SnorGen: 웹 기반 시그니쳐 자동 생성 시스템

(SnorGen: Web-based Automatic Signature Generation System)

구영훈, 심규석, 정우석, 김명섭

고려대학교 컴퓨터정보학과

{gyh0808, kusuk007, hary5832, tmskim}@korea.ac.kr

요 약

오늘날 네트워크 기능을 사용하는 응용이 급속도로 생성되고, 기존에 있는 응용들 또한 트래픽 발생 패턴이 변화하고 있다. 따라서 네트워크 관리자는 새로운 응용 및 기존 응용을 탐지 및 관리하는 것이 매우 어렵다. 본 논문은 네트워크 트래픽을 바탕으로 응용을 탐지 할 수 있는 시그니쳐를 사용자에게 신속하게 제공하기 위한 웹 기반 시그니쳐 자동 생성 시스템 (SnorGen)을 제안한다. SnorGen 은 단일 응용에서 발생시킨 패킷 데이터를 입력으로 받아 1 차적으로 단일 단어로 구성되는 Content 시그니쳐를 생성하고 2 차적으로 동일한 패킷에서 추출되는 Content 시그니쳐의 집합인 Packet 시그니쳐를 생성한다. 또한 SnorGen 은 사용자가 입력한 트래픽의 정보(Flow, Packet, Byte, Payload)를 제공함으로써 어떤 트래픽 플로우에서 시그니쳐가 생성되었는지 확인할 수 있다. 본 논문에서 제시하는 시스템의 목적은 인터넷이 연결된 어느 곳에서든지 시그니쳐 자동 생성 시스템을 이용할 수 있고, 사용자가 수집한 트래픽의 정보를 빠른 시간 안에 보여주는데 있다.

Keywords: Torrent, P2P, Traffic Classification, Network Management, Analysis

1. 서론[1]

본 논문에서는 네트워크 관리 목적을 효율적으로 충족시키기 위해 응용 시그니쳐 자동 생성 시스템을 개발하였다. 응용 시그니쳐는 응용에서 발생하는 트래픽의 고유한 특성이다. 이러한 시그니쳐를 이용하여 네트워크 관리자는 트래픽 분석을 실시한다[1,2]. 트래픽 분석은 트래픽을 발생시킨 응용을 판별하는 것으로써, 분석 결과는 네트워크 정책뿐만 아니라, 용량계획(capacity planning), 네트워크 권한 설정, 트래픽 엔지니어링, 고장 진단등과 같이 다양한 분야에서 활용될 수 있다.

[1] 이 논문은 2015 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2015R1D1A3A01018057).

헤더 시그니쳐는 응용 서버의 IP 및 포트를 이용하여 트래픽을 발생시킨 응용을 판별한다. 이러한 분석은 Internet Assigned Number Authority(IANA)에서 지정한 포트 정보를 이용한 트래픽 분석 방법이다. 포트 정보와 대응하는 서비스를 트래픽에서 분류할 수 있다. 적은 메모리 사용으로 매우 빠르게 분석 할 수 있는 장점이 있지만, 오늘날 네트워크를 사용하는 많은 응용들은 방화벽 및 IPS 장비를 통과하기 위해 포트 번호를 임의로 설정하여 트래픽을 발생시킬 뿐만 아니라 포트 사용자가 포트 번호를 임의로 설정할 수 있고, 매 실행 시 임의의 포트번호를 사용하기 때문에 더 이상 포트 번호가 특정 서비스, 프로토콜을 의미하지 않는다.

이러한 문제를 해결하기 위해 페이로드 기반 트래픽 분석 방법이 연구되었다. 페이로드 기반 분석방법은 패킷의 페이로드 내에서 응용마다 가지는 특정 문자열의 포함 유무를 통해 트래픽을 분석하는 방법이다. 현재의 네트워크 트래픽 분류 방법 중 트래픽의 페이로드를 직접 검사하기 때문에 가장 신뢰도 있는 분석 방법이다. 하지만 페이로드 기반 시그니쳐를 사용하기 위해서는 분류 대상 응용 트래픽의 특징을 파악하고 시그니쳐를 생성하는 전처리 과정이 필수적이다. 페이로드 기반 시그니쳐는 패킷의 페이로드 내용을 확인하여, 공통적인 문자열을 찾아 시그니쳐로 생성하여 특정 응용을 분류할 수 있다. 그러나 지속적으로 업데이트되고 개발되는 현대 사회의 응용 특징에 의해 시그니쳐의 생성 및 관리가 어렵고 최근 암호화된 트래픽이 증가하고 보안 기술이 발전하면서 네트워크 관리자가 직접 눈으로 공통 문자열을 찾기에는 많은 어려움이 존재한다.

따라서 본 논문에서 제안하는 웹 기반 시그니쳐 자동 생성 시스템은 네트워크 관리자뿐만 아니라 네트워크에 대한 이해가 없는 사용자도 쉬운 방법으로 시그니쳐를 생성하는 것이 가능하다. 본 시스템은 사용자가 본 시스템의 입력인 해당 응용의 트래픽을 수집하여, 웹 기반 시스템에 입력을 하게 되면 트래픽을 탐지할 수 있는 시그니쳐를 출력으로 한다. 본 시스템은 단어로 되어있는 Content 시그니쳐와 동일한 패킷에서 추출되는 Content 시그니쳐의 집합인 Packet 시그니쳐를 제공한다. 또한 사용자가 입력한 트래픽의 정보(Flow, Packet, Byte, Payload)를 확인할 수 있다. 본 시스템을 사용하면 인터넷이 연결된 어느 곳에서나 실시간으로 빠르게 시그니쳐를 생성할 수 있으므로 생성된 시그니쳐를 이용하여 해당 네트워크의 특정 응용에 대한 모니터링이 가능하다. 또한 생성된 시그니쳐를 관리 대상 네트워크의 방화벽 규칙으로 설정함으로써 잘 알려지지 않은 네트워크 공격에 빠른 대응이 가능하며 내부 네트워크의 사용자 트래픽을 제어함으로써 QoS 정책을 효율적으로 세울 수도 있다.

본 논문은 본 장 서론에 이어, 2 장에서 관련연구에 대해 언급하고, 3 장에서 SnorGen 페이로드 시그니쳐 자동 생성 방법에 대한 간략한 설명을 한다. 이어 4 장에서는 페이로드 시그니쳐 자동 생성 방법을 이용할 수 있는 웹 사이트에 대한 설명과 5 장에서 결론 및 향후 연구로써 본 논문을 마친다.

2. 관련 연구

페이로드 시그니쳐 자동 추출 방법론은 다양한 방법으로 연구되고 있다[4,9]. 기존 페이로드 시그니쳐 자동 추출 방법으로는 LCS(Longest Common String) 또는 Smith-Waterman 알고리즘을 이용하여 공통적인 스트링을 추출하는 연구가 진행되었다. LASER (LCS-based Application Signature ExtRaction)[6]은 LCS 알고리즘을 응용 트래픽 시그니쳐 추출 목적에 맞게 변형한 알고리즘이다. LASER 를 이용한 시그니쳐 추출방법은 두 개의 스트링을 비교하는 Matrix 에서 Backtracking 을 이

용하여 공통 문자열을 찾는 방법이다. 이 방법은 시그니쳐를 추출하는 시간이 오래 걸리고, 많은 계산과정을 포함하기 때문에 시스템 부하가 있을 수 있다는 단점이 존재한다. LCS 알고리즘과 마찬가지로 Smith-Waterman 알고리즘[5,8]도 시그니쳐 자동 생성 연구에서 많이 다루어지고 있다. Smith-Waterman 알고리즘은 본래 DNA 의 유사도를 판단하는 목적으로 발표된 알고리즘이다. 이 방법 또한 두 개의 스트링을 비교하여 Matrix 에 표시하고 다시 Backtracking 으로 최대, 최적의 시그니쳐를 찾아내는 작업을 한다. 위의 LCS 알고리즘과 비교했을 때 가장 큰 차이는 Backtracking 방법이다. LCS 방법보다는 시간 복잡도와 계산 복잡도가 적지만 이 방법도 시그니쳐 추출에 많은 시간과 계산 과정이 필요하다.

다른 형태의 시그니쳐 자동 추출 방법인 AutoSig[7]는 시그니쳐의 가능성이 있는 모든 공통 문자열을 추출하고 추출된 문자열을 구조화하여 시그니쳐를 생성하는 방법이다. 이 방법은 가능성 있는 모든 공통 문자열을 추출할 때, 너무 많은 문자열이 계산과정에 포함된다. 예를 들면, 20 개의 문자로 되어 있는 문자열에서 길이가 4 인 문자열을 추출한다면 16 개의 부분 문자열이 추출되고, 16 개의 문자열은 모두 계산과정에 포함된다. 따라서 추출된 부분 문자열의 개수가 많고 범위가 넓기 때문에 메모리 사용률 및 처리 시간의 단점을 가지고 있다.

기존 방법들은 특정 두 문자열을 비교하여 실제 트래픽에 적용하기에는 많은 한계가 존재한다. 적용할 트래픽의 순서를 정하거나 그룹화시키는 전처리 과정과 생성된 부분 문자열을 하나의 규칙으로 통합시키는 후처리 과정이 필요하다. 또한 생성하는 환경이 다르다면 생성되는 시그니쳐가 달라질 수 있기 때문에 신뢰도는 떨어진다. 따라서 본 논문의 웹 사이트에서 사용될 페이로드 시그니쳐 자동 생성 시스템은 공통 문자열을 순차 패턴 알고리즘[10,11]을 이용하여 자동으로 추출하는 방법을 사용한다. 또한 가장 상용화되어 있는 트래픽 분석 시스템인 Snort 의 룰 형태로 추출한다. 웹 사이트에서 사용될 페이로드 시그니쳐 자동 생성 시스템은 3 장에서 설명한다.

3. SnorGen 페이로드 시그니쳐 자동 생성 방법

본 장에서는 SnorGen 시스템에서 사용하는 순차 패턴 알고리즘을 이용한 페이로드 시그니쳐 자동 생성 방법을 설명한다. 본 논문에서 제안하는 SnorGen 방법에서 응용에서 발생하는 패킷 데이터 PCAP 형식으로 입력받는다. 패킷 데이터는 Netmonitor나 WireShark와 같은 트래픽 수집 도구를 이용하여 수집할 수 있다. SnorGen은 최소 2개 이상의 트래픽(cap, pcap, winpcap 등) 데이터를 입력 받아야만 동작한다. 입력 받은 트래픽에서 공통적으로 추출되는 단어 중 하나의 패킷 내에 존재하는 단어를 Content 시그니쳐라 명명하고, 동일한 패킷 내에서 존재하는 Content 시그니쳐를 조합하여 Packet 시그니쳐를 명명한다. 최종적으로 생성된 Content 시그니쳐와 Packet 시그니쳐를 Snort[3] content 규칙으로 변형한다. Content 시그니쳐를 사용할 시 생성되는 시그니쳐의 개수가 증가되어 False Negative를 줄일 수 있지만 네트워크 트래픽의 특성을 고려한 Packet 시그니쳐를 사용할 시 불필요한 중복을 제거하여 시그니쳐의 개수를 최소화함으로써 시스템의 부하를 감소시켜 효율성을 증대할 수 있다.

Snort content 규칙은 다양한 구성 요소를 표기할 수 있지만, 본 논문에서는 헤더 정보와 페이로드 정보만으로 구성되는 규칙을 대상으로 한다. 그림 1은 Content 시그니쳐를 이용한 Snort 규칙 예시로서, 의미는 프로토콜은 TCP, 목적지 포트 번호는 80을 사용하는 패킷 중 “mmmmX”이란 content가 페이로드 5번째 바이트와 20번째 바이트 사이에 위치하는 경우 알림을 울리라는 것이다.

```
Action Protocol SrcIP SrcPort → DstIP DstPort (Payload)
Alert tcp any any → any 80 (content:”mmmmX”; offset:5;depth:20;)
```

그림 1. Snort content 규칙과 예시

또한 본 방법은 단일 문자로 구성된 Content 시그니쳐 뿐만 아니라 트래픽 특성에 따라 패킷 단위로 이루어진 Packet 시그니쳐도 생성된다. Packet 시그니쳐는 하나의 패킷에서 동일하게 나오는 Content 시그니쳐의 집합이다. 그림 2는 Packet 시그니쳐를 이용한 Snort 규칙의 예이다.

Content Signature(in one packet)	Alert tcp any any → x.x.x.0/24 80 (content:"mmmmX"; offset:5; depth:20;)
	Alert tcp any any → x.x.x.0/24 80 (content:"vvvvC"; offset: 23; depth:40;)
Packet Signature	Alert tcp any any → x.x.x.0/24 80 (content:"mmmmX"; offset:5; depth:20; content:"vvvvC"; offset:23; depth:40;)

그림 2. Snort Packet 규칙과 예시

다음과 같이 시그니쳐를 자동 생성하기 위해 그림 3과 같은 과정을 수행한다. 최초에 호스트 별로 분석 대상 응용 및 서비스, 혹은 악성 코드가 발생한 트래픽을 수집한다. 수집된 패킷 집합 트래픽을 플로우로 구성하고 단일 플로우에서 전송방향이 같은 패킷들을 조합하여 하나의 Sequence를 구성한다. Sequence는 Content 시그니쳐를 추출하기 위한 순차 패턴 알고리즘의 입력으로 수집한 트래픽에서 모든 페이로드의 연속이다. 이를 Content Sequence라 명명하고 Content 시그니쳐를 추출하기 위한 입력으로 사용한다. 또한 추출된 모든 Content 중 동일한 패킷에서 추출되는 Content의 연속을 Packet Sequence라 명명하고 이를 Packet 시그니쳐를 추출하기 위한 입력으로 사용한다.

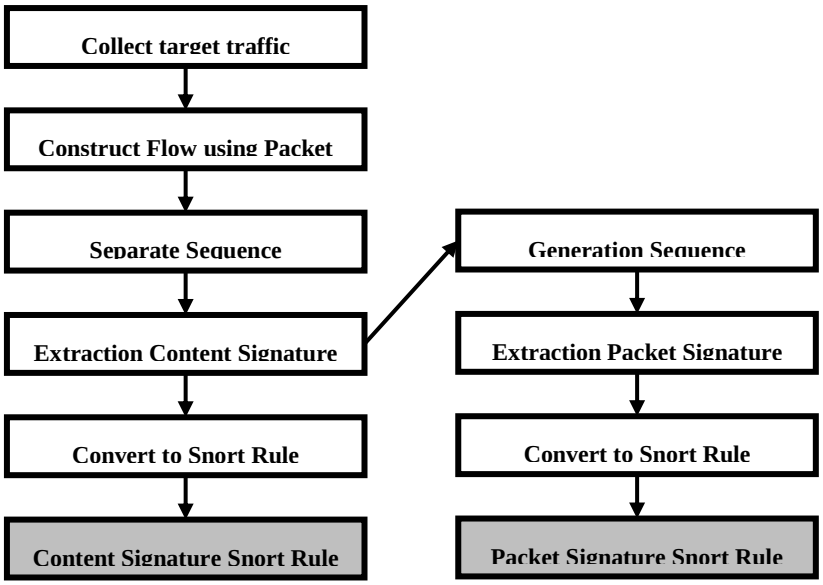

그림 3. 시그니쳐 자동 생성 프로그램 순서도

그림 4와 같이 순차 패턴 알고리즘은 입력 받은 Sequence에서 길이가 1인 후보 Content를 시작으로 길이를 증가시키면서 후보 Content를 찾고 최종적으로 일정 수준이상의 지지도를 가지는 Content를 추출한다. 단순히 Content만을 규칙으로 사용할 경우 오탐(탐지 대상이 아닌 트래픽을 탐지)의 가능성이 높기 때문에 추가적인 정보를 분석하여 Snort 규칙에 기술한다. 페이로드 시그니쳐 자동 생성 프로그램에서 사용한 추가 정보는 IP 주소와 포트 번호와 같은 헤더 정보와 Content의 위치 정보이다. 추출된 Content를 입력 트래픽에 적용하여 해당 Content가 매칭되는 트래픽을 그룹화하고 해당 그룹의 공통된 헤더 정보와 위치정보를 분석한다. 최종적으로 생성된 Snort Content 규칙은 Snort 엔진이 탑재된 네트워크 장비에 적용할 수 있다.

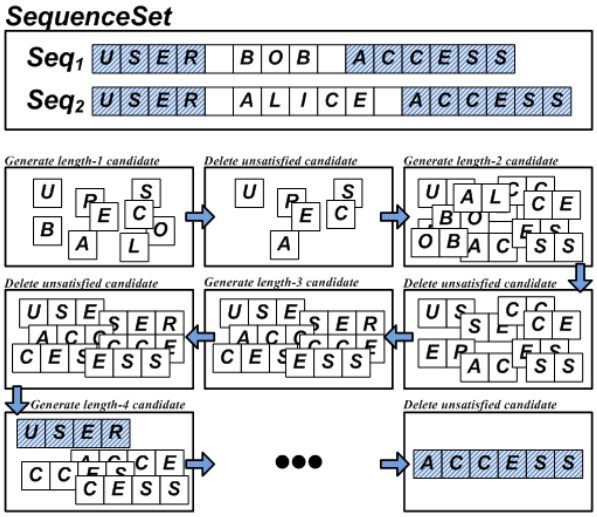

그림 4. 순차 패턴 알고리즘을 이용한 Content 시그니쳐 생성 과정

Packet Sequence 또한 순차 패턴 알고리즘에 Content Sequence와 동일한 방법으로 입력된다. Content 시그니쳐를 생성할 시 길이 1의 의미가 문자 1개라면, Packet 시그니쳐를 생성할 때 길이 1의 의미는 Content 1개를 의미한다. 동일한 방법으로 Packet 시그니쳐를 생성한다. Packet 시그니쳐는 Content 시그니쳐보다 추출되는 개수가 적고, 하나의 패킷에서 여러 개의 단어가 매칭되어야 하기 때문에 오탐률이 적어 정확성이 향상된다. 그러나 분석률이 약간 감소될 수 있다.

4. 웹 기반 시그니쳐 자동 생성 시스템

순차 패턴 알고리즘을 이용한 시그니쳐 자동 생성 시스템을 이용하기 위해서는 다음 URL (http://snorgen.korea.ac.kr) 로 접속할 수 있다. 본 웹 사이트에서는 단어로 구성되어 있는 Content 시그니쳐와 동일한 패킷에서 추출되는 Content 시그니쳐의 집합인 Packet 시그니쳐를 확인할 수 있을 뿐만 아니라 각 시그니쳐에 대한 분석률을 확인할 수 있다. 분석률은 각 시그니쳐를 사용하여 최초에 입력된 트래픽을 분석했을 때 전체 Flow, Packet, Byte의 양에서 분석할 수 있는 Flow, Packet, Byte의 양의 비율이다. 또한, 사용자가 입력한 트래픽의 정보(Flow, Packet, Byte, Payload)를 확인할 수 있다. 그림 5는 본 웹사이트의 개념도이다. 네트워크 관리자는 특정 응용에 대한 트래픽을 수집하거나 관리 대상 네트워크의 트래픽을 수집하여 SnorGen 페이지의 시그니쳐 자동 생성 시스템으로 입력한다. 앞 장의 SnorGen 방법의 순서도를 거쳐 시그니쳐 리스트를 생성하고 시그니쳐들의 분석률을 검증한다. 생성된 시그니쳐를 이용하여 새로운 응용 트래픽 탐지 및 모니터링, 악성 트래픽 차단을 위한 방화벽 규칙 생성, 트래픽 제어를 위한 QoS 장비 규칙 생성을 통해 네트워크를 효율적으로 관리할 수 있다.

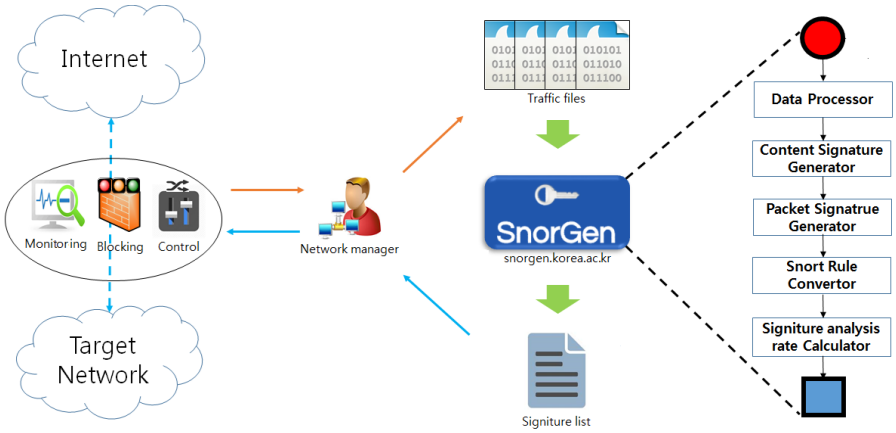

그림 5. 웹 기반 시그니쳐 자동 생성 시스템 개념도

SnorGen 웹 사이트에서는 4가지 메뉴를 구현하였다. Home은 SnorGen 페이지의 메인페이지로 시그니쳐 자동 생성 시스템의 개념도를 포함하였으며 그림 6과 같다. 다음은 HOW TO RUN으로 다음 SnorGen 페이지에서 시그니쳐 자동 생성 시스템의 매뉴얼 페이지이다. RUN에서 시그니쳐 자동 생성 시스템을 사용할 수 있고, SAMPLE 페이지에서는 대표적 응용 프로그램의 시그니쳐를 게시하여, 사용자가 해당 응용 트래픽의 시그니쳐를 추출한 후 비교하여, 더 정확한 시그니쳐를 선택할 수 있도록 제작하였다. 부가적으로 페이지 하단에 Forum 으로 본 SnorGen 웹 사이트를 사용하는 사용자의 의견을 수렴할 수 있도록 하였다.

그림 7 는 SnorGen 웹 사이트의 RUN 페이지 중 상단 페이지이다. 본 페이지에서는 시그니쳐 자동 생성 시스템을 이용할 수 있다. 그림 7 의 ①에서는 사용자가 수집한 트래픽을 Drag & Drop 방식 또는 박스를 클릭하여 파일을 선택 후 업로드 할 수 있다. 이때, 수집한 트래픽 파일은 최소 2 개 이상이어야한다. 시그니쳐 자동 생성을 위해 순차 패턴 알고리즘을 사용하는데 이때, 최소 2 개 이상의 파일에서 공통 문자열을 추출할 수 있기 때문이다. 그림 7 의 ②는 시그니쳐 자동 생성 시스템을 실행하는 서버의 CPU 상태와 추출 과정을 가시적으로 보여준다. 여러 사용자가 사용할 수 있는 가능성이 있기 때문에 서버의 CPU 상태를 보여줌으로써 서버 부하를 예방할 수 있다. 그림 7 의 ③은 사용자가 입력한 트래픽과 생성된 시그니쳐의 통계적인 정보를 보여준다. 표 1 은 그림 7 의 ③에서 트래픽 입력시 보여주는 정보이다.

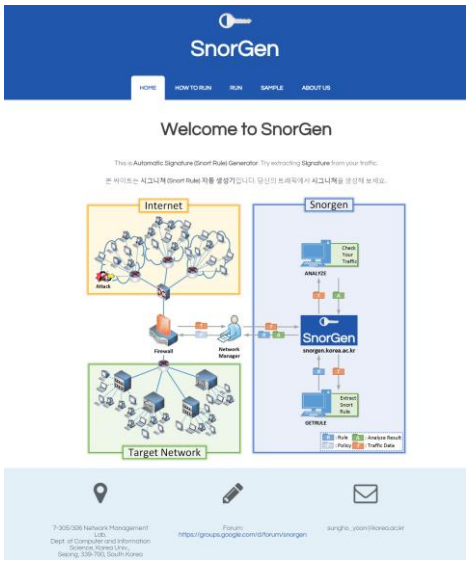

그림 6. SnorGen 웹 사이트 메인 페이지

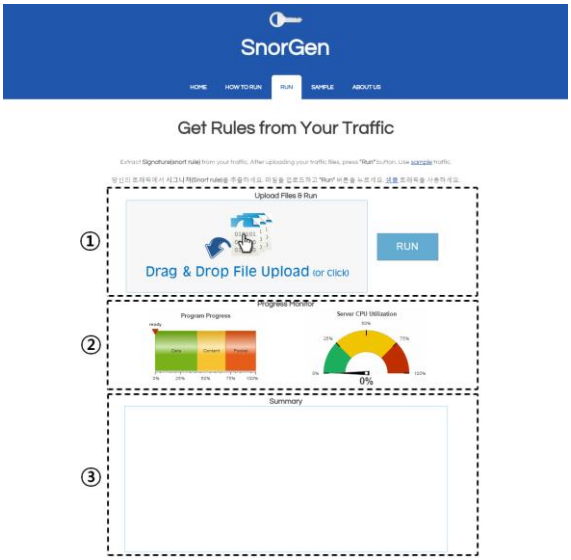

그림 7. SnorGen 웹 사이트 시그니쳐 자동생성 페이지

표 1. SUMMARY(그림 7 의 ③)에서 보여주는 통계적인 정보

종류	예시	설명
File Information	#1 file : test1.pcap – 8809 KB #2 file : test2.pcap – 9919 KB	입력된 트래픽 파일의 크기
Traffic Information	#1 file : test1.pcap – flow: 167 pkt: 11973 byte: 12817162 #2 file : test2.pcap – flow: 216 pkt: 12919 byte: 12017814	입력된 트래픽 파일에 대한 Flow, Packet, Byte 양

Rule Information	Content Signature: 84 Completeness : 100 (383/383) 100 (24892/24892) 100 (24,252KB/24,252KB) Content Process Time : 118.60s Packet Signature: 50 Completeness : 99.74 (382/383) 99.95 (24880/24892) 99.98 (24,248KB/24,252KB) Packet Process Time : 0.42s Total Process Time : 120.59s	추출된 Content 시그니쳐의 개수 추출된 Content 시그니쳐로 입력된 트래픽의 분석률 추출에 사용된 시간, 추출된 Packet 시그니쳐의 개수, 추출된 Packet 시그니쳐로 입력된 트래픽의 분석률 추출에 사용된 시간 총 Process Time

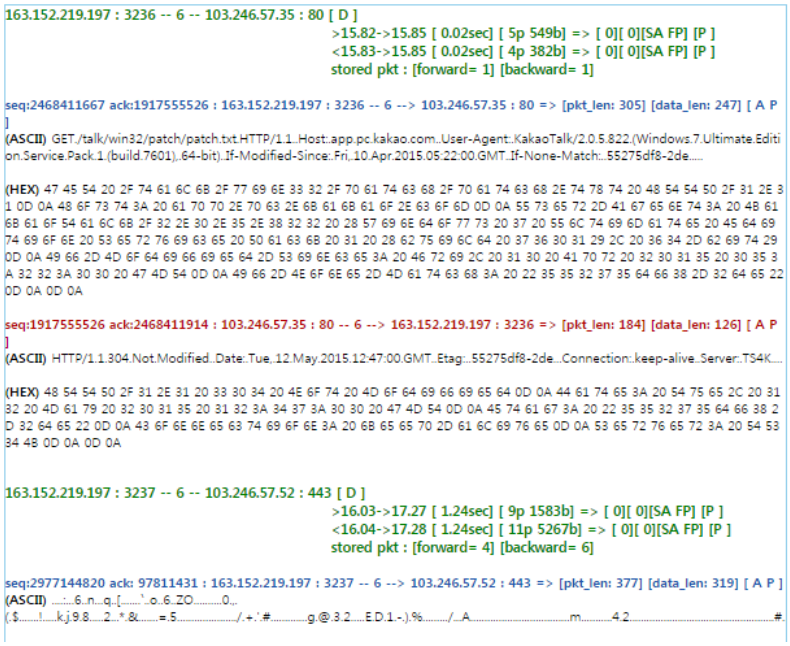

그림 8. SnorGen 에서 확인 가능한 트래픽 정보

또한 각 트래픽의 정보를 확인할 수 있다. 트래픽 정보란에 See_Detail 버튼을 생성하여 트래픽의 헤더 정보 및 페이로드 정보를 Session 단위로 확인할 수 있다. 그림 8 은 See_Detail 의 내용이다. 그림 8 에서 초록 문자는 트래픽 파일내의 하나의 Session 헤더 정보를 의미한다. 또한 해당 Session 의 페이로드는 ASCII 와 16 진수로된 두 가지 표현을 모두 표기한다. HTTP 프로토콜의 경우 ASCII 로 표현된 페이로드 내용만 필요하지만, 암호화된 내용 또는 문자로는 식별 불가능한 정보는 16 진수로 확인하기 때문이다. 본 RUN 페이지에서는 사용자가 입력한 트래픽을 바탕으로 자동으로 생성된 Content 시그니쳐와 Packet 시그니쳐를 확인할 수 있다. 표 2 는 추출된 시그니쳐에서 제공하는 정보에 대한 설명이다. 추출된 각 시그니쳐는 See_Detail

버튼을 이용하여 추출된 시그니쳐가 위치한 곳을 트래픽 정보에서 확인하는 것이 가능하다. 그림 9 은 추출된 시그니쳐에서 확인할 수 있는 See_Detail 정보이다. 또한 그림 10 은 사용자가 입력한 트래픽을 기반으로 추출된 Content 시그니쳐와 Packet 시그니쳐이다.

표 2. 추출된 시그니쳐에서 제공하는 정보

종류	예시	설명
Support	Support: 2/2 files	추출된 해당 시그니쳐가 2 개의 트래픽 파일 중 2 개의 파일에 포함
Fixed Offset		추출된 해당 시그니쳐가 항상 같은 위치에서 추출됨
F-Com	F-Com: 3.92(15/383) 24.20(6023/24892) 38.72(9615995/24834976)	추출된 해당 시그니쳐로 입력된 트래픽을 분석할 수 있는 Flow, Packet, Byte 양
Snort Rule	alert tcp 1.201.0.62 443 -> any any (sid: 1115414; content:"*.talk.kakao.com"; offset:333; depth:16;)	Snort 규칙 형태로 변경된 시그니쳐
See_Detail		추출된 해당 시그니쳐를 트래픽에서 확인 가능한 페이지

그림 9. 해당 시그니쳐를 트래픽에서 확인 가능한 See_Detail 페이지

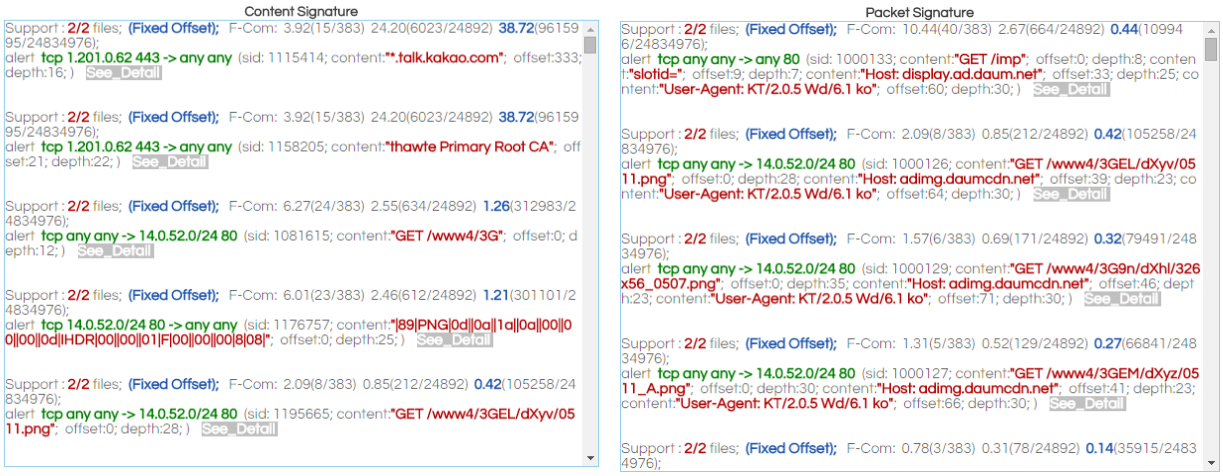

그림 10. 추출된 Content 시그니쳐와 Packet 시그니쳐

마지막으로 SnorGen 웹 사이트에서는 Sample 페이지를 제공한다. 본 Sample 페이지에서는 국내 가장 대표적인 응용과 정보보안취약점 표준(CVE), 웹 응용 15 가지를 선정하여 SnorGen 에서 추출된 시그니쳐를 게시한다. 따라서 사용자들은 SnorGen 을 이용하여 시그니쳐를 추출하고, 추출된 시그니쳐의 정확성을 파악하기 위해 Sample 페이지에서 비교 확인 가능하다.

5. 결론 및 향후 연구

본 SnorGen 웹 사이트는 사용자가 인터넷 연결이 되어 있는 모든 지역에서 네트워크 트래픽 시그니쳐 자동 생성 시스템을 손쉽게 사용할 수 있을 뿐만 아니라 사용자의 트래픽 정보를 확인하는 것이 가능하다. 또한, 해당 시스템은 순차 패턴 알고리즘을 이용하여 시그니쳐를 생성하기 때문에 사용자는 신속하게 시그니쳐를 제공 받을 수 있다. 본 SnorGen 웹 사이트는 메인 페이지, 사용자 메뉴얼 제공 페이지, 시스템 사용 페이지, Sample 페이지로 이루어져 있어서 사용자가 쉽게 사용할 수 있게 디자인 되었다.

향후 본 SnorGen 웹 사이트에서는 시그니쳐 자동 생성 시스템의 옵션을 조절할 수 있도록 제공할 계획이다. 옵션은 Minimum Support, Content 시그니쳐의 최소 길이, 하나의 플로우에서 사용할 패킷의 개수, 최대 Sequence 길이 등을 조절할 수 있도록 시스템을 개선할 계획이다. 또한, 제안하는 시그니쳐 자동 생성 시스템의 검증을 위하여 성능을 평가할 수 있는 다양한 방법들을 연구하고 이를 적용시켜 평가 후 개선점을 찾을 계획이다.

6. 참고 문헌

[1]Y. Wang, Y. Xiang, W. L. Zhou, and S. Z. Yu, “Generating regular expression signatures for network traffic classification in trusted network management,” J. Netw. Comput. Appl., vol. 35, pp. 992-1000, May 2012.

[2]B. Park, Y. Won, J. Chung, M. S. Kim, and J. W. K. Hong, “Fine-grained traffic classification based on functional separation,” Int. J. Netw. Management, vol. 23, pp. 350- 381, Sept. 2013.

[3]snort. Available: https://www.snort.org/

[4]H.-A. Kim and B. Karp, “Autograph: Toward automated, distributed worm signature detection,” in USENIX Security Symp., vol. 286, 2004.

[5]J. Newsome, B. Karp, and D. Song, “Polygraph: Automatically generating signatures for polymorphic worms,” IEEE Symp. Security and Privacy, pp. 226-241, 2005.

[6]B.-C. Park, Y. J. Won, M.-S. Kim, and J. W. Hong, “Towards automated application signature generation for traffic identification,” IEEE Network Operations and Management Symp. (NOMS 2008), pp. 160-167, 2008.

[7]M. Ye, K. Xu, J. Wu, and H. Po, “Autosig-automatically generating signatures for applications,” IEEE Int. Conf. Computer and Inf. Technol.(CIT'09), pp. 104-109, 2009.

[8]X. Feng, X. Huang, X. Tian, and Y. Ma, “Automatic traffic signature extraction based on Smith-waterman algorithm for traffic classification,” IEEE Int. Conf. Broadband Netw. Multimedia Technol. (IC-BNMT), pp. 154-158, 2010.

[9]C. MU, X.-h. HUANG, X. TIAN, Y. MA, and J.-l. Qi, “Automatic traffic signature extraction based on fixed bit offset algorithm for traffic classification,” The J. China Universities of Posts and Telecommun., vol. 18, pp. 79-85, 2011.

[10]R. Agrawal and R. Srikant, “Fast algorithms for mining association rules,” in Proc. 20th Int. Conf. VLDB, pp. 487-499, 1994.

[11]R. Agrawal and R. Srikant, “Mining sequential patterns,” in Proc. Eleventh Int. Conf. Data Eng., pp. 3-14, 1995.

[12] C. S. Park, J. S. Park, and M. S. Kim, “Automatic payload signature generation system,” J. KICS, vol. 38B, no. 08, pp. 615-622, Aug. 2013.

[13] S. H. Yoon, J. S. Park, H. M. An, and M. S. Kim, “Traffic behavior signature extraction using sequence pattern algorithm,” in Proc. KICS Int. Conf. Commun. (KICS ICC 2014), pp. 996-997, Jeju Island, Korea, Jun. 2014.

구 영 훈

2016 고려대학교 컴퓨터정보학과 졸업

2016~현재 고려대학교 컴퓨터정보학과 석사과정

<관심분야> 네트워크 관리 및 보안, 트래픽 모니터링 및 분석, 트래픽 분류

심 규 석

2014 고려대학교 컴퓨터정보학과 졸업

2014~현재 고려대학교 컴퓨터정보학과 석사과정

<관심분야> 네트워크 관리 및 보안, 트래픽 모니터링 및 분석

정 우 석
2015 고려대학교 컴퓨터정보학과 졸업
2015년~현재 고려대학교 컴퓨터 정보학과 석사과정
<관심분야> 네트워크 관리 및 보안, 트래픽 모니터링 및 분석

김 명 섭
1998 포항공과대학교 전자 계산학과 졸업
2000 포항공과대학교 컴퓨터 공학과 석사
2004 포항공과대학교 컴퓨터 공학과 박사
2006 Post-Doc. Dept. of ECE, Univ. of Toronto, Canada
2006년~현재 고려대학교 컴퓨터정보학과 부교수
<관심분야> 네트워크 관리 및 보안, 트래픽 모니터 링 및 분석, 멀티미디어 네트워크