

트래픽 패턴분석 기반 실시간 모의 트래픽 생성 및 검증 기법 연구*

강 현 중** · 김 현 철***

A Study on the Real-time Simulated Traffic Generation and Verification Methods based on the Traffic Pattern Analysis

Kang, Hyun Joong · Kim, Hyun Cheol

〈Abstract〉

High-speed services and bulk transmission were available by the development of the network and communication technologies. Moreover various next-generation converged services are undergoing a change by various services. This paper presents improved real-time simulated traffic generation and verification schemes based on the actual traffic pattern analysis. For this, we analyzed traffic patterns of actual application system and generated simulated traffics. We also suggested scheme that verify similarity of simulated traffic and actual traffic.

Key Words : Simulated Traffic Generation, Patter Analysis, Traffic Verification

I. 서론

정보화시대의 발전과 더불어 사용자들은 보다 신속하고 다양한 통신서비스를 요구하고 있으며, 이를 위해 다양한 통신기술을 기반으로 여러 형태의 초고속 전송 기술 및 통신 서비스가 속속 등장하고 있다. 이처럼 새로운 융합형 서비스의 등장과 더불어 통신망에서 제공하는 서비스 및 속도가 급속하게 발전함에 따라 사용자들의 서비스 요구는 간단한 텍스트정보의 전송에서부터 데이터, 이미지, 음성과 동영상 정보에 이르기까지 매우 다양하

게 변화하고 있고, 요구하는 대역폭도 급속도로 증가하고 있는 실정이다.

한편 통신서비스 사업자들은 이러한 사용자 요구사항의 변화를 만족시키기 위해 통신 서비스 수요 예측을 실시하고 있으나, 새로운 형태의 융합 서비스의 대중화 및 초고속화에 따른 통신서비스 패러다임의 변화를 따라가지 못하고 있다. 따라서 현재의 통신 서비스와 향후 등장할 다양한 형태의 융합 서비스를 유기적으로 제공하기 위해서는 통신서비스 패러다임의 변화에 대한 능동적인 대처와 더불어 현재의 통신 서비스와 새로운 융합 서비스에 모두 효과적으로 적용 가능한 모의 트래픽 생성 및 검증 알고리즘에 대한 연구가 절실한 상황이다.

본 논문은 통신망에서 전달되는 실제 트래픽과 동일

* 이 논문은 2008년도 서일대학 학술연구비의 지원에 의해 연구되었음.

** 서일대학 인터넷정보과 교수

*** 남서울대학교 컴퓨터학과 교수

한 패턴을 네트워크 시뮬레이션 환경에서 생성하는 알고리즘의 제안을 목적으로 하고 있다. 이를 위해 본 논문에서는 실제 응용체계의 트래픽 발생 패턴을 분석하여 이를 기반으로 모의 트래픽을 생성하고, 생성된 모의 트래픽과 실제 트래픽의 유사성을 검증하는 기법을 제안하였다.

또한 본 논문에서는 빠르게 변화하는 초고속 네트워크 환경을 시뮬레이션 환경에 반영하기 위하여 앞으로 등장할 신규 서비스의 트래픽 패턴 예측방법을 제안하였다.

II. 관련연구

모의 트래픽을 생성하기 위해서는 기존의 모의 트래픽 생성 기법과 검증 기법에 대한 연구가 반드시 선행되어야 한다. 본 절에서는 기존의 모의 트래픽 생성을 위한 대표적인 모델링 방법과 모의 트래픽 검증 방법을 조사하고 이를 분석하였다.

2.1 모의 트래픽 생성기법

트래픽은 다양한 특성을 기반으로 여러 가지 형태로 모델링이 될 수 있다. 대표적으로 CBR(Continuous Bit Rate)과 VBR(Variable Bit Rate)로 구분할 수 있으며 VBR에는 POISSON 모델, ON-OFF 모델, IPP(Interrupted Poisson Process) 모델, AR(Autoregressive) 모델, 그리고 MMPP(Markov Modulated Poisson Process) 모델 등이 있다[14].

우선 CBR은 가장 간단한 형태로서 일정 간격으로 트래픽이 발생하는 형태를 갖는다. CBR은 큰 파일의 전송이나 기존 STM(Synchronous Transfer Mode) 기반의 트래픽에서 발행하는 형태이다. 이때 셀의 발생간격의 확률 분포 함수 $F(t)$ 는 다음과 같다.

$$F(t) = U(t - T_c) \quad (1)$$

U 는 단위스텝함수를 나타낸다. 단위스텝함수 $u(t)$ 는 $t < 0$ 일 때 0의 값을 갖고, $t \geq 0$ 일 때 1의 값을 갖는다. 즉, CBR은 t 가 T_c 인 경우만 발생하는 것이다. CBR 모델은 가장 쉬운 트래픽 모델이지만 일정 간격으로만 트래픽을 생성하므로 트래픽의 다양한 특성을 반영하기는 어렵다.

포아송 모델은 발생 분포가 포아송인 형태로서 t 시간 동안 들어온 셀의 개수가 n 인 확률은 (식 2)와 같고 이 트래픽의 셀 발생률을 λ 라고 할 때 셀의 발생간격의 분포는 (식 3)과 같다.

$$P(n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad (2)$$

$$f(x) = \lambda e^{-\lambda x} U(x) \quad (3)$$

$$\lambda = \frac{1}{T_p} \quad T_p : \text{발생간격의 평균}$$

즉, 포아송 모델은 t 의 시간 간격동안 n 개의 셀이 포아송 분포로 생성되고 이 때, n 개의 셀 각각의 시간간격은 지수분포를 가진다. 포아송 모델은 대표적인 인터넷 트래픽의 특성 중 하나인 포아송 분포 트래픽을 생성하기에 적합하지만 또 다른 자기유사성의 특성을 지니는 트래픽은 생성할 수 없다는 단점이 있다.

ON-OFF 모델은 버스티니스가 큰 트래픽을 모델링하기 위한 형태로서 지수 분포를 가지는 ON 구간과 OFF 구간을 가지며 ON 구간에서는 일정 간격의 셀이 발생되고 OFF 구간에서는 셀이 발생되지 않는다. ON-OFF 모델에서는 구간을 ON 구간과 OFF 구간으로 나누어 ON 구간 안에서는 CBR 모델과 같이 일정간격으로 셀을 생성한다. 따라서 CBR 모델의 발전된 형태라고 할 수 있지만 ON 구간의 트래픽 분포의 간격을 유연하게 조정할 수 없다.

IPP 모델은 MMPP 모델의 특수한 경우로서 ON-OFF 모델과 거의 유사하며, ON 구간에 셀의 발생이 CBR이 아니라 $\lambda = IPP(1/Tipp)$ 를 가지는 포아송 분포를 가진다. 즉, ON-OFF 모델의 유연성을 가지면서 ON 구간의 트래픽에 포아송 분포를 적용 할 수 있다. 이것은 통계적으로

Hyperexponential(초지수) 분포(H2)를 가진다.

$$f_{H_2}(t) = P\mu_1 e^{-\mu_1 t} + (1-P)\mu_2 e^{-\mu_2 t} \quad (4)$$

한편 영상신호는 영상압축기법(interframe differential pulse code modulation(DPCM), intraframe DPCM, conditional replenishment, motion compensation, transform coding(예, Discrete Cosine Transform, Run Length Coding) 등)과 scene type(화상회의, 스포츠, 영화), 카메라 종류, 화면의 명암, 카메라 동작(zooming, panning 등)에 따라서 발생하는 비트율과 그 형태가 다르므로 이것은 정확히 모델링 한다는 것은 매우 어렵다. 이 중 비디오폰과 같이 화면의 변화가 적은 영상신호는 AR 모델로 모델링이 가능하다.

MMPP 모델은 가장 복잡한 형태의 모델링이며 여러 개의 데이터 소스나 ON-OFF 음성 신호의 다중화 된 트래픽을 모델링하거나 장면 변화가 많은 영상신호를 모델링하는데 사용되기도 한다. 2-state MMPP는 포아송 모델과 IPP 모델의 합으로 나타낼 수가 있다. 또한 2-state MMPP의 다중화 된 것은 2n -state MMPP 모델이 될 수 있으므로 더 복잡한 트래픽의 모델링도 가능하다.

2.2 모의 트래픽 검증기법

기존 연구에서는 네트워크 설계시 측정에 기반을 두지 않고 포아송 분포와 지수함수 분포같이 수학에 기반을 두고 연구를 수행하였다. 그러나 실제 측정된 트래픽은 수학적인 모델과 다른 특성을 가진다는 것을 최근 연구들에서 보여주고 있다.

수학적인 모델들은 수학적인 분석이나 컴퓨터를 사용한 수치해석에는 적합하지만, 이를 바탕으로 네트워크를 설계했을 때 자원의 낭비나 부족을 초래할 수 있다. 따라서 네트워크 설계시 실제 측정된 데이터를 사용하여 실제 트래픽 특성을 반영하는 일이 필요하다. 이를 위해 자기유사성 모델이 대두되었으며 연구가 수행되었다[5-6].

자기유사성이란 급수를 다수 다중화 시켰을 때도 새

로운 급수들은 원래 급수와 같은 자기 상관함수를 가지는 특성을 말한다. 즉, 어떤 데이터의 급수가 특정 자기 상관함수를 가질 때 그 데이터를 다중화 시켜 생성된 새로운 급수들도 원래 데이터 급수와 같은 자기 상관함수를 가진다는 것이다. 대표적인 방법으로는 트래픽의 다중화 정도에 따른 트래픽의 분산의 감소도를 살펴보는 Time-Variance Plot 기법과 R/ S(Rescaled Range)가 H를 가진 Power law에 따라 커지는 특성을 사용한 R/ S plot 기법이 있다.

R/S plot은 기울기가 입력부하 따라 커진다는 특성을 이용한다. 식을 살펴보면 n은 자연수로 데이터 포인트의 샘플링 간격을 의미하며, R(n)과 S(n)은 각각 조정된 범위 내에서 프로세스의 통계적 수치와 샘플(Sample)의 표준 편차를 나타낸다. 즉, R(n)은 사용자가 미리 정해 놓은 값이고 S(n)은 실제 네트워크의 트래픽이다. 샘플의 평균값을 $\bar{X}(n)$ 라 할 때, 샘플 정규화 프로세스 W_k 와 R(n)/S(n)에 관한 식은 (식 5), (식 6)과 같이 정의된다.

$$W_k = (X_1 + X_2 + \dots + X_k) - k\bar{X}(n), (k \geq 1) \quad (5)$$

$$R(n)/S(n) = \max(0, W_1, W_2, \dots, W_n) - \min(0, W_1, W_2, W_n)]/S(n) \quad (6)$$

$\log [R(n)/S(n)]$ 와 n을 log - log 그래프로 나타내면, 일반적으로 기울기가 0.5와 1사이에 존재하는 직선이 나타난다. R(n)/S(n)에 대한 수식에서 X가 증가하면 W도 증가한다. 그러므로 R(n)/S(n) 수식의 기울기가 증가하게 된다. 여기서 기울기는 1에 가까워질수록 보다 정확한 자기유사성을 가진다고 할 수 있다. 실제 인터넷 트래픽에서도 입력 부하(X)가 증가함에 따라 자기유사성의 강도가 커지므로, RS plot에 따른 검증은 실제 트래픽의 특성을 잘 반영한다는 것을 확인할 수 있다[7].

Time-Variance plot 은 관찰하고자 하는 데이터의 자기유사성을 그래프로 보여주는 방법이다. m개의 데이터를 하나로 묶어서 그 묶음의 평균값의 확률과정(Random

Process)을 $X^{(m)} = (X_k^{(m)} : k = 1, 2, 3, \dots)$ 라고 하면, 이때 $X_k^{(m)}(X_k^{(m)} = \frac{1}{m} \sum_{i=km-(m-a)}^{km} X_i)$ 의 분산은 (식 7)과 같이 정의된다.

$$\begin{aligned} \text{Var}(X^{(m)}) &= \frac{1}{n} \sum_{k=1}^n (X_k^{(m)} - \overline{X^{(m)}})^2 \\ &= E[X^{(m)^2}] - \overline{X^{(m)^2}} \end{aligned} \quad (7)$$

만약 이 확률 과정이 자기유사성을 가진다면, 위의 식에서 나온 분산들은 (식 8)을 만족한다.

$$\text{Var}(X^{(m)}) \sim \frac{\text{Var}(X)}{m^\beta} \quad (8)$$

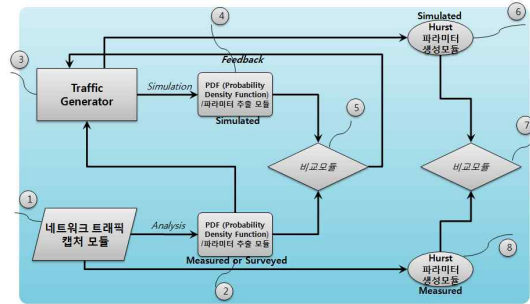
III. 패턴분석 기반 모의 트래픽 생성 및 검증 알고리즘

3.1 모의 트래픽 생성 및 검증 모듈

<그림 1>은 본 논문에서 제안하고 있는 분포함수와 파라미터를 기반으로 한 실시간 모의트래픽 생성기의 전체적인 구조를 나타내고 있다. <그림 1>에서와 같이 본 논문은 크게 네트워크 트래픽 캡처 모듈, PDF (Probability Density Function)/파라미터 추출 모듈, 트래픽 생성 모듈, 비교 모듈, 그리고 Hurst 파라미터 생성 모듈로 구성되어 있다.

먼저 <그림 1> ①은 트래픽 캡처 모듈을 나타내고 있다. 트래픽 캡처 모듈에서는 측정하고자 하는 네트워크의 패킷들을 캡처하여 이를 플로우 정보와 패킷 헤더 정보로 분류한다. 플로우 정보로는 시간 단위 당 얼마나 많은 양의 패킷이 발생했는지, 특정 시간대의 트래픽 양이 어느 정도나 되는지 등에 대한 데이터 값들을 파악할 수 있으며, 패킷 헤더 정보에서는 헤더 내에 포함되어 있는 필드 값들을 분석해 패킷의 출발지와 목적지, 패킷의 길이, 발생 시간, 해당 응용체계 등 패킷 자체에 대한 정보

를 파악할 수 있다. 이렇게 수집된 정보들은 모의 트래픽과의 비교를 위해 각 응용체계별로 분류하여 필요한 파라미터들을 따로 분석, 관리하게 된다.



<그림 1> 패턴분석기반 모의트래픽 생성 및 검증 구성도

<그림 1> ②는 PDF/파라미터 추출 모듈을 나타내고 있고 PDF/파라미터 추출 모듈은 네트워크 트래픽 캡처 모듈에서 얻어진 정보를 통해 해당 트래픽의 확률밀도함수(PDF)와 그에 따른 파라미터 값을 추출한다. 만약, 트래픽 플로우만을 비교하려 한다면 패킷 캡처를 통해 알게 된 트래픽 량을 Q-Q Plot을 이용하여 어떠한 분포를 보이는지 파악하면 되고, 응용체계 별 세분화된 비교 군을 위해서는 패킷 헤더 내에 포함된 정보를 사용하여 별도로 트래픽 정보를 작성한 후, Q-Q Plot을 이용하여 분포를 찾아내면 된다. Q-Q Plot은 동일 비교군 내에 포함되어 있는 두 분포가 서로 일치하는지 여부를 측정하는 방법으로서, 불특정 분포가 어떠한 특정 분포와 가장 유사하고 적합한지 측정하기 위한 방법으로도 많이 쓰이고 있다. 이를 통해 최적합 확률 밀도 함수를 찾아내고, 그 분포함수를 구성하고 있는 파라미터(모수) 값들을 추출하여 분포함수-파라미터 쌍에 대한 정보를 저장한다.

<그림 1> ③은 트래픽 생성 모듈을 나타내고 있고 모의 트래픽을 생성하기 위해서 기존 응용체계별로 분류하여 트래픽들을 모델링 한다. 이렇게 기존 응용체계별로 분류하여 트래픽 모델링을 하는 과정에서는 각각의 응용 체계에 맞는 트래픽 분포 값을 입력해 줘야 한다. 시뮬레

이터를 통해 트래픽을 모델링하게 되면 특정 분포와 그에 따른 파라미터를 입력할 수 있다. 이 때, 위에서 측정된 확률 밀도 함수와 파라미터 값들을 각 응용체계에 맞게 입력해주면 된다. 이렇게 모델링된 트래픽 생성기를 통해 모의 트래픽을 발생시킨다.

<그림 1> ⑤는 실제 트래픽과 모의 트래픽간 파라미터를 비교하는 비교 모듈을 나타내고 있다. 트래픽 생성기를 통해 모의 트래픽을 발생시키게 되면 이 트래픽을 통해 마찬가지로 확률 밀도 함수와 파라미터 값을 추출해낼 수 있다. 만약, 모델링이 성공적이며 정확하게 되어 있는 상태라면 모의 트래픽을 통해 추출해낸 확률 밀도 함수와 파라미터가 실제 트래픽에서 추출하여 모델링 시 입력했던 그것들과 일치하여야 한다. 만약 두 값들이 상이한 결과를 보인다면 트래픽 생성기 내의 모듈 설계상의 오류를 살펴거나 디버깅 과정을 통해 트래픽 생성기의 수정과정을 거친다. 이러한 수정과정을 통해서도 실측 파라미터와 모의 트래픽 파라미터의 값에 오차가 생기면 다시 검증모듈의 첫 번째 과정으로 돌아가 실제 네트워크 트래픽에 대한 분석을 다시 수행할 필요가 있다. 최초로 측정된 확률 밀도 함수와 파라미터 값이 잘못되었다면 그 값의 입력을 통해 생성된 모의 트래픽이 실제 트래픽과 유사할 꺼라 기대할 수 없다.

이러한 과정들을 통해 비교된 플로우와 파라미터 값들이 큰 오차 없이 일치하게 된다면 이는 트래픽 생성기의 설계와 실제 트래픽으로부터 추출해온 파라미터들에 대한 분석이 이상 없이 이루어졌음을 증명할 수 있게 된다.

<그림 1> ⑥, ⑦ 그리고 ⑧은 Hurst 파라미터 생성 및 비교 모듈을 나타내고 있다. Hurst 파라미터 생성 및 비교 모듈에서는 트래픽의 자기유사성을 통해 실제 트래픽과 모의 트래픽의 자기유사성 정도를 비교한다. 자기유사성의 정도는 R/S Plot, VT Plot, 또는 Wavelet 방식을 통해 Hurst 파라미터 값을 구함으로써 알 수 있다. 실제 트래픽에서 측정된 Hurst 파라미터 값이 0.5 이상, 1 이하의 값을 보이면 이는 자기유사성을 보인다고 할 수 있으며, 그 수치를 통해 어느 정도의 자기유사성을 보이

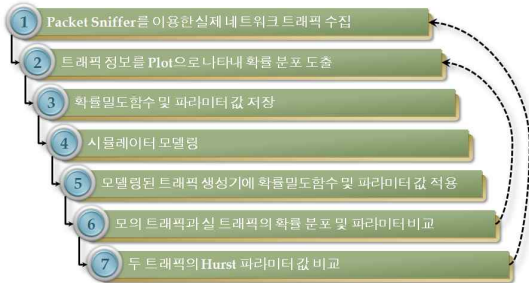
는지 파악한다. 실제 트래픽을 모델링하여 생성된 모의 트래픽은 실제 트래픽과 마찬가지로 자기유사성을 보여야 하므로, 모의 트래픽에도 마찬가지로 Hurst 파라미터 값을 구해낸다. 모의 트래픽의 Hurst 파라미터 값이 범위 내에 존재하지 않아 자기유사성을 보이지 않는다면 실제 트래픽과의 비교가 어려워짐은 물론 두 트래픽이 유사하다고 보기도 힘들다. 두 트래픽이 유사하다는 것을 보이려면 같은 특성을 보여야 하므로 자기유사성에 대해서도 공통적인 특성을 가져야 하며, 두 트래픽 모두 자기유사성을 가질 경우 Hurst 파라미터의 값 역시 큰 오차 없이 일치하는 모습을 보여야 한다.

3.2 유사도 검증 시뮬레이션

제안된 트래픽 생성 모듈의 검증에 앞서 본 논문에서 제시된 트래픽 유사도 검증방안을 위한 시뮬레이션을 수행한다. 시뮬레이션의 순서는 <그림 2>와 같다.

우선 <그림 2>의 예서와 같이 실 트래픽 정보를 수집한다. 이를 위해 Tcpdump나 Wireshark와 같은 Packet Sniffer를 이용해 해당 서버에서 발생하는 트래픽을 수집하여 정보를 분석한다. 모의 트래픽 유사도 측정에 대한 방법론을 시뮬레이션 하는 것이기 때문에 특정 응용 트래픽 하나를 선별하고 그 트래픽의 플로우 정보를 분석하여 엑셀 등의 파일로 데이터를 기록한다.

이후 <그림 2> 예서와 같이 최적 분포 및 파라미터를 도출하며 실측 트래픽의 최적 분포를 찾아내기 위해서 Q-Q Plot을 이용한다. 특정 분포와 실측 트래픽의 분포를 Q-Q Plot으로 나타내면 두 분포간의 일치 여부를 알 수 있는데, 만약 Plot 위의 점들이 일직선 형태를 보이고 있다면 비교한 두 분포가 서로 일치한다는 것을 의미한다. 이러한 방식으로 여러 표준 분포들을 대상으로 실측 트래픽의 최적 분포를 찾아낼 수 있다. Q-Q Plot을 그리기 위해 본 논문에서는 SPSS 등과 같은 통계 프로그램을 사용하여 확률 밀도 함수에 그에 따른 파라미터 값들을 추출한다.



<그림 2> 확률 밀도 함수 및 자기 유사성을 통한 시뮬레이션 과정

이후 트래픽 모델링을 수행한다. 하나의 응용체계를 실측하여 얻어진 확률밀도함수와 파라미터를 트래픽 모델링에 적용한다. 이때 시뮬레이터를 통해 하나의 응용체계에서 발생하는 트래픽을 생성할 모의 트래픽 생성기를 설계하게 되는데, 이 때 앞서 추출한 확률 밀도 함수와 파라미터 값을 입력해주면 된다. 이렇게 특정 분포를 입력하면 시뮬레이터는 해당 분포를 따르는 모의트래픽을 생성하게 된다.

마지막으로 결과 값을 비교한다. 생성된 모의 트래픽도 실측 트래픽과 마찬가지로 확률 밀도 함수와 그에 따른 파라미터를 구해낼 수 있다. 이 경우에도 Q-Q Plot을 사용한다. 이렇게 추출된 모의 트래픽의 결과 값이 트래픽 모델링 시 적용한 확률 밀도 함수와 파라미터 값과 일치하는지 여부를 판단한다. 값이 일치한다면, 두 트래픽은 유사하게 생성되었다고 가정할 수 있다. 이러한 두 트래픽이 동일한 특성을 갖는지 검증하기 위하여 자기 유사성의 정도를 나타내는 Hurst 파라미터 값을 비교한다. 본 논문에서 제시된 Hurst 파라미터 계산법에 의해 실측 트래픽과 모의 트래픽에서의 자기유사성 정도를 계산하여 이를 트래픽 특성에 대한 검증으로 사용한다. 만약, 두 값이 서로 일치하지 않거나 상반되는 특성을 갖는다는 결과를 얻게 되면, 두 트래픽은 유사하다고 볼 수 없으며 실측 트래픽의 확률 밀도 함수를 얻는 과정과 응용체계를 모델링 하는 과정으로 되돌아가 다시 시뮬레이션을 수행해야 한다.

3.3 시뮬레이션 분석

<그림 3>은 Inter-arrival 시간 계산을 위해 수집된 정보들을 나타내고 있다. A열은 각각의 패킷이 도달한 상대적인 시간 값이며, B열은 이 시간들의 차를 이용해 구한 Inter-arrival 시간을 나타내고 있다.

A	B	C	D	E
8062	1790.351056	0.101513		
8063	1790.452569	0.081321		
8064	1790.454589	0.068798		
8065	1790.020438	0.001004		
8066	1790.621532	0.129535		
8067	1790.731937	0.079506		
8068	1790.830643	0.051482		
8069	1790.832125	0.137087		
8070	1790.999812	0.094229		
8071	1791.024041	0.021679		
8072	1791.025272	0.162095		
8073	1791.188937	0.093163		
8074	1791.190133	0.217195		
8075	1791.407138	3.176521		
8076	1794.583849	0.01		
8077	1794.583849	0.204158		
8078	1794.709037	0.271396		
8079	1795.069603	0.001436		
8080	1795.071039	0.164484		
8081	1795.255532	0.060187		
8082	1795.281191	0.101411		
8083	1795.383421	1.101012		
8084	1796.099833	0.083578		
8085	1796.792132	0.453718		
8086	1797.248993	0.094426		
8087	1797.343154	0.104873		
8088	1797.446237	0.095046		
8089	1797.943778	0.150459		
8090	1797.693696	0.099603		
8091	1797.791279			
8092				
8093				

<그림 3> 추출된 트래픽에서의 Inter-arrival 시간 계산

트래픽 모델링에는 OPNET Modeler를 사용한다. 이 때, 트래픽 모델링은 실제 트래픽에서 도출해낸 분포와 파라미터 값의 입력만을 통해 모의 트래픽 생성해주도록 간략하게 이루어진다.

본 시뮬레이션에서는 가장 일반적으로 사용되는 R/S Plot을 사용하여 자기유사성의 정도를 측정한다. R/S Plot에 사용되는 수식은 트래픽의 시간당 패킷 도달률을 기반으로 계산된다.

이러한 과정을 거쳐 R/S Plot의 과정에서 계산되는 몇몇 수식의 결과값들과 함께 최종적으로 도출된 Hurst Parameter 값 0.7304는 0.5와 1 사이에 존재하는 값이므로 본 시뮬레이션을 통해 측정된 트래픽의 시간 당 패킷 도달률은 자기유사성을 가지고 있음을 확인할 수 있다.

IV. 결론

인터넷이 발전하면서 점점 더 많은 시스템들이 네트워크에 연결되고 있으며, 다양한 서비스의 출현 및 사용 인구의 급격한 증가로 인해 네트워크 트래픽은 점점 복잡한 양상을 보이고 있다.

일반적으로 트래픽의 흐름은 응용 서비스별로 서로 다른 특성들을 보이고 있으며 이는 수학적인 분포로도 모델링된다. 대부분의 네트워크 시뮬레이션 도구의 트래픽 생성기에서는 이러한 트래픽 흐름을 모델링하기 위해 지수 분포나 포아송 분포 같은 하나의 수학적 확률 분포만을 사용하여 시뮬레이션용 트래픽을 생성한다. 그러나 실제 네트워크의 다양한 서비스와 이에 따른 트래픽 특성 때문에 이와 같은 특정한 분포로는 실제 트래픽의 특성을 제대로 반영하지 못하는 부분이 있다. 따라서 이를 바탕으로 네트워크를 설계하여 구축하였을 때 자원의 낭비나 부족을 초래할 수 있으므로, 실제 각 응용 서비스별로 생성되는 트래픽의 특성과 흐름을 잘 파악하여 모델링하여야 한다.

본 논문에서는 통신망에서 전달되는 실제 트래픽과 동일한 패턴을 네트워크 시뮬레이션 환경에서 생성하는 알고리즘을 제안하였다. 이를 위해 본 논문에서는 실제 응용체계의 트래픽 발생 패턴을 분석하여 이를 기반으로 모의 트래픽을 생성하고, 생성된 모의 트래픽과 실제 트래픽의 유사성을 검증하는 기법을 제안하였다.

또한 본 논문에서는 빠르게 변화하는 초고속 네트워크 환경을 시뮬레이션 환경에 반영하기 위하여 앞으로 등장할 신규 서비스의 트래픽 패턴 예측방법을 제안하였다.

참고문헌

- [1] A. Veres et al, "On the Propagation of Long-Range Dependence in the Internet," *Computer Communication Review*, Vol. 30, No. 4, Oct. 2000.
- [2] A. Veres, Zs. Kenesi, S. Molnar and G. Vattay, "On the Propagation of Long-Range Dependence in the Internet," *SIGCOMM*. 2000.
- [3] C. You and K. Chandra, "Time Series Models for Internet Data Traffic," *24th Conference on Local Computer Networks*, 1999.
- [4] Mark E. Crovella, Azer Bestavros, "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes," *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, December. 2000, pp. 835-840.
- [5] S. Y. Wang, C. L. Chou, C. H. Huang, C. C. Hwang, Z. M. Yang, C. C. Chiou, and C. C. Lin, "The design and implementation of the NCTUns 1.0 network simulator," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, Volume 42, Issue 2, 2003, pp 175-197.
- [6] W. Willinger, M. Taqqu, R. Sherman, D. Wilson, "Self-Similarity Through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level," *IEEE/ACM Transactions on Networking*, Vol. 5, No. 1, Jan. 14. 1997, pp. 71-86.
- [7] Won-Ki Hong, S. U. Park, Y. M. Kang and J. T. Park, "Enterprise Network Traffic Monitoring, Analysis and Reporting Using Web Technology," *Journal of Network and Systems Management*, Plenum Press, March, 2001, pp. 89-111.

■ 저자소개 ■



강 현 중
Kang, Hyun Joong

1989년 3월~현재
서일대학 인터넷정보과 교수
1982년 3월~1989년 2월
한화종합금융(주) 전산팀장
1979년 11월~1982년 2월
한국과학기술연구소(KIST) 연구원
1996년 2월 성균관대학교 정보공학과(공학박사)
1986년 2월 연세대학교 전자계산학과(공학석사)
1980년 2월 성균관대학교 수학교육학과(이학사)

관심분야 : 데이터통신, 프로그래밍언어
E-mail : college@dreamwiz.com



김 현 철
Kim, Hyun Cheol

2006년 9월~현재
남서울대학교 컴퓨터학과 교수
2005년 8월~2006년 8월
한국정보통신대학교(ICU) 연구교수
1992년 1월~2002년 2월
한국전자통신연구원(ETRI) 연구원
2005년 8월 성균관대학교
전기전자컴퓨터공학과(공학박사)
1992년 2월 성균관대학교 정보공학과(공학석사)
1990년 2월 성균관대학교 정보공학과(공학사)

관심분야 : 초고속통신망, 차세대융합망
E-mail : hckim@msu.ac.kr

논문접수일 : 2009년 11월 1일
수 정 일 : 2009년 11월 20일 (1차)
2009년 12월 5일 (2차)
게재확정일 : 2009년 12월 9일