Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.24, NO.3, Jun. 2014 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2014.24.3.491

트래픽 분석을 통한 악성코드 감염PC 및 APT 공격탐지 방안*

손 경 호,[1†] 이 태 진,[1] 원 동 호[2‡]
[1]한국인터넷진흥원, [2]성균관대학교

Design for Zombie PCs and APT Attack Detection based on traffic analysis*

Kyungho Son,[1†] Taijin Lee,[1] Dongho Won[2‡]
[1]Korea Internet & Security Agency, [2]Sungkyunkwan University

요 약

최근, 지능화된 공격기법을 통한 사이버테러가 지속적으로 발생하고 있으며 특히, 알려지지 않은 신종 악성코드를 사용하기에 탐지 및 대응이 매우 어렵다. 본 논문에서는 대용량 데이터 분석을 통해, 악성코드 침투단계 이후에, 좀비PC와 공격자와 통신을 사전탐지, 대응하는 알고리즘 개발 및 상용환경에서 검증하였다. 향후, 알고리즘의 고도화, 대용량 데이터 처리기술 적용을 통해, APT 공격의 탐지성능이 향상될 것으로 예상한다.

ABSTRACT

Recently, cyber terror has been occurred frequently based on advanced persistent threat(APT) and it is very difficult to detect these attacks because of new malwares which cannot be detected by anti-virus softwares. This paper proposes and verifies the algorithms to detect the advanced persistent threat previously through real-time network monitoring and combinatorial analysis of big data log. In the future, APT attacks can be detected more easily by enhancing these algorithms and adapting big data platform.

Keywords: Advanced Persistent Threat, APT, big data

I. 서 론

최근 들어, 사이버 공격은 기존 해커에 의한 호기심·자기과시 목적에서 벗어나 국가나 단체에 의한 경제적 이익과 사회혼란을 목적으로 진화하고 있다. 최근 발생한 '11년 3.4 DDoS공격,'13년 3.20 사이버테러는 공격자가 특정대상을 겨냥해 명확한 목표를 두고 지능적, 지속적으로 은밀히 공격을 가하여 정보를 수집하고 유출하는 방법인 지능형 지속위협(APT, Advanced Persistent Threat) 공격을 통해 이루어졌다. APT공격은 알려지지 않은 다양한 공격기법을 활용하여 기존이 보안솔루션(예, Anti-Virus제품 등)을 우회하고, 장기간에 걸쳐 취약점을 찾는 등을 통해 이뤄지는 공격으로 기존의 보안기술 및 제품으로 이를 찾아내고 대응하는 것은 어려운 실정이다. Fig. 1.은 일반적인 APT 공격 프로세스를 나타낸다.

〔1단계〕 APT 공격은 공격자가 내부 사용자가 방문하는 웹사이트 또는 이메일을 통해, 악성코드를 감염시켜 내부에 감염PC를 만들고, 〔2단계〕 조직내부 네트워크 등 인프라 정보를 수집하고, 〔3단계〕 계정정보

접수일(2013년 9월 10일), 수정일(1차: 2014년 4월 9일, 2차: 2014년 6월 12일), 게재확정일(2014년 6월 13일)
* 본 연구는 산업통상자원부 및 정보통신기술진흥센터의 산업융합원천기술개발사업(정보통신)의 일환으로 수행하였음. 〔10044938, 악성코드 프로파일링 및 대용량 보안이벤트 분석을 통한 공격징후 탐지기술 개발〕
† 주저자, khson@kisa.or.kr
‡ 교신저자, dhwon@security.re.kr(Corresponding author)

탈취, 악성코드 감염등을 통해 취약한 시스템·서버 등에 침입 한 후, [4단계] 공격자의 명령을 받아, 내부 정보유출 및 시스템 파괴 등의 단계로 이뤄지고 있다.

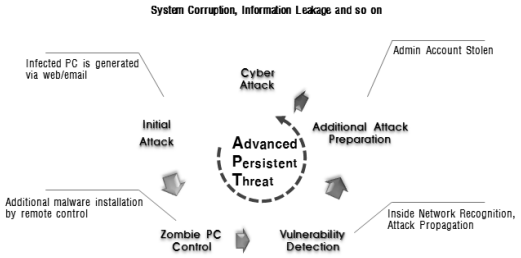

Fig. 1. APT Attack Process

'13년 3월 20일 발생한 해킹사고도 최초 해커가 피해회사의 취약한 서버 또는 직원 PC를 장악한 후, 백신업데이트 서버 등에 S/W 업데이트 파일로 위장한 악성코드(하드디스크 파괴형) 설치하고, 백신업데이트 서버는 내부 PC에 악성코드를 유포하였고, 감염된 내부 PC에서 특정 시각 이후에 하드디스크 손상을 발생시킨 전형적인 APT공격의 양상을 보였다. 이런 APT공격의 공통적인 특성으로 ① HTTP 백도어 통신, ② 시스템 내부에서의 감염 확산, ③ 동시 업데이트(예: P2P), ④ SSL통신 및 USB 등을 통한 정보수집 및 전달 형태로 APT 공격은 "탐지하지 어려운, 조용한 공격"으로 알려져 있다. 이러한 공통적인 형태에 대해 시스템 측면에서 대책을 수립하기 위해서는 먼저 조직 내의 시스템에 대한 공격의 구체적인 흐름을 분석해야 한다. 또한 이 공격의 흐름이 시스템의 네트워크와 어떻게 연관되어 있는지 이해하여야 시스템 설계자가 APT 공격의 영향을 분석하고 보안대책을 고려하여 시스템을 설계할 수 있다.

기술적으로는 최초 악성코드 침투단계에서는 악성URL 탐지기술, 악성URL에서 수집된 악성코드 및 이메일 첨부파일 대상 악성코드 분석기술, 악성코드 유입이 탐지될 경우, 관련된 내부 감염PC 탐지기술, 악성코드 분석을 통한 C&C(Command & Control;공격명령제어서버) 탐지 및 C&C로 접속하는 내부 감염PC 추가 탐지/조치 기술 등 조직 내부로 유입되는 전체널 대상 악성코드 탐지가 필요하다. 또한, 내부에 악성코드 감염PC를 확보한 단계에서는, 내부망 구성도 파악, 주요 시스템 접속계정 습득, 추가 악성코드 설치 등을 위해 수일에서 수개월 이상 탐색 실시함으로 조직 내 알려지지 않은 취약점/악성코드 등에 대한 탐지를 위해 대용량 보안 이벤트 연관분석을 통해 감염PC에서의 이상징후 탐지/대응 필요하다.

이에 본 논문에서는 보안 이벤트 연관분석 기술의 장단점을 비교하고, APT공격을 효과적으로 탐지하기 위한 알고리즘을 제안하고, 제안된 알고리듬을 실제 네트웍에 대해서 그 효과를 검증한다. 본 논문의 구성은 다음과 같다. 2장에서는 보안관제 측면에서의 보안로그 연관분석 연구동향 및 장단점을 분석하고 3장에서는 본 논문이 제안하는 APT 이상공격 징후 탐지 모델을 설명한다. 그리고, 4장에서는 실제 시스템 구현을 통해 제안한 알고리즘 검증하고, 마지막으로 5장에서 결론을 맺는다.

II. 관련 연구

보안로그 연관분석을 통한 공격탐지 연구는 이기종 보안시스템 및 IT시스템에서 발생하는 로그들을 분석하기 위한 방법으로 로그간의 연관성(Correlation)을 분석함을 의미한다. 이런 연관분석 방법에는 보안로그들의 유사속성을 비교한 유사도 기반 접근방법[1,2], 공격이 성공하기 위해 사전에 필요한 이벤트를 미리 설정해 탐지하는 시나리오 기반 접근방법[3,4,5], 유사도에 근거한 방법보다는 상위 응용 수준이며, 알려진 시나리오 기반 방법보다는 하위 분석 수준에서 동작하는 개별공격의 선행조건과 결과에 근거한 방법[6,7,8], 시간 흐름을 기반으로 공격시나리오를 분석하는 통계적 원인분석 방법[9,10], 개별적 단일 보안경보 간의 연관성을 가지고 침해시도여부를 탐지하는 경보 클러스터링 방법[11,12] 등 다양한 방법들이 제시되고 있다. Table 1은 보안로그 연관분석을 통한 공격탐지 연구의 특징을 비교하였다.

Table 1. Comparison of Related Research

Method	Pros	Cons
Similarity based	duplicate alarm reduction from security systems	multi-level attack detection difficulty
Scenario based	duplicate alarm reduction from security systems	high false positive and unknown attack detection
Attack based	multi-level attack detection	high false positive due to automatic rule
Statistics based	do not need pre-defined knowledge	complete correlation analysis difficulty

또한, 기존 방식들은 실시간 탐지를 목표로 하는 것이 대부분이며, 탐지시간이 3분/10분/1시간 수준의 매우 작은 윈도우를 기반으로 한다. 이벤트 빈도가 낮은 APT 공격에는 빈도를 이용한 분석으로 처리하기에는 한계점이 존재한다. 효과적인 APT공격 탐지를 위해서는 아래 같은 요구사항을 만족시켜야 한다.

- 대용량 보안로그를 분석하기 위한 로그 병합이 수행되어야 함
- 최대한 로그 간의 연관성을 부여할 수 있는 요소를 식별하여야 함
- 다단계 공격을 식별 및 탐지 가능해야 함
- 잘못된 공격(Nonrelevant Positive)을 감소시킬 수 있어야 함
- 대용량 보안로그를 처리하기 위한 하부구조가 구성되어야 함

이런 요구사항을 해결하기 위해, 데이터마이닝 기법을 이용한 방법[13,14]이나 다단계 공격을 탐지하기 위한 방법[15,16,17]들이 제안되고 있다. 또한, 상용제품들의 보안이벤트 연관성 분석 들에 대한 비교[18] 연구도 진행되었다. 기존 연구들에서 이미 로그를 교차·연관 분석함으로써 탐지율을 높이고, 오탐 또는 과도한 경보의 비율을 낮춘다고 판단하고 있다는 점과 기존의 대부분의 방법론에서 대상으로 삼은 데이터 양이 빅데이터 영역에 속함에도 불구하고, 실제적인 실험 또는 분석시스템의 규모가 매우 작음을 알수 있다. 이는 대용량 보안로그를 처리하기 위한 실질적인 인프라 또는 플랫폼이 제대로 갖춰지기 이전에 연구들로 본 연구에서는 빅데이터 플랫폼을 구축해 로그의 교차·상관 분석을 검증함으로써 APT 공격을 좀 더 효율적으로 탐지하기 위한 방안을 제시하고자 한다.

III. 이상 공격징후 탐지 방안

3.1 침해사고 발생 모델

본 논문에서 제안하는 이상 공격징후 탐지를 위한 침해공격 모델을 기술한다. 사이버 침해공격은 웹이나 이메일 등을 통해 내부PC를 악성코드에 감염시키는 1단계와 악성코드 감염PC에서 나타나는 모든 행위를 모니터링하여 내부망 구성, 주요 시스템에 대한 접속계정, 그 외 감염PC에서 쓸만한 정보들을 수집하는 2단계로 구성된다. 2단계 준비가 완료되면, 특정 시점에 개인정보 유출, 서버파괴 등의 목적을 가진 악성코드를 추가 설치하여 원하는 목적을 달성한다. 본 논문에서는 2단계에서 이상징후를 탐지하고, 감염PC를 조치함으로써 침해사고를 사전에 예방하는 것을 목표로 한다. Fig. 2.은 본 논문에서 다루고 있는 침해공격 탐지모델을 나타낸다.

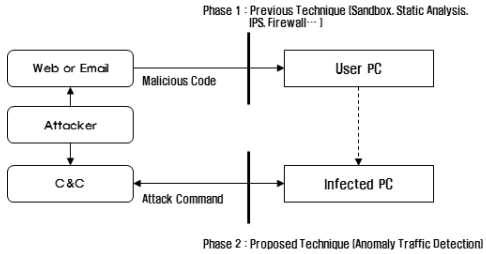

Fig. 2. Cyber Attack Model

2단계에 진입한 공격자는 악성코드 감염PC에서의 키보드 입력내용, 주요 시스템 접속정보, PC내 중요문서 등의 정보를 수집하고, 공격전략 수립이 완료되면 의도한 시점에 대량의 개인정보 탈취, 시스템 파괴, DDoS 공격 등을 수행한다. 즉, 감염PC를 통해 공격에 필요한 여러 정보를 수집해야 하는데, 이 과정은 통상 많은 시간이 소요된다. 지난 3.20 전산망마비 공격, 농협APT 공격의 경우 7개월로 추정되며, '10년 발생한 해외 eBay 해킹사고는 약 2년이 소요되었다. 따라서, 2단계에서 공격징후 탐지는 시간적 성능 요구사항은 민감하지 않다.

3.2 시스템 구성

본 논문에서 제안한 트래픽 기반 이상징후 탐지 시스템 구성은 다음과 같다. 우선, TAP 장비를 통해, 특정 조직의 내외부로 연결된 트래픽을 모두 수집할 수 있어야 하며, 수집된 트래픽 정보는 전처리를 거쳐, 원본데이터는 빅데이터 플랫폼인 HDFS (Hadoop File System)에 저장한다. 이와 동시에, 이상징후 탐지로직에서 사용할 데이터는 별도의 전처리를 거쳐 NoSQL에 저장한다. NoSQL에 저장된 데이터는 이상징후 탐지로직에 따른 결과를 효과적으로 분석하기 위해 개발된 Scheme으로 Batch 작업을 통해, 이상징후 관련 정보를 조회·분석을 수행하고, 이상징후로 판단될 때 원본 데이터가 저장된 HDFS를 통해 검증을 수행한다. 악성코드 감염PC로

판단될 경우, 해당 IP를 조치함으로써 침해사고가 실제 발생하기 전에 예방할 수 있다. Fig. 3.은 논문에서 제안하는 시스템 구성을 나타낸다.

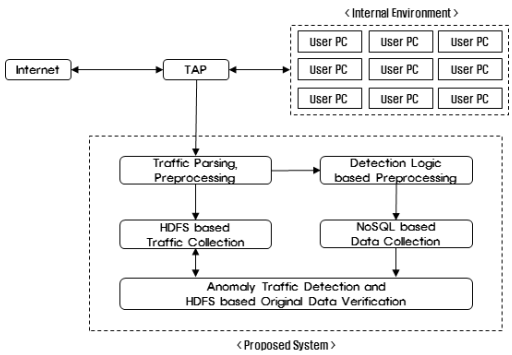

Fig. 3. System Architecture

개발된 시스템이 적용된 상용환경 A업체의 경우, 직원수가 45명 정도 수준(PC, 스마트폰 포함 약 100대)이며, 일평균 Outbound 패킷수는 14만~37만, Inbound 패킷수는 8~25만 정도이며, Outbound 트래픽에서 볼 때, Destination IP의 수는 일평균 3만~8만 정도 수준이다. 이러한 트래픽 환경에서 분석서버는 5-Tuples, HTTP Request URI, HTTP Response Filetype, 세션당 PPS, BPS 등 Anomaly 탐지로직에 필요한 데이터를 추출한다.

3.3 이상공격 탐지 알고리즘

3.3.1 주기적 접속 기반 이상징후 탐지

공격자는 악성코드에 감염된 PC를 통해 내부망 구성, 주요 시스템 접속정보 등을 탈취한다. 이러한 정보들은 감염PC와 공격자가 운영하는 C&C서버와 통신하면서 얻게 되므로, 일반적인 경우 주기적인 통신이 발생하게 된다. Fig. 4.는 주기적 접속 탐지를 위한 기본 데이터 생성 알고리즘을 나타낸다.

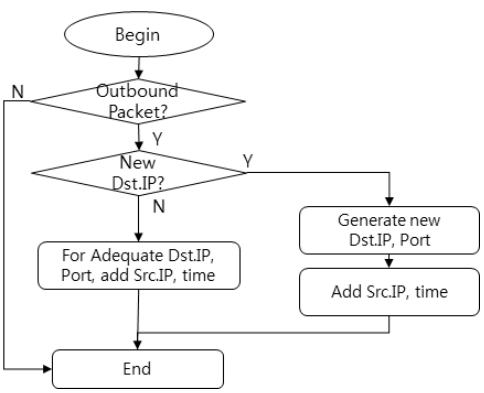

Fig. 4. Data Extraction for Repetive Connection

우선, Outbound Packet을 대상으로, Destination IP, Port가 동일한 packet을 대상으로 Source IP별 하위 분류를 한다. Source IP별 하위 분류가 완료되면, 10분/1시간/1일/한달 등의 단위로 의심 Destination IP, Port에 대해 주기적 접속 여부를 분석할 수 있는 기반이 마련된다. Fig. 5.는 생성된 데이터를 기반으로 일정주기마다 주기적 접속 여부를 점검하는 알고리즘을 나타낸다.

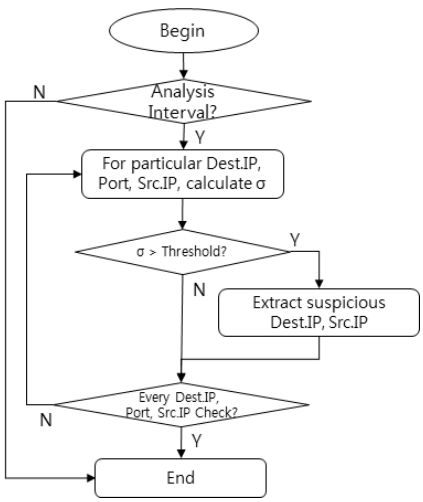

Fig. 5. Detection of C&C and Zombie PC

Destination IP, Port, Source IP에 따른 시간값에 대해, 표준편차 방식, 평균편차 방식 등을 적용하여, 특정 임계값 이하의 경우, 주기적 접속이 발생한 것으로 판단할 수 있다. 일반적인 트래픽 발생패턴을 보면, 각 값들의 편차가 상당한 차이로 나타나게 되는데, 표준편차는 많은 연산이 발생하고, 전반적인 분포수준이 아닌 일정한 분포모양을 띄고 있는지 확인하기 위함이므로, 여기서는 평균편차 방식을 통해, 임계값 5초 이상의 평균편차를 보일 경우, 이때의 Destination IP는 C&C, Source IP는 내부의 감염PC로 의심할 수 있다. 앞서 대용량 데이터 처리를 위한 NoSQL DB 스키마는 이러한 주기적 접속여부 탐지를 효과적으로 지원한다.

3.3.2 스캐닝 시도 감염의심 IP 탐지

공격자는 악성코드에 감염된 PC를 확보하면 내부망의 구성 및 주요 시스템을 파악하고, 이들 시스템에 침입해 들어갈 수 있는지 확인하게 되는데, 이는 결국 오픈된 Port가 있는지 검색하는 형태로 나타나게 된다. 이러한 스캐닝 시도는 외부에서 내부로도 가능하지만, 내부에서 내부로도 가능하다. 탐지를 위한 데이터는 다음과 같이 생성한다. Fig. 6.은 스캐닝 시도탐지를 위한 기본적인 데이터 생성과정을 나타낸다.

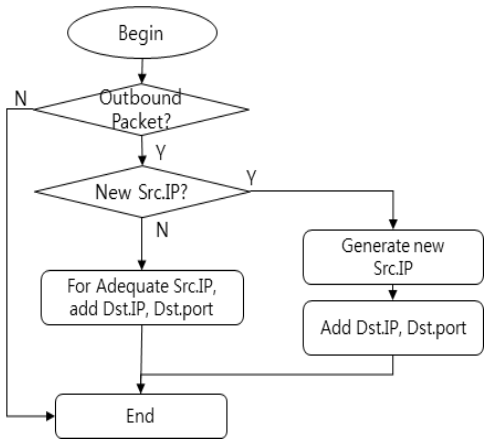

Fig. 6. Scanning Attack Detection Algorithm

위와 같이 생성된 데이터들을 대상으로, 주기적으로 스캐닝 시도가 발생했는지 점검해 나가면 된다. 특정 Source IP에서 다수의 IP들로 지속적인 접속시도가 발생하거나, 특정 Source IP에서 특정 IP의 여러 Port에 대한 접속시도가 발생할 경우, Source IP는 악성코드에 감염된 PC로 의심해볼 수 있다.

3.3.3 대용량 데이터 처리 플랫폼

이상징후 탐지·분석을 위해서는 최소 한달 정도의 트래픽 정보를 수집·관리할 수 있어야 하는데, 기존 관계형 데이터베이스에서 처리하는데 한계가 있어, 확장성을 가진 HDFS 시스템을 활용한다. 네트워크 용량이 1Gbps인 환경에서 트래픽을 수집하는 빅데이터 플랫폼을 구성하려면, 평상시 네트워크 트래픽 발생 비율을 평균 40 %로 가정할 때 초당 50Mbytes 즉, 분당 3GBytes의 데이터 수집이 필요하다. HDFS 기반 대규모 트래픽 수집·저장을 위해서 아래와 같이 1개의 Master Node, 6개의 Slave Node로 구성했으며, 데이터 관리규모에 따라 확장 가능하다. Fig. 7.은 빅데이터 플랫폼에 데이터를 관리하는 구조를 나타낸다.

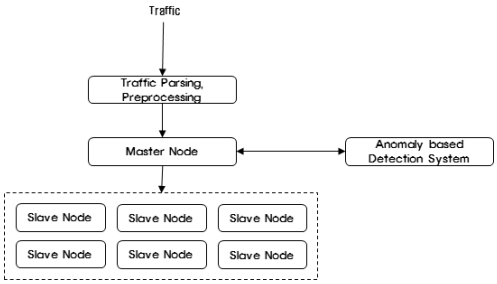

Fig. 7. HDFS Data Management Scheme

수신된 트래픽은 이상징후 탐지로직에 따라, 아래 표의 데이터를 실시간 수집·처리하여 빅데이터 플랫폼인 HDFS에 저장한다. Destination IP를 Row key로 하며, Source IP, Port 등을 포함한 나머지 데이터 항목을 각 Column의 값으로 정의한다. Row key인 Destination IP를 기준으로 Source IP와 Destination Port등 여러 Column의 항목들을 하나의 Column Family로 묶어 정의가 가능 하다. 각 셀들마다 시간에 따라 구분하는 Time Stamp가 존재하며, 이를 기준으로 시간에 따른 데이터의 추출이 가능하다. Table 2는 NoSQL 데이터 관리 포맷을 나타낸다.

Table 2. NoSQL Data Format

- Time : start time, end time - IP/Port : Src IP/Port, Dst IP/Port - Flag : Inbound/Outbound - Protocol : TCP/UDP - Session Info : Session Duration, PPS, BPS

IV. 시험결과

본 논문에서 제안한 시스템을 상용환경의 A업체에 적용하여, 시험결과를 도출하였다. 시험환경은 직원수가 45명 정도 수준(PC, 스마트폰 포함 약 100대)이며, 일평균 Outbound 패킷수는 14만~37만, Inbound 패킷수는 8~25만 정도이며, Outbound 트래픽에서 볼 때, Destination IP의 수는 일평균 3만~8만 정도 수준이다. 이러한 트래픽 환경에서 분석서버는 5-Tuples, HTTP Request URI, HTTP Response Filetype, 세션당 PPS, BPS 등

Anomaly 탐지로직에 필요한 데이터를 추출하여 검증하였다.

4.1 주기적 접속 기반 이상징후 탐지 성능

주기적 접속여부에 대한 탐지 결과는 다음과 같다. 시험결과, IP, Port 기준으로 IP별 일평균 약 300건 수준의 커넥션 시도가 발생하였다 Fig. 8.은 IP, Port 기준 주기적 접속이 발생했던 Destination, Source, 접속주기 정보를 나타낸다. 예를 들어, 첫 번째 줄의 경우, 내부의 10.0.0.107에서 외부의 164.124.101.2 으로 20분마다 주기적 접속이 발생했음을 의미한다.

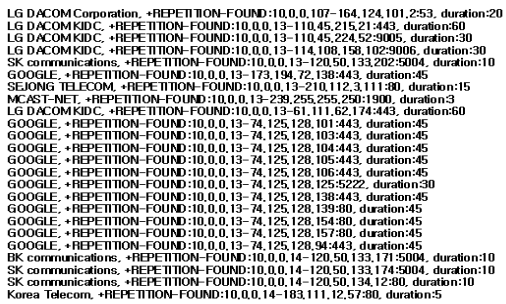

Fig. 8. Repetive Connection Detection Result

세부내역을 분석해보면, 대부분의 경우 잘 알려진 서비스들에 접속하는 통신으로 화이트리스트 처리가 가능하다. 이러한 주기적 접속 내역 중, 기존에 없던 전혀 새로운 Destination으로 트래픽이 발생하면 이는 이상 트래픽으로 의심해 볼 수 있다. 특히, 주기적 접속내역 분석결과, 특정 211.254.228.26 (update.windowupdate.org)로의 주기적 접속은 알려지지 않은 의심 Destination으로, Source IP를 대상으로 백신S/W 분석결과 실제 악성코드가 탐지되었다. 이 악성코드를 제거 후, 해당 Destination으로의 트래픽은 더 이상 발생하지 않음을 확인하였다. 이와 같이 주기적 접속 기반 분석은 가능하지만, 일평균 300건에 이르는 등 많은 오탐지 정보를 같이 수반하기 때문에, 신뢰할 수 있는 Destination 대상 white-list에 대한 처리 및 다른 알고리즘과 연관분석이 필요하다. Table 3은 주기적 접속탐지에 대한 배치분석 성능을 나타낸다. 주기적 접속분석은 10분, 1시간마다 배치처리로 분석하는 데이터로, 분석시간은 무리가 없는 것으로 판단된다.

Table 3. Repetive Connection Detection

Division	num of pkts	data size	time	performance
10 minutes	1,500,000	864MByte	7.1sec	121.7MByte/s
1 hour	5,270,000	3,035MByte	28sec	108.4MByte/s

스캐닝 시도의 경우, Source IP, Destination IP가 동일하고, 각각 다른 20개 이상의 Port번호가 사용된 경우를 분석하였는데, 이번 시험환경에서는 탐지되지 않았다. 이는 대규모 빅데이터 환경에서 장기간 데이터에 대해 분석하는 것이 필요할 것으로 예상된다.

4.2 대용량 데이터 처리성능

대용량 네트워크 트래픽에서의 이상공격 탐지를 위해, 트래픽 수집 서버, HDFS 기반 데이터 저장을 위한 Master Node 1대, Slave Node 6대로 구성했을 때 데이터 처리성능을 측정하였다. 시험결과, 각 flow를 개별적으로 수집하면, 24,210 packets/s가 처리가능한데, 10개 패킷을 묶어서 한번에 처리할 경우, 101,748 packets/s 처리 성능을 보였다. 이는 패킷의 평균크기를 576Byte로 가정해보면[20], 약 58MByte/s의 데이터 수집이 가능하였다. 1Gbps 환경에서 평균 트래픽이 40%일 경우, 3MByte/s 수준에 해당하므로, 1Gbps 환경에서 제안한 시스템이 원활히 동작함을 확인하였다.

V. 결 론

본 논문에서는 특정 조직·기업 대상으로 발생하는 APT 공격을 탐지하기 위해, 악성코드 감염PC와 공격자가 운영하는 C&C서버와의 비정상 트래픽을 탐지하는 3가지 탐지 알고리즘을 제시하고, 검증을 수행하였다. 또한, 1Gbps 환경에서 탐지 알고리즘 구축을 위해 HDFS 기반 플랫폼 및 배치처리를 위한 NO SQL을 이용하였다. 개발 결과물은 상용환경에서 모두 검증해야 하지만, 알고 있는 공격유형의 set을 발생시키기 어려운 한계가 있었다. 향후, APT공격에 사용된 실제 샘플을 같이 발생시켜 검증하는 과정이 필요하며, 시험결과에 대한 통계, 시각화 기능 보완이 필요하다.

또한, 장기간 데이터를 안정적으로 분석하기 위해 주기적 이상징후 분석을 위한 시간주기를 설정하는 등의 배치처리 정책은 누적된 데이터의 크기, 분석 Scheme의 복잡성, 분석 요구사항, 비용 등에 따라 달라진다. 향후 빅데이터 플랫폼에서 장기간 운영하면서 이상징후 탐지할 수 있는 기술개발을 진행할 예정이다.

References

[1] Elshoush, H. Tagelsir, and I. M. Osmank, "Alert correlation in collaborative intelligent intrusion detection systems – A survey," Applied Soft Computing In Press, 2011.

[2] K. Julisch, "Mining alarm clusters to improve alarm handling efficiency," Proceedings of the 17th Annual Conference on Computer Security Applications, 2001.

[3] S. Cheung, U. Lindqvist, and M.W. Fong, "Modeling multistep cyber attacks for scenario recognition," DARPA Information Survivability Conference and Exposition, pp.284-292, 2003.

[4] H. Debar, and A. Wespi, "Aggregation and correlation of intrusion detection alerts," Proceedings of the International Symposium on Recent Advances in Intrusion Detection, pp.85-103, 2001.

[5] B. Morin, L. Me, H. Debar, and M. Ducasse, "M2D2: A formal data model for IDS alert correlation," Proc. Recent Advances in Intrusion Detection, pp.115-137, 2002

[6] P. Ning, Y. Cui, and D. Reeves, "Analyzing intensive intrusion alerts via correlation," Proceedings of the International Symposium on the Recent Advances in Intrusion Detection, pp. 74-94. 2002.

[7] P. Ning, Y. Cui, and D.S. Reeves, "Constructing attack scenarios through correlation of intrusion alerts," Proc. ACM Conf. Computer and Comm. Security, pp. 245-254, 2002.

[8] F. Cuppens, "Managing alerts in a multi-intrusion detection environment," 17th Annual Computer Security Applications Conference, 2001.

[9] X. Qin, and W. Le, "Statistical causality of infosec alert data," Proceedings of Recent Advances in Intrusion Detection, 2003.

[10] W.L. Xinzhou Qin, "Statistical causality analysis of infosec alert data," Lecture Notes in Computer Science, 2003.

[11] A.Valdes and K. Skinner, "Probabilistic alert correlation," RAID 2001.

[12] O.Dain and R.Cunninghan, "Building scenarios from a heterogeneous alert stream," IEEE Workshop on Information Assurance and Security, 2001.

[13] Munsun Shin, Eunhui Kim, Hosung Mun, Keunho Ryu and Kiyoung Kim, "Data mining based alarm data analysis implementation," KCC : database 31(1), 2004.2.

[14] F. Xiao, S. Jin and X. Li, "A novel data mining-based method for alert reduction and analysis," Journal of Network, vol. 5, no. 1, 2010, pp. 88-97.

[15] Ning P and Cui Y (2002), "An intrusion alert correlator based on prerequisites of intrusions," TR-2002-01

[16] S. Noel and S. Jajodia, "Correlating intrusion events and building attack scenarios through attack graph distance," In Proceedings of the 20th Annual Computer Security Applications Conference (ACSAC'04), 2004.

[17] C. Abad, Y. Li, K. Lakkaraju, X. Yin, and W. Yurcik. "Correlation between netFlow system and network views for intrusion detection in workshop on link analysis," Counter-terrorism, and Privacy held in conjunction with the SIAM International Conference on Data Mining, 2004.

〔18〕 Suhyung Lee, Hyochan Bang, Byunghwan Jang and Jungchan Na, "Security event processing for effective analysis," Electronics and Telecommunications Trends, 22(1), 2007.2.
〔19〕 A. Rao and S. Zang, "HBase-0.20.0 performance evaluation, "http://cloudepr.blogspot.com/2009_08_01_archive.html
〔20〕 Rishi Sinha, et.al, "Internet packet size distributions: some observations," Technical Report ISI-TR-2007-643, USC/Information Sciences Institute, May, 2007

〈저자소개〉

손 경 호 (Kyung Ho Son) 정회원
2001년 2월 : 성균관대학교 전기전자컴퓨터공학과 학사
2004년~현재 : 성균관대학교 컴퓨터공학과 석·박사과정 수료
2001년 1월~현재 : 한국인터넷진흥원 정보보호산업기획팀
〈관심분야〉 침해사고대응기술, 융합보안, 네트워크보안, 보안 시험·평가, 클라우드·빅데이터 보안

이 태 진 (Tai Jin Lee) 정회원
2003년 2월 : 포스텍 컴퓨터공학과 학사
2008년 2월 : 연세대학교 컴퓨터공학과 석사
2003년 ~ : 한국인터넷진흥원 책임연구원
〈관심분야〉 네트워크 보안, 시스템 보안, 악성코드 탐지 및 분석

원 동 호 (Dongho Won) 종신회원
1976년~1988년: 성균관대학교 전자공학과 (공학사, 공학석사, 공학박사)
1978년~1980년: 한국전자통신연구원 전임연구원
1985년~1986년: 일본 동경공업대 객원연구원
1988년~2003년: 성균관대학교 교학처장, 전기전자 및 컴퓨터공학부장, 정보통신대학원장,정보통신기술연구소장, 연구처장
1996년~1998년: 국무총리실 정보화추진위원회 자문위원
2002년~2003년: 한국정보보호학회장
2002년~2003년: 감사원 IT 감사 자문위원
2002년~2004년: 대검찰청 컴퓨터 범죄 수사 자문위원
2002년~2004년: 산학연 정보보안협의회 회장
2011년~2014년: 스마트그리드보안 워크샵 조직위원장
현재 : 성균관대학교 정보통신공학부 교수, 한국정보보호학회 명예회장
〈관심분야〉 정보보호, 암호이론, 정보이론