 ISSN : 2586-0238 (Online)

상관 관계 분석 방법을 이용한 개선된 Skype 응용 트래픽 탐지 시스템

이성호, 구영훈, 박지태, 지세현, 김명섭
고려대학교

{gaek5, gyh0808, pjj5846, sxzer, tmskim}@korea.ac.kr

Improved Skype Application Traffic Detection System Using Correlation Analysis Method

Sung-Ho Lee, Young-Hoon Goo, Ji-Tae Park, Se-Hyun Ji and Myung-Sup Kim
Korea Univ.

요 약

오늘날 네트워크 환경이 증가하며 응용 및 서비스 별로 발생시키는 트래픽 패턴에 다양한 종류가 생성됨에 따라 트래픽의 응용 및 서비스 별로 분류할 수 있는 트래픽 분석 방법이 연구되고있다. 특히 Skype 는 Microsoft 사에서 서비스하는 P2P 기반의 VoIP 서비스로서 현재 국제적으로 가장 광범위하게 사용되고있다. 이러한 이유로 Skype 트래픽 탐지는 네트워크 망 관리 측면에 있어 그 중요성이 대두되고있다. 기존의 Skype 응용 탐지 시스템이 갖고있었던 한계점 극복과 다양한 Skype 트래픽 패턴을 보다 정확하게 분석하고 탐지하기 위해 본 논문은 패턴, 리스트, 시그니쳐 기반의 응용 탐지 방법을 통한 개선된 Skype 트래픽 탐지 시스템을 제안한다. 제안된 시스템은 학내 네트워크망을 통해 수집한 다양한 시나리오의 Skype 트래픽에 적용하여 정확성과 탐지율을 검증하였다.

Ⅰ. 서 론

최근 인터넷 사용자의 증가와 고속 네트워크의 보급으로 네트워크 트래픽이 급증하였다. 이것은 단 순히 WWW, FTP, SMTP, DNS 와 같은 전통적인 인터 넷 서비스뿐만 아니라 멀티미디어, P2P (peer-to-peer), P2P 기반으로 게임, 스마트 폰 응용 등의 다양한 서비스의 증가로 인한 부분도 큰 몫을 차지하기 때문이다. 이에 따른 트래픽이 급증함에 따라 효과적인 네트워크 관리를 위해 트래픽 모니터링 및 분석의 중요성이 커지고 있다.

Skype 의 트래픽들은 기본 적으로 암호화가 되어있고, 해당 응용 설치 시 동적 포트 번호가 할당되고, 일반적인 프로토콜을 사용하지 않아 Skype 의 트래픽을 정확하게 탐지하는 것은 일반적인 분석 방법론으로는 불가능하다. 하지만 Skype 응용 설치 때마다 달라지는 클라이언트 포트를 알아낼 수 있다면 대부분의 트래픽 을 쉽게 분류할 수 있으며 각 호스트의 응용 사용 시간 측정 및 제어 등 다양한 방면에서 활용이 가능하다.[1][2] 기존의 Skype 응용 트래픽 탐지 시스템은 트래픽 탐지율 측면에서는 만족할 만한 성능을 보여줬지만, 상대적으로 높은 Type 2 문제(False Positive)가 발생하는 문제점이 존재했다. 본 논문에서는 이러한 문제를 해결하고 보다 대용량의 다양한 시나리오의 Skype 응용 트래픽을 분류할 수 있는 개선된 Skype 응용 트래픽 분류 시스템을 제안한다.

이 논문은 2016 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 과학기술인문융합연구사업 (No.NRF-2016M3C1B6929228)과 2015 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2015R1D1A3A01018057).

Ⅱ. 본론

본 논문에서 제안하는 Skype 트래픽 탐지 그림 1 과 같다. 시스템은 두 단계로 구성되어 있다. 첫 단계 (Program1,2,3)는 Pre-Processing 단계이다. 두번째 단계는 Skype 응용 탐지 단계로서 크게 3 가지 모듈로 구성되어 있다. Login Detection 모듈은 Skype 응용을 사용하는 호스트의 IP(SC-IP)와 Skype 클라이언트에서 호스트에 할당하는 동적 포트(SC-Port)를 탐지하고 {SC-IP, SC-Port} 정보를 리스트로 구축한다.

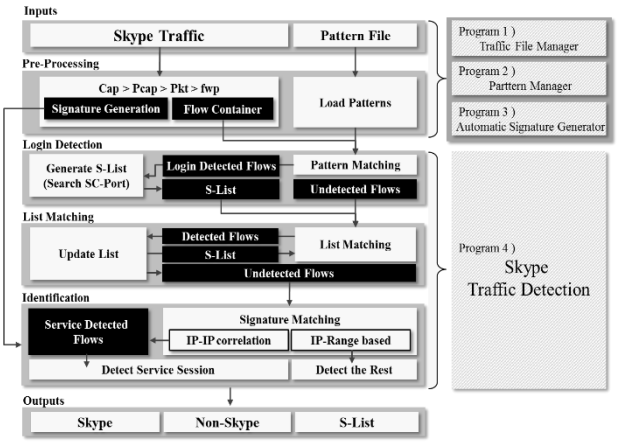

Figure 1. Sky-Scope 시스템 구조

이 과정에서는 패턴 기반 탐지 방법이 적용된다. List Matching 모듈에서는 {SC-IP,SC-Port} 리스트 정보를 바탕으로 Skype 응용 호스트를 탐색한다. Identification 모듈에서는 리스트를 통해 탐지된 호스트에서 사용하는

Skype 응용의 서비스 기능들을 분류하고 트래픽을 탐지한다. Identification 모듈에서는 리스트, 시그니쳐 기반의 탐지 방법이 적용되고 추가적으로 IP 간의 상관 관계를 이용한 correlation 방법과 IP 대역을 기반으로 한 IP range 방법을 적용 한다.

Skype 트래픽 중 기존의 시스템에서 가장 탐지하기 어려웠던 flow 는 TLS flow 이다. Skype 트래픽을 분석했을 때 대용량의 TLS flow 들이 발생하게 되는데 기존의 Skype 트래픽 탐지 시스템에서는 시그니쳐와 IP 대역을 기반으로 TLS flow 들을 탐지했다.

개선된 Skype 트래픽 탐지 시스템에서는 그림 2 와 같이 추가로 IP 간의 상관 관계를 이용한 방법을 적용한다. 그림 2 는 그림 1 의 Identification 모듈에서 적용되는 알고리즘으로 기존에 없던 리스트 기반의 IP 상관 관계 분석 방법을 추가해 이전의 시스템 보다 오탐을 줄이고 분석률을 높일 수 있다.

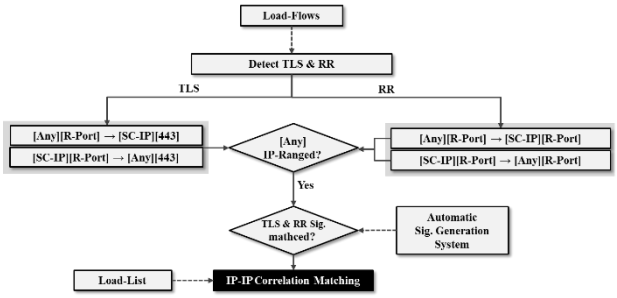

Figure 2. TLS flow 탐지 알고리즘

Ⅲ. 실험

기존의 Sky-Scope 시스템을 통해 Skype 응용 트래픽을 탐지했을 때 결과는 표 1 과 같다.

Detection Method	Flow Completeness	Pkt Completeness	Byte Completeness
1. List	39.6% (578/1,461)	26.7% (2,501/9,380)	6.6%
2. Signature & IP-IP Correlation	56.3% (824/1,461)	69.4% (6,514/9,380)	91.5%
3. Signature &IP-Range	2.6% (39/1,461)	3.6% (338/9,380)	1.7%
Total Completeness	98.6%	99.7%	99.8%

Table 1. Skype 응용 트래픽 적용 실험 결과

기존의 Sky-Scope 시스템의 문제점은 높은 분석률과 비교해 상대적으로 오탐 FP(False Positivie) 비율 또한 높았다는 점이다. 이러한 문제를 개선하기 위해 실제로 대부분의 트래픽을 탐지하는 Signature 와 IP-Range 기반의 탐지 방법을 개선해 오탐률을 기존의 10~15%에서 약 5%까지 낮추는 결과를 얻을 수 있었다. 개선된 Sky-Scope 의 실험 결과는 표 2 와 같다.

본 실험에서는 기존의 로그인, 서비스 트래픽으로 구분된 트래픽이 아닌 그림 3 와 같은 시나리오를 구성해 트래픽을 수집하고 실험을 진행했다.

개선된 Sky-Scope 시스템에 그림 3 의 과정을 통해 수집한 트래픽을 적용했을 탐지 결과 모든 트래픽 트레이스에서 95% 이상의 분석률을 보였다.

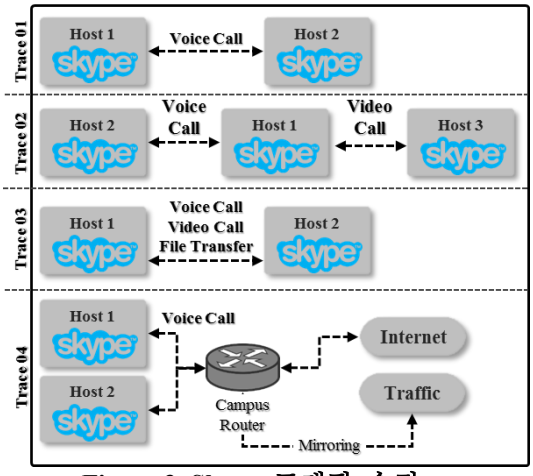

Figure 3. Skype 트래픽 수집

Trace #	Identified Skype Traffic Completeness		
	Flow	Pkt	Byte
Trace 01	100% (243/243)	100% (1,854/1,854)	100% (941K/941K)
Trace 02	100% (390/390)	100% (2,659/2,659)	100% (1,351K/1,351K)
Trace 03	99.5% (670/673)	99.8% (4,798/4,807)	99.9% (2,741K/2,742K)
Trace 04	96% (1,382/1,440)	98.7% (11,728/11,882)	99.8% (8,429K/8,446K)

Table 2. 개선된 Sky-Scope 시스템 탐지 결과

Ⅳ. 결론 및 향후 연구

본 논문에서는 기존의 Skype 응용 트래픽 탐지를 위한 Sky-Scope 시스템의 구조를 개선해 4 가지 시나리오의 트래픽에 적용해보고 그 타당성을 검증했다. 향후 연구로 실시간 트래픽 수집 시스템에 적용해 연속적으로 발생하는 대용량의 트래픽에서 Skype 응용 트래픽을 정확하게 탐지할 수 있는 방향으로 시스템을 개선 시킬 계획이다.

참 고 문 헌

[1] Sung-Ho Yoon, Kyu-Seok Shim, Su-Kang Lee, and Myung-Sup Kim , "Framework for Multi-Level Application Traffic Identification," Proc. of the Asia-Pacific Network Operations and Management Symposium (APNOMS) 2015, Busan, Korea, Aug. 19-21, 2015, pp.424-427.

[2] Z. Yuan, C. Du, X. Chen, D. Wang, Y. Xue,"SkyTracer: Towards fine-grained identification for Skype traffic via sequence signatures", Computing, Networking and Communications (ICNC) 2014 International Conference, Feb. 3-6. 2014.