 경우 네트워크에 바로 적용되기에는 몇가지 제한점이 있다. 첫째로 H 값만으로는 실제 네트워크의 복잡한 상관 구조를 설명하기에 충분하지 않다. 더 나아가 단기 상관계수가 버퍼링을 할 때의 영향과 의미있는 시간 척도를 찾는 것의 중요성을 강조 하지 않을 수 없다^[14]. 두번째로 FGN 모델의 Gaussian 특성은 현실과 정확히 대응되지 않는다. 예를 들어 평균제공편차가 평균보다 크다고 한다면 FGN의 결과는 음수의 값들을 갖게 된다. 세번째로 많은 실제 네트워크 어플리케이션들은 Gaussian을 따르지 않고 시간 척도가 작을 수록 더 그러하다. 트래픽이 갖는 복잡한 상관 구조 때문에 FARIMA가 FGN에 비해서 연산량이 많음에도 불구하고 더 선호된다.

Paxson은 빠른 푸리에 변환 기법을 사용하여 자

기 유사성 표본 경로를 생성하는 기법을 만들었다^[19]. 이 알고리즘은 멱 스펙트럼 밀도를 periodogram 기법을 사용하여 계산을 한다. Periodogram은 어떤 주파수 값에 대해 멱 스펙트럼이 독립적인 지 수 확률변수를 통해 표현된다. 이것을 구하기 위해서는 먼저 복소수의 크기는 정규 분포를 따라야 하고 빠른 푸리에 역변환을 거쳐야 한다. 그림 10은 자기 유사성 수열이 어떤 순서로 FFT를 사용하여 얻어 졌나를 나타내고 있다.

FGN 프로세스는 다음과 같이 정의 되었다.

$$X(t) = \mu + \sigma * Z_H(t) \text{ 이때 } Z_H(t) \text{는 히스트 변수 } H$$

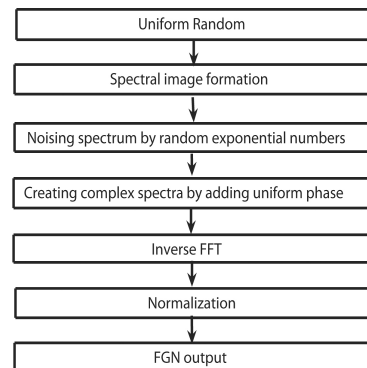


그림 10. Fast Fourier Transform Algorithm for FGN

를 갖는 FGN 프로세스를 나타낸다. $Z_{\mu}(t)$ 가 정규 분포이기 때문에 μ 와 σ 는 $X(t)$ 의 평균과 편차를 나타낸다. 알고리즘은 그림 10에 나와 있다. 이 모델의 결과는 μ 와 σ 위에 나타난 실제 무선 랜에서 얻은 4개의 트래픽 데이터의 총합된 트래픽에 적용하여 보였다. 주어진 4개의 데이터를 갖고 시행한 LRD 실험에서 얻은 허스트 변수의 예측 값은 표 2에 나와 있다.

7.3 Self-Similar Process from Fractional-ARIMA

프랙탈 프로세스들을 사용하는 다른 모델링 기법들보다 좋은 성능을 갖고 있다고 알려진 FARIMA (fractional autoregressive integrated moving average) 기법을 이용하여 분석을 한다. FBM은 소스의 ON/OFF 상태 변화가 자주 일어나는 경우를 모델링 할 때 주로 사용된다. FBM의 경우 하나의 변수만을 갖고 있고 이 하나의 변수가 상관 함수를 제어한다. 그렇기 때문에 FBM으로는 단기 의존적인 것을 모델링하는데는 한계가 있다. ON과 OFF 기간이 무한의 분산을 갖는 많은 수의 ON/OFF 소스들의 합은 장기 의존성 성질의 띠도록 만들고 장기 의존적인 트래픽의 특성을 접근적 표현이 가능하게 한다. 하지만 단기 현상을 시뮬레이션 하고자 한다면 그것은 가능하지가 않다.

FARIMA(p,d,q) 모델의 경우 3개의 변수 p, d, 그리고 q가 존재 한다. 이 변수들은 상관 구조를 제어한다. 그러므로 FBM 과 달리 장기 의존성 뿐만 아니라 단기 의존성도 표현할 수가 있다. 단기 의존성과 장기 의존성 그리고 다른 종류의 분포까지도 표현할 수 있는 모델을 생성해내는 것은 매우 필요한 일이다. FARIMA(p,d,q) 프로세스에서 d는 정수가 아닌 값을 갖고 있으며 차이의 정도를 표현하고, p는 autoregression order를 표현한다. 그리고 q 는 이동 평균의 값을 나타낸다. 패킷 트래이스의 특성을 가공하기 위해서 fractional ARIMA 프로세스를 사용한다. FARIMA(0,d,0) 을 사용하여 임의의 수열을 만들었다. 이때 fractional differencing order d 값은 $0 < d < 1/2$ 이고 이 값은 $H = d + 1/2$ 를 만족한다. Hosking의 알고리즘^{[8][9]}을 사용하여 FARMIA 모델을 이용한 자기 유사성 프로세스를 생성하였다. 실제로는 장기 수열을 생성해 내는 것에는 많은 연산 작업이 필요로 한다.

VIII. 결 론

본 연구에서는 802.11 네트워크 트래픽에 대한 포괄

적인 분석을 진행했다. 분석을 위해서 실제하는 무선 랜 환경에서 수집한 패킷 트래이스를 사용하였고 4개의 데이터 집합들인 총합 된 트래픽, 업스트림 트래픽, 다운 스트림 트래픽 그리고 TCP만의 총합 된 트래픽을 분석하였다. 먼저 기초 통계 분석을 하였고 모든 데이터에 대해 표본 자기 상관계수를 구했다. 이 분석을 통해서 업스트림 트래픽의 패킷 크기 분포는 다른 세개의 데이터들과는 다른 분포를 갖고 있음을 확인했다. 데이터 패킷의 일부분으로 이해 할 수 있는 최대 전송 가능한 크기의 페이로드를 갖는 패킷이 업스트림 트래픽에서의 분포가 적음을 알았다. 패킷의 도착 시간 간격은 Pareto 분포와 유사했다. 또한 트래픽 트래이스의 패킷 갯수 분포 (packets/sec)와 바이트크기 프로세스 (bytes/sec)에 대한 분석도 하였다. 이 분석들을 통해서 트래픽 트래이스의 자기 상관 계수는 매우 느리게 감소하는 것을 알 수가 있었고 허스트 변수가 0.5보다 큰 경우에는 트래픽 데이터가 장기 의존성 성질을 포함하고 있다는 것을 알았다. FGN과 FARIMA 를 이용하여 장기 의존성 성질을 모델링하였고 FARIMA를 이용했을 때 트래픽의 장기 기억 특성을 더욱 정확하게 표현할 수 있음을 확인하였다.

IX. Appendix: Proof of Long Range Dependence

장기 의존성의 증명은 다음과 같다. 정수배의 수로 이루어진 표본의 증분 프로세스 $X(t)$ 로 이루어진 새로운 프로세스 $Y(t) = X(t) - X(t-1)$ 를 정의한다. 증분 프로세스의 장기 의존성 특성은 프로세스 $Y(t)$ 의 공분산을 분석함으로써 얻을 수 있다.

$$\begin{aligned} E[Y(t+k)Y(t)] &= E[(X(t+k) - X(t+k-1)) \\ &\quad (X(t) - X(t-1))] \\ &= E[X(t+k)X(t)] \\ &\quad - E[X(t+k)X(t-1)] \\ &= E[X(t+k-1)X(t)] \\ &\quad + E[X(t+k-1)X(t-1)] \end{aligned}$$

위의 식은 다음의 식으로 간단화 될 수 있다.

$$\begin{aligned} \gamma_Y(k) &= E[Y(t+k)Y(t)] \\ &= \frac{\rho^2}{2} (k-1)^{2H} - 2k^{2H} + (k+1)^{2H} \end{aligned}$$

증분 프로세스 $\rho(k)$ 의 자기 상관계수는 다음의 식으로 나타낸다.

$$\rho_Y(k) = \frac{\sum_Y Y(k)}{\sigma^2}$$

$$= \frac{1}{2} \{ (k-1)^{2H-2k^{2H}} + (k+1)^{2H} \}$$

$\rho(k)$ 의 점근적 성질은 테일러 급수를 전개함으로써 얻을 수 있고 이 식은 다음과 같이 변형될 수 있다.

$$\rho(k) = \frac{k^{2H}}{2} \left[\left(1 + \frac{1}{k}\right)^{2H} - 2 + \left(1 - \frac{1}{k}\right)^{2H} \right]$$

$$= \frac{k^{2H}}{2} g(k^{-1})$$

이 때 $g(x) = [(1+x)^{2H} - 2 + (1-x)^{2H}] \circledast$ 이다.
 $g(x)$ 의 일차 미분과 이차 미분을 구하면 다음과 같다.

$$g'(x) = 2H[(1+x)^{2H-1} - (1-x)^{2H-1}]$$

$$g''(x) = 2H(2H-1)[(1+x)^{2H-2} + (1-x)^{2H-2}]$$

$g(x)$ 를 테일러 급수 전개 하면 다음과 같이 표현된다.

$$g(x-x_0) = g(x_0) + xg'(x_0) + \frac{x^2}{2}g''(x_0) + \dots$$

$$g(x) = g(0) + xg'(0) + \frac{x^2}{2}g''(0) + \dots$$

$g'(x)$ 와 $g''(x)$ 에 대한 식을 치환하면 $g(x)$ 를 전개 했을 때 첫번째 나타나는 영이 아닌 항은 다음과 같다.

$$g(x) = \frac{x^2}{2} 2H(2H-1)$$

그러므로 $\rho(k)$ 에 대한 표현은 다음과 같이 된다.

$$\rho(k) = \frac{k^{2H}}{s} g(k^{-1}) = k^{2H} H(2H-1) k^{-2}$$

$$\lim_{k \rightarrow \infty} \frac{\rho(k)}{H(2H-1)k^{2H-2}} \rightarrow 1$$

참 고 문 헌

[1] A. Balachandran, G. Voelker, P. Bahl, and P. Rangan. "Characterizing user behavior and network performance in a public wireless lan," In Proc. *ACM SIGMETRICS'02*, June 2002. NY, USA.

[2] Jan Beran. "Statistics for long-memory processes," *Jeffrey Glosup Technometrics*, 39(1):105-106, 2007.

[3] J.J. Collins and C.J De Luca. "Upright, correlated random walks: a statistical-biomechanics approach to the human postural control system,"

Chaos, 5(1):57-63, 1994.

[4] Cox D. "Computer Experiments with Fractional Gaussian Noises," *A Review*, In Statistics and Appraisal Iowa State Statistical Library, 1984.

[5] R. F. Felter. "Processes stochastiques fractals avec applications These de doctorat." PhD thesis, *Univ.Paris*,1998.

[6] U. Frisch. "Turbulence: the legacy of a.n. kolmogorov," *Cambridge University Press*, 1995.

[7] Matthias Grossglauser and Jean- Chrysostome Bolot. "On the relevance of long-range dependence in network traffic," *IEEE/ACM Transactions on Networking*, 7(5):629-640, 1999.

[8] Hosking J. "Fractional differencing," *Biometrika*, 68(1):165-176, 1981.

[9] Hosking J. "Modeling persistence in hydrological time series using fractional differencing," *Water Resoures Research*, 20(12):1898-1908, 1984.

[10] M.S. Taquq J. Beran, R. Sherman and W. Willinger. "Long-range dependence in variable-bit-rate video traffic," *IEEE TRans. on Communications*, 5(43):1566-1579, 1995.

[11] W. Leland M. Taquq, W. Willinger and D. Wilson. "On the self-similar nature of ethernet trafficc (extended version)," *IEEE/ACM Transactions on Networking*, Vol.2, pp.1-15. 1994.

[12] Bruce A. Mah. "An empirical model of http network traffic," In *Proceedings of the INFOCOM '97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution*, page 592, Washington, DC, USA, 1997.

[13] M. McNett and G. Voelker. "Access and mobility of wireless pda users," *Technical Report CS'04*, 2004.

[14] W. Willinger M.S. Taquq, V. Teverovsky. "Estimators for long-range dependence: an empirical study, fractals," In *Fractal Geometry and Analysis, World Scientific*, Vol.3, pp.785-798, 1996. Singapore.

[15] Tudjarov A. Temkov D. Janevski T. Firfov. O. "Empirical modeling of internet traffic at middlelevel burstiness," In *Proceedings of the 12th IEEE Mediterranean*, Vol.2, pp.535 - 538, 2004.

[16] O.Shiluhin, S.Smolskiy, and A.Ocin. "Fractal

process in the telecommunication,” *Radio tehnika*, 2003 in Russian.

[17] O.Shiluhin, S.Smolskiy, and A.Ocin. “Modeling for information system,” *Radio Technika*, Moscow, 2005 in Russian.

[18] Kihong Park, Gitae Kim, and Mark Crovella. “On the relationship between file sizes, transport protocols, and self-similar network traffic,” Technical Report 1996-016, 30, 1996.

[19] Vern Paxson. “Fast, approximate synthesis of fractional gaussian noise for generating self-similar network traffic,” *SIGCOMM Comput. Commun. Rev.*, 27(5):5-18, 1997. NY, USA.

[20] J M Pitts and J A Schormans. “Introduction IP and ATM Design and Performance With Applications Analysis Software”. *British Library Cataloguing in Publication Data*, 2000.

[21] Vinay J. Ribeiro, Rudolf H. Riedi, and Richard G. Baraniuk. “Multiscale queuing analysis of longrange-dependent network traffic,” *IEEE/ACM Trans. Netw.*, 14(5):1005-1018, 2000. Piscataway, NJ, USA.

[22] Jong-Suk R.Lee and Hyoung-Woo Park (Korea) Hae-Duck J.Jeong (New Zealand), editors. “Simulation Run-lengths of Self-similar Queuing Processes,” In Proceeding of IEEE, Vol.1. 2003.

[23] William Stallings. “High-Speed Networks and Internets: Performance and Quality of Service,” *Prentice Hall Upper Saddle River*, New Jersey, 2-th edition, 2002.

[24] Shahram Teymori and Weihua Zhuang. “Queue analysis and multiplexing of heavy-tailed traffic in wireless packet data networks,” *Mob. Netw. Appl.*, 12(1):31-41, 2007. Hingham, MA, USA.

[25] Boris Tsybakov and Nicoals D. Georganas. “On self-similar traffic in atm queues: definitions, overflow probability bound, and cell delay distribution.” *IEEE/ACM Trans. Netw.*, 5(3):397-409, 1997. Piscataway, NJ, USA.

[26] Vikram M.Gadre Urali Krishna P and Uday B. Desai. “Multifractal based network traffic modeling.” *kluwer academic publishers*, Boston/Dordrecht/London, 2003.

[27] P. Abry and D. Veitch. “Wavelet analysis of

long-range-dependent traffic.” *IEEE Transactions on Information Theory*, Vol. 44. No.1, pp.2-15, 1998.

[28] M. Gospodinov, E. Gospodinova “The graphical methods for estimating hurst parameter of self-similar network traffic,” Proceedings of the 2005 International Conference on Computer Systems and Technologies, 2005.

대쉬도즈 압힌 (Dashdorj Yamkhin)

정회원



1991년 Novosibirsk State Technical University of Novosibirsk, Russia. Computer System Engineering 졸업

1996년 Mongolian Univ. of Science and Technology, Mongolia 공학석사

1991년~1997년 Mongolian Univ. of Science and Technology, Computer Science and Management School 전임 강사

1997년~2003년 Mongolian University of Science and Technology, Network Administrator

2004년~현재 한양대학교 전자컴퓨터통신공학과 박사과정

<관심분야> Network Traffic, Network Modeling and Analysis, Programming Language

이 성 진 (Seongjin Lee)

정회원



2006년 2월 한양대학교 전자전 기컴퓨터공학과 졸업

2008년 2월 한양대학교 전자컴퓨터통신공학 석사

2008년 3월~현재 한양대학교 전자컴퓨터통신공학과 박사과정

<관심분야> Network Traffic Modeling and Analysis,

Traffic Engineering, Performace Measurement

원 유 집 (Youjip Won)

정회원



1990년 2월 서울대학교 계산통계학과 졸업

1992년 2월 서울대학교 전산학과 석사

1997년 7월 Univ. of Minesota 졸업(박사)

1997년9월~1999년2월 Intel 연구원

1999년 3월~현재 한양대학교 전자컴퓨터통신 공학부 부교수

<관심분야> 운영체제, 컴퓨터 네트워크, 성능평가