학내 네트워크 트래픽의 관리 및 제어를 위한 HTTP 트래픽 분석에 관한 연구

정우석, 권재범, 심규석, 김명섭

고려대학교 컴퓨터정보학과

{hary5832, lhmagic, kusuk007, tmskim}@korea.ac.kr

요 약

최근 학내 네트워크 트래픽이 급격히 증가하고 있다. HTTP 프로토콜을 사용하는 웹 기반의 서비스와 응용 프로그램이 증가하는 추세이다. 또한, 스마트 디바이스의 급속한 성장과 다양한 응용 및 서비스들의 등장으로 무선 트래픽이 급격하게 증가하면서 전체 네트워크 트래픽이 증가하고 있다. 이에 따라 학내 네트워크 트래픽을 관리 및 제어할 수 있는 방안에 대한 연구가 요구된다. 본 논문에서는 Bro 트래픽 분석기에 대해 언급하고, Bro 트래픽 분석기를 이용하여 학내 HTTP 트래픽을 분석한다. 또한 분석 결과를 통해 사용되는 브라우저의 종류, 응용의 종류, 데이터형태, 스마트 디바이스의 사용 비율 등 학내 네트워크 사용 경향을 파악한다.

1. 서론

HTTP 프로토콜을 사용하는 웹 기반의 서비스와 응용 프로그램의 증가뿐만 아니라 특정 P2P 서비스에서 네트워크 방화벽에 의한 차단을 피하기 위하여 HTTP 프로토콜을 사용하기 때문에 HTTP 트래픽은 급격하게 증가하고 있다. 또한, 스마트폰의 보급이 확대되고, 관련연구[1]에 따르면 학내 스마트 디바이스가 발생하는 트래픽 중 70% 이상이 HTTP 프로토콜이다. 이러한 요인들로 학내 네트워크 HTTP 트래픽이 증가하였다. 따라서 학내 네트워크 HTTP 트래픽을 관리 및 제어 할 수 있는 방법론이 요구된다.[2]

본 논문에서는 학내 네트워크 트래픽을 수집하고, 트래픽을 분석할 수 있는 다양한 분석기 중 Bro 트래픽 분석기를 이용하여 분석한다. Bro 는 오픈 소스 기반의 네트워크 트래픽 분석기로써 트래픽 분석에 대해 소스 IP 주소, 목적지 IP 주소, 연결 상태를 포함한 각 TCP 연결의 자세한 정보들을 제공하여 트래픽 분석에 자주 쓰인다.[3] 분석한 결과를 토대로 학내 HTTP 트래픽의 증가를 증명하기 위해 학내 네트워크 사용자들이 사용하는 브라우저의 종류, 스마트 디바이스가 차지하는 비중 그리고 주로 사용되는 데이터 형태를 파악한다.

본 논문의 구성은 다음과 같다. 본 장의 서론에 이어, 본론에서는 실험에 사용된 학내 네트워크 트래픽의 수집 방법과 Bro 트래픽 분석기에 대해 언급한다. 또한 분석의 결과를 토대로 학내 HTTP 트래픽의 증가를 증명하기 위해 학내 네트워크 사용자들이 사용하는 브라우저의 종류, 응용의 종류, 스마트 디바이스가 차지하는 비중 그리고 주로 사용되는 데이터 형태에 대해 파악한다. 마지막으로 결론 및 향후 연구에 대해 언급한다.

이 논문은 2012 년 정부(교육과학기술부)의 재원으로 한국연구재단(2012R1A1A2007483) 및 2013 년도 정부(미래창조과학부)의 재원으로 한국연구재단-차세대정보.컴퓨팅기술개발사업(2010-0020728)의 지원을 받아 수행된 연구임

2. 본론

본 장에서는 분석에 사용된 학내 네트워크 트래픽의 수집방법과 Bro 트래픽 분석기에 대해 언급하고 분석한 결과를 보여준다.

실험에 사용된 트래픽은 TCP 트래픽 중 port 번호 80 번을 사용하는 트래픽을 HTTP 트래픽으로 간주한다. 학내 네트워크에서의 6 시간(2014 년 4 월 18 일 15 시부터 21 시까지)동안 HTTP 트래픽을 수집한 것이다. 수집된 HTTP 트래픽을 Bro 트래픽 분석기를 통해 분석한다.

Bro 는 오픈 소스 기반의 네트워크 트래픽 분석기이며 성능 측정 및 trouble-shooting 등 넓은 범위의 트래픽 분석 작업을 지원한다. Bro 의 가장 큰 장점 중 하나는 네트워크 활동을 분석해 udp, http, icmp 등으로 분류하여 각각의 로그파일로 기록된다는 것이다.[4] 본 논문에서는 host, uri, user agent 등의 HTTP 정보를 제공하는 HTTP 로그만을 이용한다.

그림 1 은 학내 네트워크에서 발생한 HTTP 트래픽 중 브라우저에 대한 정보를 분석한다. 그림 1 을 통해 과거 매우 높은 점유율을 보였던 Internet Explorer 의 학내 네트워크망 점유율이 50% 이하로 떨어졌음을 알 수 있다. Safari 의 점유율은 31%이다. 웹로그 분석 전문업체인 '넷애플리케이션'의 통계에 따르면 2014 년 4 월 기준으로 스마트 디바이스에서 사용되는 브라우저 중 safari 의 점유율이 53.91%이다. 이를 통해 학내 HTTP 트래픽 중 스마트 디바이스의 사용량 이 증가했음을 알 수 있다.

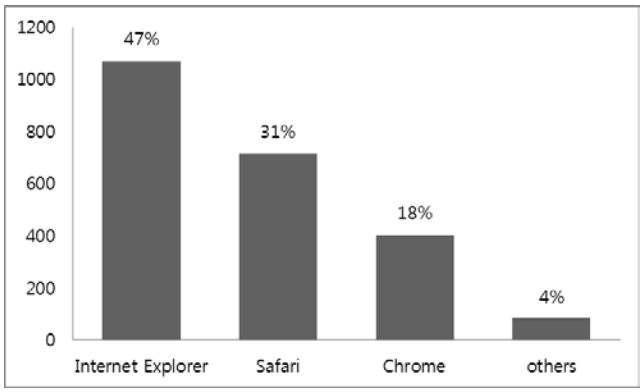

그림 1. 학내 네트워크 HTTP 트래픽의 브라우저 종류

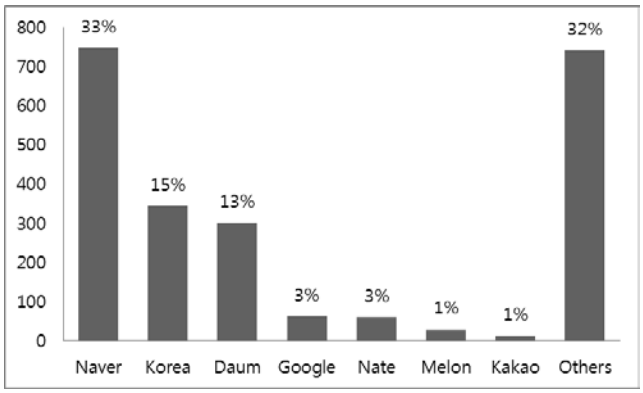

그림 2. 학내 네트워크 HTTP 트래픽의 응용 종류

그림 2 는 학내 네트워크 HTTP 트래픽을 이용한 응용 분류 결과이다. 대형 포탈 사이트들이 선호 되었고, 학내 네트워크 트래픽의 특성상 본교 홈페이지의 사용자 수가 다음으로 많다. 또한, 대형 포탈들의 모바일 버전 응용의 사용이 전체 HTTP 트래픽중 10%로 높은 비중을 차지 하였다.

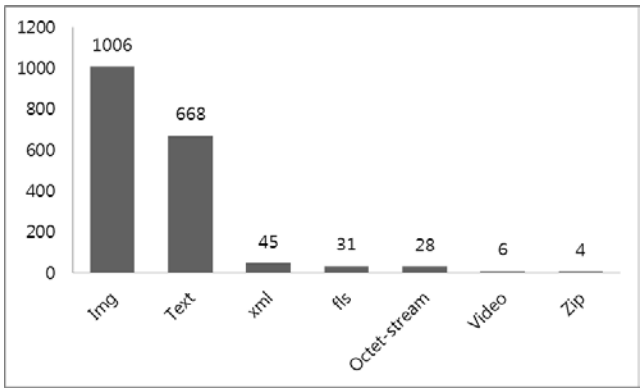

그림 3. 학내 네트워크 HTTP 트래픽의 데이터의 형태

그림 3 은 네트워크에서 사용되는 데이터의 형태의 비율이다. Img, Text 의 형태가 대부분을 차지한다. HTTP 트래픽 급증의 많은 부분을 차지하는 스마트 디바이스 관련한 트래픽이 대부분 웹서핑에 집중되어 있기 때문이다.

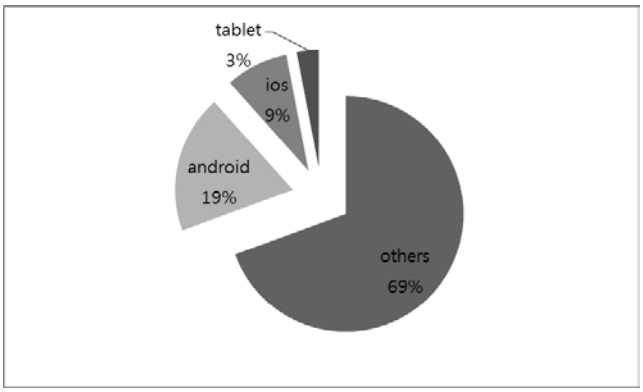

그림 4. 학내 전체 네트워트 HTTP 트래픽 중 스마트 디바이스의 사용 비율

그림 4 는 학내 전체 HTTP 트래픽 중 스마트 디바이스가 차지하는 비율을 보여준다. 학내의 전체 HTTP 트래픽 중 스마트 디바이스에서 발생하는 트래픽이 약 31%로 많이 발생한다. 스마트폰 보급의 확장과 어디서나 간편하게 네트워크에 접속할 수 있는 스마트 디바이스의 특성 때문이다.

3. 결론 및 향후 연구

본 논문에서는 학내 네트워크 트래픽을 Bro 트래픽 분석기를 통해 분석 했었다. 위 분석을 통하여 학내 HTTP 트래픽에서 사용되는 브라우저의 종류, 응용의 종류, 데이터 형태의 비율, 전체 HTTP 트래픽 중 스마트 디바이스의 사용 비율을 알 수 있었다. 그 결과 최근에 급격하게 증가한 학내 네트워크 트래픽은 스마트 디바이스 보급과 사용의 증가 그리고 HTTP 프로토콜을 사용하는 웹 기반의 서비스와 응용 프로그램의 증가가 원인임을 알 수 있었다.

향후 연구에서는 학내 트래픽 증가의 주된 요인인 스마트 디바이스가 발생하는 HTTP 트래픽을 효율적으로 관리 및 제어하는 방안 마련이 필요하다.

4. 참고문헌

[1] Sang-Woo Lee, Jun-Sang Park, Hyun-Shin Lee, and Myung-Sup Kim, "A Study on Smart-phone Traffic Analysis," Proc. of the Asia-Pacific Network Operations and Management Symposium (APNOMS) 2011, Taipei, Taiwan, Sep. 21-23, 2011.

[4] 육종환, 스마트교육을 위한 학교 무선인터넷 환경 구축 타당성분석(2012), Retrieved April, 16, 2012, from http://www.keris.or.kr.

[3] Jaeyeon Jung, Emil Sit, “An Empirical Study of Spam Trafic and the Use of DNS Black Lists”, 4th ACM SIGCOMM conference on Internet measurement, Pages 370-375, Taormina, Sicily, Italy, Oct. 25-27, 2004

[4] Vern Paxson, “Bro: A System for Detecting Network Intruders in Real-Time” 7th USENIX Security Symposium San Antonio, Texas, January 26-29, 1998