

논문 2010-47CI-6-12

# 멀티미디어 데이터 특성 모델링에 기반한 네트워크 트래픽 생성기의 구현

( Design of Traffic Generator Based on Modeling of Characteristic of  
Multimedia Data )

김진혁\*, 신광식\*\*, 윤완오\*, 이창호\*, 최상방\*\*\*

( Jin-Hyuk Kim, Kwang-Sik Shin, Wan-Oh Yoon, Chang-Ho Lee, and Sang-Bang Choi )

## 요약

네트워크 트래픽 분석 및 모델링은 그 중요성에 따라 많은 연구가 진행되어온 영역이다. 기존의 네트워크 트래픽 분석 및 모델링 관련 연구는 간단한 패킷 스트림을 전송하거나 HTTP와 같은 특정 어플리케이션의 트래픽 특성을 반영하는데 초점이 맞춰져 있다. 본 논문에서는 멀티미디어 데이터의 특성을 반영한 네트워크 트래픽 생성기를 제안한다. 멀티미디어 데이터의 특성을 반영하기 위해 대표적인 멀티미디어 콘텐츠인 온라인 게임의 트래픽을 분석하기 위해서 온라인 게임의 패킷 전송간 시간과 패킷 크기 확률변수들에 따른 분포를 모델링 하였으며 그 모델을 입력으로 갖는 네트워크 트래픽 생성기의 구조를 제시하였다. 제안된 네트워크 트래픽 생성기를 이용하여 L4D(Left4Dead), WoW(World of Warcraft)의 트래픽을 생성해 본 결과 실제 데이터와 유사도가 높은 분포를 갖는 것을 확인 할 수 있었다.

## Abstract

A study on network traffic analysis and modeling has been exclusively done due to its importance. However, conventional studies on network traffic analysis and modeling only focus on transmitting simple packet stream or traffic features of specific application, such as HTTP. In this paper, we propose a network traffic generator, which reflects the characteristics of multimedia data. To analyze the traffics of online game, which is one of the most popular multimedia contents, we modeled the distribution according to the time between packets and packet size random variable and designed the traffic generator which has the model for input. We generated the traffics of L4D(Left4Dead), WoW(World of Warcraft) with proposed network traffic generator and we found that the generated traffics have similar distributions with real data.

**Keywords :** traffic generator, traffic analysis, analytical model, online gaming traffic

## I. 서론

인터넷 인구의 증가와 컴퓨팅 기술의 발달로 멀티미

디어 서비스의 수요는 지난 몇 년간 꾸준히 증가하는 추세이다. 특히 최근에는 온라인 게임과 같은 고화질의 3D 멀티미디어 콘텐츠에 대한 수요가 늘어나고 있다<sup>[1]</sup>.

이는 필연적으로 멀티미디어 트래픽의 증가를 유발시키며 월드 와이드 웹 트래픽에 큰 영향을 미치고 있다. 이와 같은 변화는 여러 가지 형태로 네트워크 행동 양식에 변화를 주고 있으며 네트워크의 성능과 QoS에 큰 영향을 끼친다. 온라인 게임 트래픽과 같은 멀티미디어 트래픽은 HTTP, FTP 등의 데이터 트래픽과는 다른 특징을 보인다. 일반적으로 온라인 게임은 짧은 지연시간을 요구하며 다양한 게임 상황의 갱신을 위하

\* 학생회원, \*\*\* 평생회원, 인하대학교 전자공학과  
(Dept. of Electronic Engineering Inha University)

\*\* 정회원, 한국전자통신연구원 HD게임연구팀  
(Electronics and Telecommunications Research  
Institute)

※ 본 연구는 지식경제부(MKE, Korea) 산업원천기술  
개발사업(KI002095, 대규모 가상유저의 시나리오 기  
반 제어를 통한 온라인 게임 품질검증 기술개발)의  
지원으로 수행하였음.

접수일자: 2010년9월4일, 수정완료일: 2010년10월25일

여 크기가 작은 패킷을 빈번하게 전송한다<sup>[2]</sup>.

기존의 네트워크 트래픽 생성기 관련 연구는 간단한 패킷 스트림을 전송하거나 웹과 관련된 인터넷 어플리케이션의 트래픽 특성을 반영하는데 초점이 맞춰져 있기 때문에 멀티미디어 트래픽과 같이 예측 불가능한 요청이 빈번한 트래픽에는 적용시킬 수 없다. 따라서 기존의 트래픽 생성기들은 인터넷 서비스의 트래픽으로 인한 네트워크 변화 분석이나 시스템의 안정성 테스트 용 도구로 사용하기에 적합하지만 온라인 게임과 같은 멀티미디어 데이터로 인한 트래픽에 적용시키기에는 부적절하다. 이에 본 논문에서는 멀티미디어 데이터의 특성을 반영한 트래픽 생성기를 구현하기 위하여 새로운 모델링 방법을 제안하고 제안된 방법으로 생성된 모델을 입력으로 가지는 트래픽 생성기 구조를 설계하였다. 본 논문에서는 실제 패킷을 캡처하여 분포의 특징을 분석한 뒤 데이터 분포의 가중치를 이용하여 데이터 영역을 분리하고 변환하는 방법을 사용하여 멀티미디어 데이터의 특성을 잘 표현하는 해석적 모델을 정립하였고 변환 테이블을 이용하여 트래픽 생성시 생길 수 있는 오차를 최소화 하였다. 본 논문은 다음과 같이 이루어져 있다. II장에서는 기존 트래픽 생성기들을 설명하고 III장에서는 멀티미디어 트래픽의 특성을 분석하여 모델링한다. IV장에서는 본 논문에서 제안하는 트래픽 생성기의 구조를 소개하고 V장에서는 트래픽 생성결과 및 분석을 다룬다. 마지막으로 VI장에서 결론을 맺는다.

## II. 관련 연구

네트워크 트래픽 분석 및 모델링은 그 중요성에 따라 많은 연구가 진행되어온 영역이다. 현재까지의 트래픽 생성기와 관련된 연구들은 간단한 패킷 스트림을 전송하거나 특정한 어플리케이션의 트래픽 특성을 반영하는데 초점이 맞춰져 있다. 이 장에서는 대표적인 트래픽 생성기에 대해 설명한다.

### 2.1 패킷 스트림을 이용한 트래픽 생성기

간단한 패킷 스트림을 전송하는 방법을 이용하는 것 중 가장 대표적인 iperf<sup>[3]</sup>는 일정 시간간격으로 순차적으로 패킷을 전송한다. 이 방법은 라우터의 성능을 측정하는데 적합한 방법으로 RFC 2544, RFC2889에도 정의되어 있다. 패킷 스트림을 이용하는 또 다른 대표적인 예는 무한정의 데이터를 FTP로 전송하는 방법이 있

다. 이 방법은 주로 ns-2<sup>[4]</sup>와 같은 시뮬레이터에서 트래픽을 생성 시킬때 사용하는 방법이다. 이렇게 간단한 패킷 스트림을 전송하는 방법은 간단한 네트워크의 대역폭을 측정하는데는 효율적일 수 있으나 복잡하고 상호적인 실제 네트워크의 행동을 분석하기에는 부족하다<sup>[5]</sup>. 따라서 실제 네트워크에 미치는 영향을 분석하기 위해서는 네트워크에 존재하는 실제 트래픽을 모델링하여 생성하는 것이 필요하다.

### 2.2 트래픽 모델링을 통한 트래픽 생성기

SURGE (Scalable URL Reference Generator)<sup>[6]</sup>는 웹 서버의 성능 및 안정성 테스트를 위하여 HTTP 트래픽을 발생시켜 네트워크 부하를 생성한다. 웹서버는 연결 요청이 많을수록 프로세서 부하에 의하여 성능이 감소된다. SURGE는 이러한 특성을 이용하여 웹서버의 부하를 테스트 할 수 있게 설계되었다. 또한 SURGE는 서버의 파일크기 분포, 요청 크기 분포, 동일 파일의 요청 회수 분포에 따른 모델링을 통하여 실제 사용자가 웹서버에 접속하여 사용하는 것 같은 웹 부하를 생성시킬 수 있다.

D-ITG(Distributed Internet Traffic Generator)<sup>[7]</sup>는 TCP, UDP, ICMP, DNS, Telnet, VoIP(G.711, G.723, G.729, Voice Activity Detection, Compressed RTP)등 어플리케이션 레이어의 다양한 프로토콜을 이용하여 트래픽을 생성할 수 있는 도구이다. D-ITG는 패킷 출발 시간과 패킷 크기 분포함수를 이용하여 정의된 패턴을 정확하게 생성해 낼 수 있다. 또한 생성된 트래픽을 통하여 OWD(One Way Delay), RTT(Round Trip Time) 등의 종단간 지연을 측정 할 수 있으며 패킷 손실, 지터 및 전송률을 측정 할 수 있다. D-ITG는 분산 시스템을 이용하여 로그 서버를 분리하여 로그로 인한 작업 부하가 패킷 생성 또는 각종 측정 부하에 영향을 미치지 않게 하였다.

T-mix<sup>[8]</sup>는 ns-2 시뮬레이터용 트래픽 생성기이다. 실제로 트래픽을 생성하지는 않지만 P2P(Peer-to-Peer) 파일 공유, 이메일 서버 접근등의 HTTP 트래픽을 모델링하여 시뮬레이터 내에서 웹부하를 생성시킨다. 트래픽 생성기로 인해 생성되는 트래픽은 실제 네트워크에 존재하는 트래픽과 유사해야 한다. T-mix는 실제 네트워크에 존재하는 TCP 패킷의 헤더정보를 입력받아 분석하여 입력받은 네트워크 내에 존재하는 TCP연결을 모델링한다. 이 과정에서 실제 네트워크의 소켓과

T-mix ns-2 모듈과의 연결을 위하여 connection vector를 사용한다.

### III. 멀티미디어 트래픽 특성분석 및 모델링

이 장에서는 멀티미디어 트래픽의 특성을 이해하기 위하여 대표적인 온라인 멀티미디어 서비스인 온라인 게임의 패킷 크기와 패킷 출발 간 시간(IDT-Inter Departure Time)을 분석하여 새로운 모델링 기법을 제안한다.

#### 3.1 멀티미디어 트래픽 특성분석

그림 1은 온라인 게임의 대표적인 게임인 Quake3, WoW, Quake4, L4D(Left4Dead)의 패킷 크기 분포를 나타낸다. 그림 1의 (a)와 같이 1999년 발표된 Quake3의 경우 패킷 크기의 종류가 한정적이고 분포역시 대체적으로 가우시안 분포에 가까운 것을 볼 수 있다. 이러한 분포는 기존의 웹 트래픽 분포와 비슷하여 쉽게 분석이 가능하고 모델링할 수 있다. 하지만 그림 1의 (b), (c), (d)와 같이 최근에 발표된 WoW(2004), Quake4(2005), L4D(2008)의 경우에는 확연히 다른 양상을 보인다. 이들의 패킷 크기 분포는 특정 분포를 따르지 않

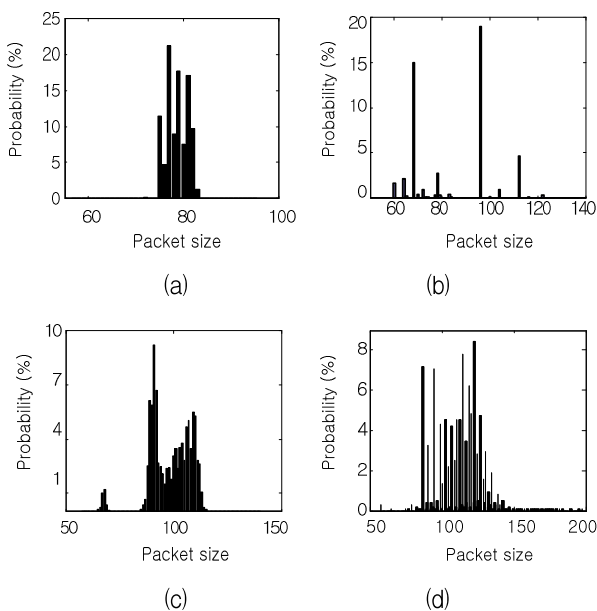


그림 1. 온라인 게임트래픽의 패킷 크기 분포  
 (a)Quake3, (b)WoW, (c)Quake4, (d) L4D  
 Fig. 1. Packet size distribution of online gaming traffic.  
 (a)Quake3, (b)WoW, (c)Quake4, (d) L4D

으며 부분적인 유사성을 갖기 보다는 불규칙적인 데이터 분포를 갖는것을 알 수있다. 이러한 불규칙한 분포를 하나의 분포 함수로 모델링하거나 특정 범위를 갖는 몇 개의 그룹으로 나누어 모델링하기에는 어려움이 존재한다. 이에 본 논문에서는 불규칙한 분포를 갖는 멀티미디어 트래픽을 위한 모델링 방법을 제안한다.

#### 3.2 제안된 멀티미디어 트래픽 모델링

온라인 게임과 같은 멀티미디어 콘텐츠에 대한 수요가 늘어나면서 멀티미디어 트래픽의 모델링 방법에 대한 연구가 수행되었다. 멀티미디어 트래픽 분포를 하나의 분포함수로 모델링 하는 방법은 웹 트래픽을 분석하는 방법과 유사하게 멀티미디어 트래픽의 특성을 모델화 하는 것이다<sup>[9-11]</sup>. 하지만 이들 모델들은 멀티미디어 트래픽의 특성을 전체적으로 반영하는 것이 아니라 주요 특징만을 반영 할 수 있었다. 이에 따라 멀티미디어 트래픽 분포를 지역적으로 몇 개의 데이터 집합으로 나누어 각각에 맞는 분포함수로 모델링 하는 연구가 진행되었다<sup>[12-13]</sup>. 이러한 방법은 지역적으로 분포가 유사한 데이터를 모델링 하기에는 적합하지만 지역적으로 분포의 유사성이 존재하지 않는 데이터들의 전체 특성을 나타내기에는 부족하다.

본 논문에서 제안하는 모델링 방법은 패킷 캡처 및 통계 분석, 패킷 분포 변형, 커브 피팅, 트래픽 생성의 4 단계를 갖는다. 이중 패킷 캡처 및 통계분석은 기존의 통계 분석방법과 유사하며 커브 피팅 방법은 피팅 알고리즘의 수행으로 이루어졌다. 따라서 패킷 분포 변형과 트래픽 생성이 본 논문에서 집중하고자 하는 부분이며 트래픽 생성은 VI장에서 다룰 것이기 때문에 이번 장에서는 패킷 분포 변형 과정에 대해서 자세히 다룬다.

앞 절에서 언급했듯이 멀티미디어 트래픽은 하나의 분포 함수로 모델링하거나 특정 범위를 갖는 몇 개의 그룹으로 나누어 모델링하기에는 어려움이 존재한다. 따라서 본 논문에서는 분포의 가중치에 따라서 전체 데이터 집합을 부분적인 데이터 집합으로 분리한 후 각각의 데이터 집합을 변환하여 부분적인 데이터 영역을 생성하는 방법을 사용하였다. 데이터 집합의 분리는 이동평균법(Moving Average)을 사용하였다. 전체 데이터 집합의 pdf(Probability Density Function)에 분포 이동평균법을 적용한 후 이동평균값을 임계치로 하여 전체 데이터 집합을 주 데이터 집합과 부 데이터 집합으로 분리하였다. 본 논문에서 사용된 이동평균법은 수식 1과 같이 정의된다.

$$M_t = \frac{W_{\lfloor \frac{n}{2} \rfloor} * P_{t-\lfloor \frac{n}{2} \rfloor} + \dots + W_0 * P_t + \dots + W_{\lfloor \frac{n}{2} \rfloor} * P_{t+\lfloor \frac{n}{2} \rfloor}}{n * \sum W_n} \quad (1)$$

$$W_0 = \lfloor \frac{n}{2} \rfloor$$

$$W_{t+1} = W_t - 1$$

수식 1에서  $M_t$ 는 t번째 데이터에서의 이동평균값을 나타내며  $n$ 은 윈도우 크기를  $P_t$ 는 t번째 데이터 값을  $W$ 는 가중치 값을 나타낸다. 이동평균법은 고르지 못한 시계열 자료들을 부드럽게 하여 전체 구간의 변화를 살펴보고자 할 때 사용하는 방법으로 슬라이딩 윈도우를 이용하여 윈도우 내의 데이터들에 가중치를 곱하여 평균을 내는 방법이다. 이동 평균을 계산하는 데 사용되는 윈도우 내의 데이터들은 위치에 따라 다른 가중치를 부여받는다. 윈도우 중앙에 위치한 데이터는 가장 큰 가중치를 부여 받으며 윈도우의 양 끝으로 갈수록 작은 가중치가 적용된다. 이렇게 계산된 이동평균값은 전체 데이터 집합에서 주 데이터 집합과 부 데이터 집합을 분리하는 기준이 된다.

그림 2는 그림 1의 (d)에 나타난 L4D의 패킷 크기 분포에 수식 1을 적용하여 이동평균법을 적용한 모습을 보여준다.

주 데이터 집합과 부 데이터 집합을 분리한 후 각각의 데이터 집합을 이용하여 pdf를 작성한다. 이렇게 생성된 pdf에서 확률밀도 값을 갖지 않는 확률변수들을 제외시켜 인덱싱한 후 원래 확률변수와 인덱스 값의 변환 테이블을 작성한다. 변환 테이블은 인덱스 정보 이외에 현재 테이블내의 확률변수가 전체 데이터 집합에서 차지하는 비율을 포함한다. 변환 테이블을 작성한 후 변환 테이블의 인덱스 값과 확률밀도사이의 함수를 작성

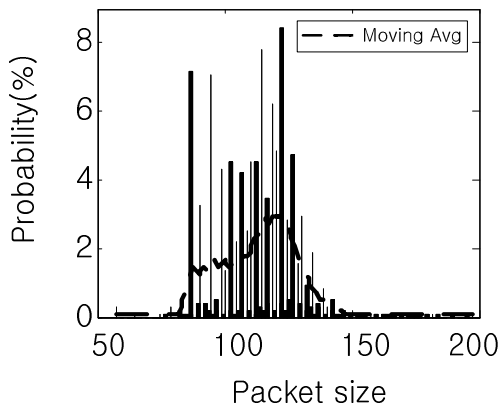


그림 2. L4D 패킷 크기 분포 함수에 이동평균법 적용  
Fig. 2. Moving Average of L4D Packet size distribution.

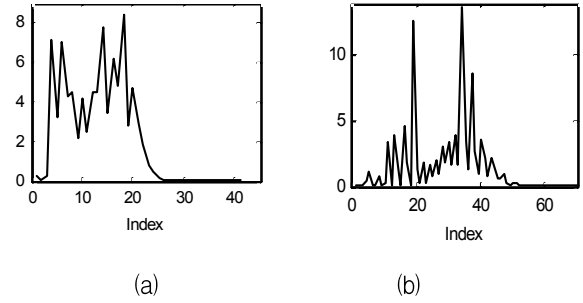


그림 3. 각 영역의 변환된 pdf  
(a)주 데이터 영역, (b)부 데이터 영역  
Fig. 3. Transformed pdf of each data domain.  
(a)Major data domain, (b)Minor data domain

하여 변환 pdf를 생성한다. 그림 3은 변환된 주 데이터 영역과 부 데이터 영역의 변환 pdf를 보여준다.

위와 같이 데이터 영역을 변환한 뒤 각 영역의 변환된 pdf를 이용하여 커브피팅을 수행한다. 본 논문에서는 MATLAB의 curve fitting toolbox를 이용하여 Levenberg-Marquardt 알고리즘을 적용하였다. Levenberg-Marquardt 알고리즘은 주요 피크들을 각각 가우시안 분포 함수로 피팅하여 전체 함수를 가우시안 분포함수의 합으로 나타낸다. 피팅을 마친 후 오차를 측정하여 오차가 최소가 되도록 반복해서 수행한다. 커브 피팅을 통해 모델링된 함수는 다음과 같이 정의 된다.

$$y = y_0 + \sum \frac{a_i}{c_i \times \sqrt{\frac{\pi}{2}}} \times e^{-2(x-b_i)^2/c_i^2} \quad (2)$$

수식 2에서 생성되는 가우시안 분포함수의 수는 Levenberg-Marquardt 알고리즘으로 결정된 주요 피크의 개수에 의해 정해진다.  $y_0$ 는 가우시안 분포함수의 초기값이며 각각의 가우시안 함수의 파라미터  $a_i, b_i, c_i$ 는 커브피팅을 통해 구해진다.

#### IV. 트래픽 생성기 구조

이 장에서는 본 논문에서 제안하는 멀티미디어 데이터의 특성을 반영한 트래픽 생성기의 구조에 대해서 설명한다. 그림 4는 멀티미디어 데이터 특성을 반영한 트래픽 생성기의 구조를 나타낸다. 그림 4에서 입력으로 주어진 해석적 모델(analytical model)은 III장에서 제안된 분포의 가중치를 이용하여 데이터 영역을 분리하고 변환하는 방법을 이용하여 정립한 모델이다. 해석적 모델은 변환 테이블과 변환된 pdf의 쌍으로 이루어진 부

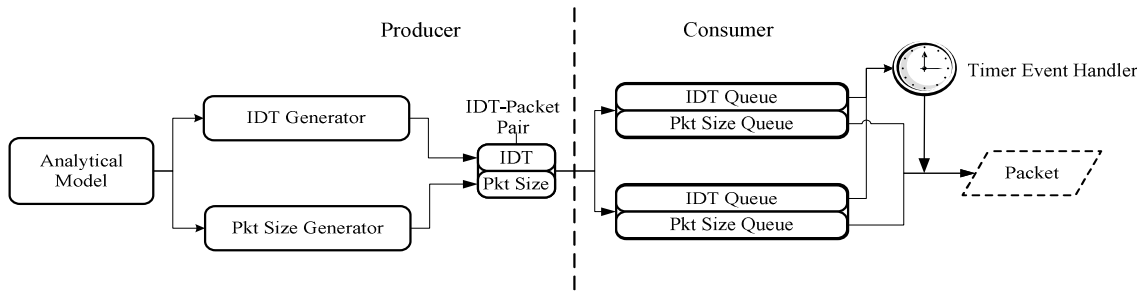


그림 4. 제안된 트래픽 생성기 구조  
Fig. 4. The architecture of the proposed traffic generator.

분적인 데이터 영역들로 이루어져 있다. 패킷 출발 간 시간 생성기(IDT Generator) 및 패킷 크기 생성기(PktSize Generator)는 온라인 게임의 패킷 크기와 패킷 출발 간 시간 모델을 이용하여 멀티미디어 데이터의 특징을 가진 온라인 게임 패킷 확률변수를 생성한다. 패킷 출발 간 시간 생성기는 멀티미디어 데이터 패킷 발생기능을 가진 타이머 이벤트 핸들러를 동작시키기 위한 패킷 출발 간 시간 간격을 생성해낸다. 패킷 출발 간 시간이 만료되면 타이머 이벤트 핸들러가 활성화되고 타이머 이벤트 핸들러는 현재 가지고 있는 패킷 크기 값을 갖는 온라인 게임 데이터 패킷을 발생시킨다. 패킷을 발생시킨 후에 타이머 이벤트 핸들러는 패킷 큐에서 새로운 확률변수를 꺼내어 타이머를 설정한다. 패킷 출발 간 시간 생성기 및 패킷 크기 생성기에서 생성된 IDT와 PktSize는 서로 짝을 이루어(IDT-PktSize 짝) 패킷 큐에 저장된다. 패킷 큐는 이중 버퍼구조로 설계되었기 때문에 타이머 이벤트 핸들러는 확률변수 생성기가 확률변수 생성을 하는 도중에도 패킷 큐에서 IDT-PktSize 짝을 이용하여 멀티미디어 데이터 패킷을

생성 할 수 있다.

제안된 트래픽 생성기는 크게 확률변수를 생성(Producer)하는 부분과 소비(Consumer)하는 두 부분으로 나눌 수 있다. 그림 5는 생성부분의 프로시저 의사코드를 나타내며 그림 6은 소비부분의 프로시저를 나타낸다.

확률변수를 생성하는 프로시저는 초기화 단계로 두 개의 패킷 큐 중에서 저장 작업을 할 버퍼를 결정하고 해당 큐가 비었는지 확인한다. 만약 해당 큐가 비어있지 않다면 프로시저를 중단하고 비어있는 패킷 큐를 찾을 때까지 다시 수행한다. 비어있는 큐를 선택하여 쓰기 버퍼로 선택한 후에는 입력으로 주어진 해석적 모델을 이용하여 패킷 출발 간 시간, 패킷 크기 확률변수를

```

Producer Thread Procedure
1  Set Q = Available Write Packet Queue
2  if Q is empty then
3    while Q is not full
4      Generate IDT & PktSize Random Variable
5      Pairing IDT & PktSize Random Variable
6      Enqueue IDT-PktSize pair
7    endwhile
8  endif
    
```

그림 5. 확률변수 생성 부분의 의사코드  
Fig. 5. Pseudo code of phase of producing random variable.

```

Consumer Thread Procedure
1  Set Q = Available Read Packet Queue
2  if Q is empty then
3    Set Q = Available Write Packet Queue
4    Set other Q = Available Read Packet Queue
5  else
6    while Q is empty
7      Dequeue IDT-PktSize pair
8      Set Timer event
9      while Timer is not expired
10     continue
11    endwhile
12    Packet Transmission
13  endwhile
14 endif
    
```

그림 6. 확률변수 소비 부분의 의사코드  
Fig. 6. Pseudo code of phase of consuming random variable.

생성하고 생성된 패킷 출발 간 시간 확률변수와 패킷 크기 확률변수를 한 쌍의 그룹으로 묶어 큐에 추가한다. 해당 큐가 가득 찰 때까지 IDT-PktSize 쌍의 생성 및 큐에 추가하는 과정을 반복해서 수행하고 생성 프로시저는 종료된다.

확률변수를 소비하는 프로시저는 생성 프로시저와 비슷하게 초기화 단계에서 읽기 작업을 할 패킷 큐를 결정하고 큐가 비었는지 확인한다. 만약 큐가 비어있다면 해당 큐는 읽기 큐로 사용할 수 없기 때문에 해당 큐를 쓰기 큐로 설정하고 나머지 큐를 읽기 큐로 설정한 후 큐가 비었는지 다시 확인한다. 결정된 큐가 비어있지 않다면 해당 큐에서 IDT-PktSize 쌍을 꺼내어 패킷 출발 간 확률변수를 이용하여 타이머 이벤트를 설정한다. 타이머의 종료로 인하여 타이머 이벤트 핸들러가 활성화되면 타이머 이벤트 핸들러는 패킷 크기 확률변수를 이용하여 패킷을 생성하고 전송한다. 소비 프로시저는 IDT-PktSize 쌍을 이용하여 패킷을 생성하고 전송하는 과정을 읽기 큐가 완전히 빌 때까지 반복하고 큐가 완전히 비게 되면 읽기 큐와 쓰기 큐를 바꾸어 다시 수행하게 된다.

## V. 성능 분석 및 평가

이번 장에서는 제안된 트래픽 생성기의 성능 분석 및 평가를 위하여 트래픽 생성기로 발생한 결과와 분석 대상인 온라인 게임 트래픽을 비교한다. 본 논문에서 성능 분석을 위하여 사용된 데이터는 FPS장르의 대표적

인 게임인 L4D와 MMORPG장르의 대표적인 게임인 WoW의 실제 게임 패킷을 사용하였다. Quake4의 경우 같은 FPS장르의 L4D의 분포보다 단순한 분포를 갖기 때문에 본 논문에서는 좀 더 복잡한 분포를 갖는 L4D를 선택하여 성능 분석 및 평가에 사용하였다. 성능 분석을 위해 L4D와 WoW의 클라이언트 단에서 캡처한 패킷의 패킷 출발 간 시간 및 패킷 크기에 대하여 분포를 분석한 후 III장에서 제안된 모델링 기법을 이용, 해석적 모델을 정립하여 트래픽을 생성한 후 원본 데이터와 생성된 결과를 비교하였다. 본 논문에서는 분포 그래프 비교와 Q-Q 플롯(Quantile-Quantile plot)을 이용한 분포 적합도 비교를 수행하였다. 분포 그래프 비교를 위하여 L4D와 WoW의 캡처된 원본 데이터와 트래픽 발생기로 생성한 결과 데이터의 패킷 출발 간 시간과 패킷 크기 분포를 그래프로 표현한 뒤 원본 데이터와 생성된 결과의 표준편차를 구하여 원본 데이터와의 차이를 수치적으로 제시하였다. Q-Q 플롯은 두 자료 사이에 유사성을 비교하기 위해 두 분포의 분위수들을 쌍으로 직교좌표에 타점한 그림이다. 따라서 Q-Q 플롯내의 타점이  $y=x$  그래프 위에 존재할수록 두 분포는 유사하다 라는 결론을 내릴 수 있다.

그림 7의 (a)는 L4D의 실제 패킷을 캡처하여 패킷 출발 간 시간 분포를 히스토그램으로 나타낸 것이고, (b)는 트래픽 생성기로 생성한 패킷을 캡처하여 패킷 출발 간 시간 분포를 나타낸 것이다. 주요 특징인 큰 피크값주변이 동일하게 분포하는 것을 확인할 수 있고 주요특징 이외의 IDT 50주변 작은 피크들도 0.001 이하의

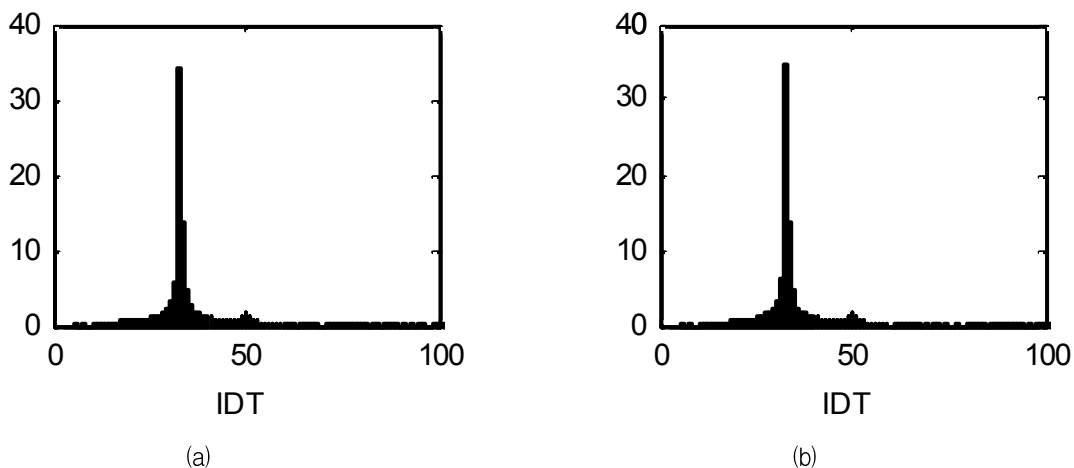


그림 7. L4D 패킷 출발 간 시간 분포 비교 (a) 원본데이터 분포, (b) 생성된 결과 분포  
Fig. 7. Comparison of L4D IDT distribution (a) Original data, (b) Produced result.

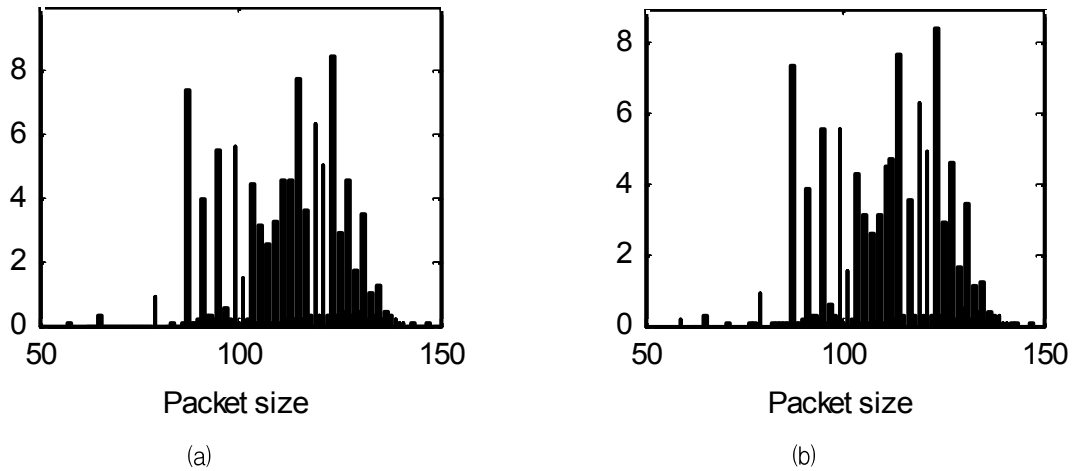


그림 8. L4D 패킷 크기 분포 비교 (a) 원본데이터 분포, (b) 생성된 결과 분포  
 Fig. 8. Comparison of L4D packet size distribution (a) Original data, (b) Produced result.

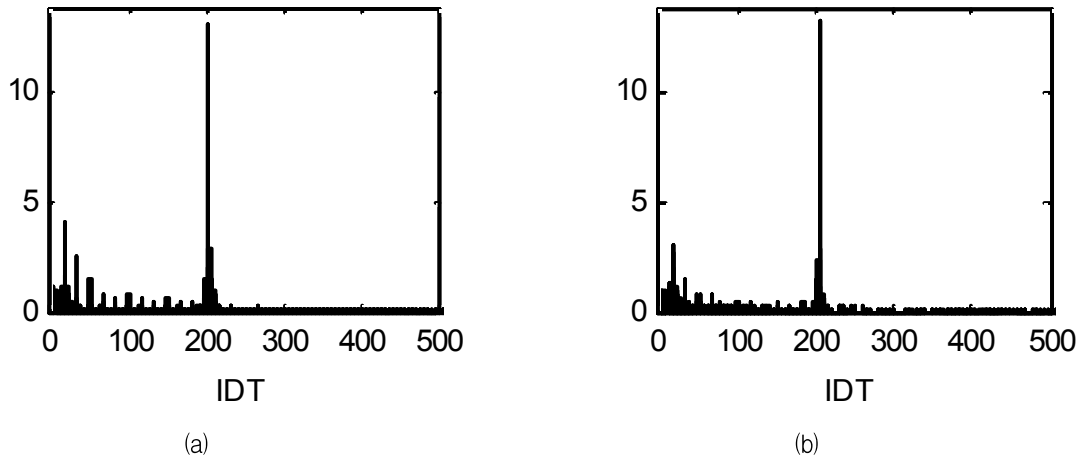


그림 9. WoW 패킷 출발 간 시간 분포 비교 (a) 원본데이터 분포, (b) 생성된 결과 분포  
 Fig. 9. Comparison of WoW IDT distribution (a) Original data, (b) Produced result.

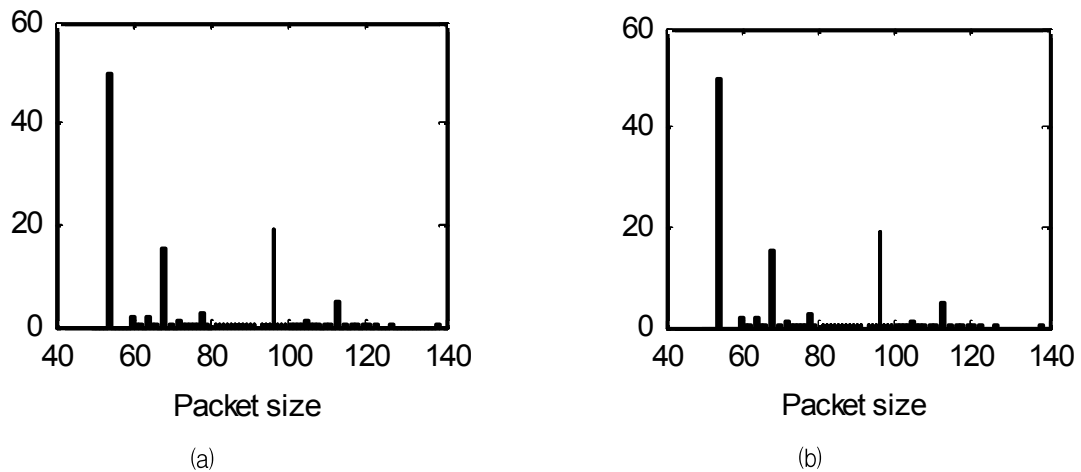


그림 10. WoW 패킷 크기 분포 비교 (a) 원본데이터 분포, (b) 생성된 결과 분포  
 Fig. 10. Comparison of WoW packet size distribution (a) Original data, (b) Produced result.

오차를 갖는 것을 확인할 수 있다. 그림 8은 L4D의 패킷 크기 분포 비교를 보여준다. L4D 패킷 크기 분포는

IDT분포보다 주요 특징점의 수가 많고 복잡한 분포를 갖는다. 제안된 방법을 사용하였을 때 복잡하고 이산적

인 분포를 가질 때에도 전체적인 분포 뿐만 아니라 주요 특징과 세부 특징이 유사하게 표현되는 것을 확인할 수 있다. L4D 트래픽과 생성된 결과와의 패킷 출발 간 시간 분포의 표준편차는 1.364이고 패킷 크기 분포의 표준편차는 11.623으로 데이터의 자유도에 비해 낮은 수치를 보인다. 패킷 출발 간 시간 분포의 경우 패킷 크기 분포보다 분포 함수가 단순하기 때문에 표준편차 값이 작은 특성을 보인다. 그림 9와 그림 10은 WoW의 패킷 출발 간 시간 분포 비교와 패킷 크기 분포 비교를 보여준다. WoW의 패킷 출발 간 시간 분포는 데이터의 분포 범위가 넓어 자유도가 매우 높게 나타남에도 불구하고 2.789의 작은 표준편차 값을 갖는다. 이는 제안된 방법을 이용하였을 때 확률변수의 생성 범위가 넓어도 원본 데이터의 특성을 잘 나타낼 수 있다는 것을 보여준다. WoW의 패킷 크기 분포의 표준편차는 6.565으로 작은 값을 가지며 원본 데이터의 특성을 잘 재현하는 것을 그림 10을 통해 확인할 수 있다. 그림 11, 12는 L4D와 WoW의 원본 데이터와 생성된 결과 데이터 사이의 Q-Q 플롯을 보여준다. 원본 데이터의 분포와 지역적으

로 차이가 있는 부분들이 존재하지만 나머지 부분들의 기울기가  $y = x$ 를 따르는 것은 작은 지역적 오차만 있다는 것을 알 수 있다. 또한 확률 분포의 주요 특징점 뿐만 아니라 머리 및 꼬리 특성까지 유사한 특성을 갖는 것을 확인할 수 있다.

### VI. 결 론

본 논문에서는 멀티미디어 데이터의 특성을 반영한 트래픽 생성기를 제안하였다. 멀티미디어 데이터 트래픽을 발생시키기 위하여 대표적인 멀티미디어 데이터인 온라인 게임의 트래픽을 캡처하여 특성을 분석하였으며 패킷 전송 간 시간과 패킷 크기 확률변수들에 따른 분포를 분석하고 모델링하여 온라인 게임 트래픽의 특성을 반영한 해석적 모델을 정립하였다.

또한 정립한 모델을 이용하여 트래픽을 효율적으로 생성하기 위해서 트래픽 생성기의 구조를 설계하였으며 원본 데이터와 분포를 비교하였다.

원본 데이터와 생성된 결과를 분포 그래프로 나타내었을 때 대체적으로 편차가 적은 것을 확인할 수 있었으며 두 데이터 사이의 유사성을 판단하는 도구인 Q-Q 플롯을 이용하여 원본 데이터와 생성된 데이터 분포 사이에 높은 유사성을 갖는 것을 확인하였다.

### 참 고 문 헌

- [1] 한국 콘텐츠 진흥원, “2009 대한민국 게임백서”, 538-543쪽, 2009년 9월.
- [2] W. C. Feng, F. Chang, W. C. Feng, et.al, “A Traffic Characterization of Popular On-line Games”, *IEEE/ACM Transactions on Networking*, vol. 13, no. 3, pp.488-499, Jun. 2005.
- [3] The iperf TCP/UDP Bandwidth Measurement Tool, <http://dast.nlanr.net/Projects/Iperf>
- [4] Ucb/lbnl/vint network simulator ns-2 <http://www.isi.edu/nsnam/ns/>
- [5] Sommers, J., Kim, H., Barford, P., “Harpoon: a flow-level traffic generator for router and network tests”, *In Proceedings of the Joint international Conference on Measurement and Modeling of Computer Systems*, pp. 392-392, Jun. 2004.
- [6] Barford, P., Crovella, M. “Generating representative Web workloads for network and server performance evaluation”, *In Proceedings*

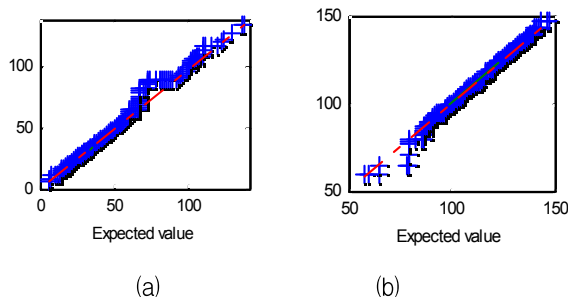


그림 11. L4D 데이터에 대한 Q-Q 플롯  
(a) 패킷 출발 간 시간 분포 비교,  
(b) 패킷 크기 분포 비교

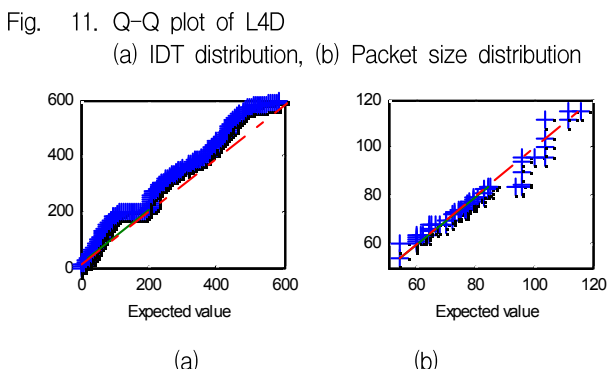


그림 12. WoW 데이터에 대한 Q-Q 플롯  
(a) 패킷 출발 간 시간 분포 비교,  
(b) 패킷 크기 분포 비교

Fig. 12. Q-Q plot of WoW  
(a) IDT distribution, (b) Packet size distribution



- of the 1998 ACM SIGMETRICS Joint international Conference on Measurement and Modeling of Computer Systems*, pp. 151-160, Jun. 1998.
- [7] "Analysis and Experimentation of an Open Distributed Platform for Synthetic Traffic Generation." *In Proceedings of the 10th IEEE international Workshop on Future Trends of Distributed Computing Systems*, pp. 277-283, May. 2004.
- [8] Weigle, M., Adurthi, P., Hernández-Campos, F., Jeffay, K., Smith, F., "Tmix: a tool for generating realistic TCP application workloads in ns-2", *SIGCOMM Comput. Commun. Rev.* 36, pp. 65-76, Jul. 2006.
- [9] J. Färber, "Network Game Traffic Modeling," *Proceedings of the 1st workshop on Network and system support for games*, pp.53-57, Apr. 2002.
- [10] T. Lang, G. Armitage, P. Branch, et al., "A Synthetic Traffic Model for Half-Life," *Proceedings of the Australian Network and Telecommunications Conference*, 2003.
- [11] T. Lang, P. Branch, and G. Armitage, "A Synthetic Traffic Model for Quake3," *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*, vol. 74, pp.233-238, 2004.
- [12] M.S. Borella, "Source Models of Network Game Traffic," *Computer Communications*, vol. 23, no. 4, pp. 403-410, Feb. 2000.
- [13] R. Antonello, S. Fernandes, J. Moreira, et al., "Traffic analysis and synthetic models of SecondLife", *Springer Multimedia Systems Journal*, vol. 15, no. 1, pp. 33-47, Feb. 2009.

저 자 소 개



**김진혁**(학생회원)  
 2009년 인하대학교 전자공학과  
 학사 졸업.  
 2009년~현재 인하대학교  
 전자공학과 석사과정.  
 <주관심분야 : 멀티미디어 통신,  
 무선 통신, 컴퓨터 네트워크, 병렬  
 및 분산 컴퓨팅>



**신광식**(정회원)  
 2001년 인하대학교 전자공학과  
 학사 졸업.  
 2003년 인하대학교 전자공학과  
 석사 졸업.  
 2008년 인하대학교 전자공학과  
 박사 졸업.  
 2008년~2010년 9월 한국전자통신연구원 HD  
 게임연구팀 Post doc.  
 2010년 9월~현재 한국전자통신연구원 HD 게임  
 연구팀 선임연구원  
 <주관심분야 : 멀티미디어 통신, 무선 통신, 컴퓨  
 터 네트워크, 병렬 및 분산 컴퓨팅>



**윤완오**(학생회원)  
 2000년 경기대학교 전자공학과  
 학사 졸업.  
 2002년 인하대학교 전자공학과  
 석사 졸업.  
 2010년 인하대학교 전자공학과  
 박사 졸업.  
 <주관심분야 : 분산 처리 시스템, 병렬프로그래  
 밍, 컴퓨터 구조, 무선 통신, 컴퓨터 네트워크>



**이창호**(학생회원)  
 2008년 청주대학교 전자공학과  
 학사 졸업.  
 2010년 인하대학교 전자공학과  
 석사 졸업.  
 2010년~현재 인하대학교 전자공  
 학과 박사과정  
 <주관심분야 : 컴퓨터 구조, 컴퓨터 네트워크, 무  
 선 통신, 병렬 및 분산 처리 시스템>



**최상방**(평생회원)  
 1981년 한양대학교 전자공학과  
 학사 졸업.  
 1981년~1986년 LG 정보통신(주).  
 1988년 University of washinton  
 석사 졸업.  
 1990년 University of washinton  
 박사 졸업.

1991년~현재 인하대학교 전자공학과 교수  
 <주관심분야 : 컴퓨터 구조, 컴퓨터 네트워크, 무  
 선 통신, 병렬 및 분산 처리 시스템>