https://doi.org/10.33778/kcsa.2021.21.3.013

MITRE ATT&CK 및 Anomaly Detection 기반 이상 공격징후 탐지기술 연구★

황 찬 웅*, 배 성 호**, 이 태 진***

요 약

공격자의 무기가 점차 지능화 및 고도화되고 있어 기존 백신만으로는 보안 사고를 막을 수 없으므로 endpoint까지 보안 위협이 검토되고 있다. 최근 endpoint를 보호하기 위한 EDR 보안 솔루션이 등장했지만, 가시성에 중점을 두고 있으며, 이에 대한 탐지 및 대응 기술은 부족하다. 본 논문에서는 보안 관리자 관점에서 효과적인 분석과 분석 대상을 선별하기 위해 실 환경 EDR 이벤트 로그를 사용하여 지식 기반 MITRE ATT&CK 및 AutoEncoder 기반 Anomaly Detection 기술을 종합적으로 사용하여 이상 공격징후를 탐지한다. 이후, 탐지된 이상 공격징후는 보안 관리자에게 로그 정보와 함께 alarm을 보여주며, 레거시 시스템과의 연계가 가능하다. 실험은 5일에 대한 EDR 이벤트 로그를 하루 단위로 탐지했으며, Hybrid Analysis 검색을 통해 이를 검증한다. 따라서, EDR 이벤트 로그 기반 언제, 어떤 IP에서, 어떤 프로세스가 얼마나 의심스러운지에 대한 결과를 산출하며, 산출된 의심 IP/Process에 대한 조치를 통해 안전한 endpoint 환경을 조성할 것으로 기대한다.

MITRE ATT&CK and Anomaly detection based abnormal attack detection technology research

Chan-Woong Hwang*, Sung-Ho Bae**, Tae-Jin Lee***

ABSTRACT

The attacker's techniques and tools are becoming intelligent and sophisticated. Existing Anti-Virus cannot prevent security accident. So the security threats on the endpoint should also be considered. Recently, EDR security solutions to protect endpoints have emerged, but they focus on visibility. There is still a lack of detection and responsiveness. In this paper, we use real-world EDR event logs to aggregate knowledge-based MITRE ATT&CK and autoencoder-based anomaly detection techniques to detect anomalies in order to screen effective analysis and analysis targets from a security manager perspective. After that, detected anomaly attack signs show the security manager an alarm along with log information and can be connected to legacy systems. The experiment detected EDR event logs for 5 days, and verified them with hybrid analysis search. Therefore, it is expected to produce results on when, which IPs and processes is suspected based on the EDR event log and create a secure endpoint environment through measures on the suspicious IP/Process.

Key words : EDR, MITRE ATT&CK, Anomaly Detection, AutoEncoder, Process Analysis

접수일(2021년0 7월 30일), 게재확정일(2021년 09월 15일)

★ 본 연구는 2020년도 호서대학교의 재원으로 학술연구비 지원을 받아 수행되었습니다(2020-0419)

* 호서대학교 정보보호학과 대학원생(주저자)

** 호서대학교 정보보호학과 학부생(공동저자)

*** 호서대학교 컴퓨터공학부 교수(교신저자)

1. 서 론

2020년 코로나19가 전 세계적으로 유행하면서 사이버 공간에서도 관련된 위협이 지속되고 있다. 한국 인터넷진흥원의 2020년 사이버 위협 동향 보고서[1-2]에 따르면, 코로나19, 미국 대선 등 사회적 이슈를 악용하여 스마트폰, PC를 노리는 피싱, 스미싱이 증가하고, 재택근무 등 비대면 활동이 늘어나면서 RDP(Remote Desktop Protocol) 무차별 대입 공격이 증가했다. 이뿐만 아니라 불특정 다수를 대상으로 하는 기존 랜섬웨어 공격에서 특정 목표를 겨냥하고 장기간 공격을 수행하는 APT(Advanced Persistent Threat) 위협그룹이 증가했다는 사실만으로 공격의 배후를 추적하는 것은 어려워졌다.

통계적으로 악성코드 규모는 많이 줄고 있지만[3], 공격자들이 표적을 노려 드러나지 않는 방식으로 공격하기 때문에 위협의 감소로 해석하기보다는 위협이 드러나지 않는 것이다. 알려진 공격의 감소에도 불구하고 전 세계적으로 스피어피싱, APT 공격, 랜섬웨어 등 표적 공격이 점차 늘어가는 추세이기 때문에 이를 염두에 두고 대응할 필요가 있다.

기존 보안의 초기 대응은 패턴 기반의 방화벽, Anti-Virus 및 AI 기반 악성코드 분석을 통해 endpoint로의 악성코드 유입에 대응하고 있다. 하지만 랜섬웨어 등 알려지지 않은 공격과 파일 생성 없이 시스템에 피해를 주는 파일리스(fileless)와 같은 공격 때문에 필연적으로 미탐지를 수반하게 된다. 또한 APT 공격은 침투부터 내부 검색 및 정보수집을 통한 데이터 접근까지 오랫동안 단계적으로 진행되기 때문에 즉각적인 대응이 어렵다. 이에 대응하기 위해 다수의 업체에서 EDR(Endpoint Detection and Response) 기술을 개발하고 있지만 대부분 가시성 확보에 머무르는 경우가 많으며, 정작 공격 탐지 및 대응에 필요한 정보를 제공하기 어렵다. 따라서, 백신만으로는 보안 사고를 막을 수 없고, endpoint 레벨까지 보안 위협이 검토되고 있으며, 알려지지 않은 공격(unknown attack)에 대한 대응책이 필요하다.

따라서 본 논문에서는 endpoint 환경을 고려하여 사전지식 없이 동작하는 MITRE ATT&CK (Adversarial Tactics, Techniques 및 Common Knowledge) 및 Anomaly Detection 기반 이상 공격징후 탐지 기술을 제안한다. endpoint에서 발생하는 이벤트 로그를 이용해 Unknown Attack을 탐지하여 대응함으로써 효율적이고 안전한 보안시스템을 구축하고, 새로운 공격을 처리하는 데 소비하는 시간과 인력 비용을 줄일 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 Anomaly Detection에 관한 기존 연구의 특징 및 장단점에 관해 기술하고, MITRE ATT&CK에 관해 설명한다. 3장에서는 제안하는 MITRE ATT&CK 및 Anomaly Detection 기반 이상 공격징후 탐지 기술의 세부 동작 과정을 제시한다. 4장에서는 제안하는 기법을 적용한 endpoint 환경에서의 탐지 결과 및 해석 결과를 제시한다. 마지막으로 5장에서는 결론으로 마친다.

2. 관련 연구

2.1 MITRE ATT&CK

MITRE는 취약점 데이터베이스인 CVE(Common Vulnerabilities and Exposures)를 감독하는 비영리 단체로 ATT&CK이라는 사이버 공격 Tactics 및 Techniques에 대한 정보를 기반으로 개방하여 모든 사람이나 조직이 무료로 사용할 수 있는 보안 프레임워크를 제공한다[4]. 이는 실제 관측을 기반으로 공격자의 Tactics와 Techniques에 대한 전 세계적으로 접근 가능한 지식 기반이며, 14단계의 Tactics에 대해 기술하고 있고, Tactics을 달성하기 위한 실제 공격 방법인 Techniques을 항목별로 분류하여 제공한다. 따라서, Tactics에 따라 다양한 Techniques들이 존재할 수 있다. MITRE ATT&CK 프레임워크는 수년 동안 사용되어 왔지만, 이제는 강력한 IT 보안의 필요성을 인식하고 정보보안 프로그램의 성숙도를 높이기 위해 조직에서 더 많이 채택되고 있으며 민간, 정부, 사이버 보안 제품 및 서비스 커뮤니티에서 특정 위협 모델 및 방법론을 개발하기 위한 기반으로 사용되고 있다. 본 논문에서는 MITRE A

TT&CK 기반으로 Tactics별 Techniques들을 탐지하기 위한 Rule을 생성하여 의심 공격에 해당하는 이벤트 로그를 탐지한다.

2.2 Anomaly Detection 연구 동향

Anomaly Detection은 금융 사기 탐지, 네트워크 침입 탐지, 인간 행동 분석, 유전자 발현 분석 등 다양한 영역에서 사용하고 있고[5], 이상징후를 탐지하는 방법으로 여러 기법을 활용한 연구가 진행되었다[6-8]. Anomaly 기반 IDS(Intrusion Detection System)는 정상적인 행동에서 크게 벗어난 침입행위에 대해 탐지한다. 또한, feature로 활용하기 위해 로그의 시간 값, 페이로드 등이 어떤 의미와 활용성을 지니고 있는지에 관한 연구도 진행되고 있다[9-10].

Aljawarneha 등[11]은 IDS에서 anomaly 기반의 침입 탐지를 위한 모델을 제안했다. Anomaly 탐지를 위해 해당 논문에서는 의사결정트리, Neural Network, Nearest Neighbor 방식을 이용한 하이브리드 모델을 생성했다. KDD99 데이터를 이용하여 모델의 학습에 필요한 feature의 전처리 작업을 진행하였으며, 엔트로피의 차를 이용하는 Information Gain을 계산하여 일정 값 이상인 feature만을 선택하였다. 이를 통해 제안모델이 향상된 정확도와 탐지 시간 축소가 가능하다고 설명했다.

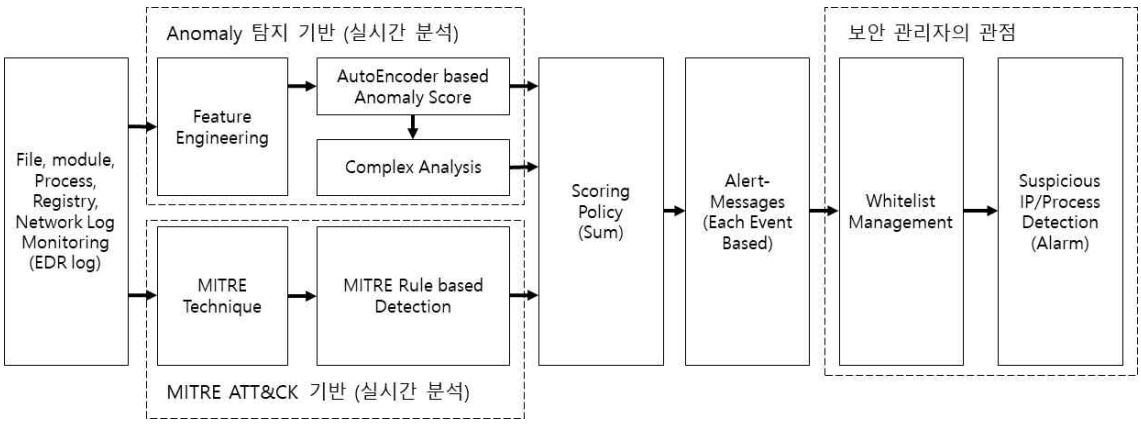
Andrew A. Cook 등[12]은 산업 IoT뿐만 아니라 Smart Energy, Smart City 등에서의 시계열 데이터에 대한 이상징후 탐지 방법들을 조사하였다. 시계열 데이터의 이상을 감지하기 위해 알고리즘과 접근 방식을 통계 및 확률, 패턴 일치, 거리 기반, 클러스터링 등 크게 6개의 그룹으로 분류하였으며 이에 따라 머신러닝을 이용한 탐지 방법을 세부적으로 정리하였다.

다크트레이스(darktrace)[13]는 데이터를 스스로 학습하고 network anomaly를 판단할 수 있는 기계학습 기반 기술을 갖추고 있다. 기존 APT 솔루션과 네트워크 포렌식 솔루션은 기존 이상 데이터에 의존하므로 데이터에서 찾을 수 없는 새로운 패턴이나 위협에 대응할 수 없다. 장치 및 네트워크 행위를 학습, 추론 및 시각화한다. 또한, 250개 이상의 위협 모델을 구축하고 이를 기반으로 위협을 탐지한다. 32개 경쟁사의 모델링에 비해 압도적인 수치이다. 이를 통해 더 정교한 위협 탐지가 가능하다. 이러한 연구 동향에서 anomaly 탐지를 위해 연구되는 기법들은 대부분 네트워크 로그 대상이 많은 것을 확인할 수 있다. 본 논문에서는 실제 상용 endpoint 환경에서 발생하는 네트워크만이 아니라 파일, 프로세스, 모듈과 같은 시스템 로그에서도 활용 가능한 Unknown Attack 탐지 기술을 제안한다.

3. 제안 모델

3.1 전체 시스템 개요

지금까지 알려진 공격에 대해서만 대응하는 백신에 의존해 왔다. 그러나, 점차 지능화되고 있는



(그림 1) 제안 모델의 구조

공격들에 대해서는 대응하지 못하며, 이미 악성코드에 감염된 endpoint에 대해서도 대응이 불가한 것이 현실이다. 본 논문에서는 EDR 환경에서 MITRE ATT&CK 및 Anomaly Detection 기반 이상 공격징후 탐지 기술을 제안한다. 제안 모델은 실제 큰 피해가 발생하기 전에 endpoint 이벤트 로그를 이용하여 이상 공격징후를 탐지 및 대응하는 것으로 추가적인 피해를 방지할 수 있다. 이상 공격징후를 탐지하기 위한 전체 시스템 구성은 그림 1과 같다. 신규 이벤트 로그의 이상 공격징후 탐지 방법은 크게 MITRE ATT&CK 및 Anomaly Detection 기반으로 탐지한다. 각각 탐지된 이벤트 로그는 하나의 분석 파일을 구성하고, 이벤트 로그별 탐지된 3개의 score를 종합판단하여 Alert를 발생시킨다. Alert를 발생시킨 이벤트 로그는 분석 정보가 없으므로 모델을 평가할 수 없다. 따라서, 보안 관리자 관점에서는 전문가 분석 결과를 바탕으로 Whitelist를 생성하고 관리할 수 있다. 본 논문에서는 Hybrid Analysis[14]를 통해 Whitelist를 생성 및 적용하고, 탐지된 이벤트 로그의 정보를 보안 관리자에게 Alarm을 발생시켜 신속한 대응을 할 수 있는 기반을 제공한다.

3.2 Anomaly Detection Approach

EDR 이벤트 로그는 라벨이 존재하지 않기 때문에 Unsupervised 방식으로 동작한다. endpoint에서 발생하는 이상 공격징후를 탐지하기 위해 기존에 발생한 이벤트 로그와 신규 발생 이벤트 로그 사이에 격차가 큰 데이터를 파악하여 anomaly를 탐지하는 방식을 이용한다. Anomaly Detection은 기존의 정상 로그 데이터에 대한 특성을 보유한 상태에서, 신규 이벤트 로그가 기존 이벤트 로그와 유사한지를 비교/파악하는 방식으로 이상 행위를 하는 이벤트 로그를 특정한다. EDR에서 수집된 이벤트 로그에서 feature engineering을 통해 유의미한 feature를 선정하여 가공한다. 이후, 딥러닝 기반의 AutoEncoder를 사용하여 학습하여 실제값과 예측값의 오차를 통계적으로 계산하여 anomaly score로 사용한다. 게다가, complex analysis는 일정 수준 이상의 anomaly score를 부여받은 네트워크 이벤트 로그를 대상으로 시스템 로그를 분석하여 이상 공격징후를 탐지한다.

3.2.1 Feature Engineering

EDR에서 수집된 이벤트 로그에서 이상 공격징후 탐지를 위한 feature vector가 필요하다. Feature Engineering은 EDR 이벤트 로그에서 유의미한 Field를 선정하고, 이를 통해 벡터화하여 신경망 학습을 위한 feature를 추출한다. 제안 모델에서는 표 1처럼 unusual behavior을 토대로 유의미한 Field를 선정하고, feature를 생성했다. 선정된 Field는 총 6개로 프로세스명, 목적지 ip, 이

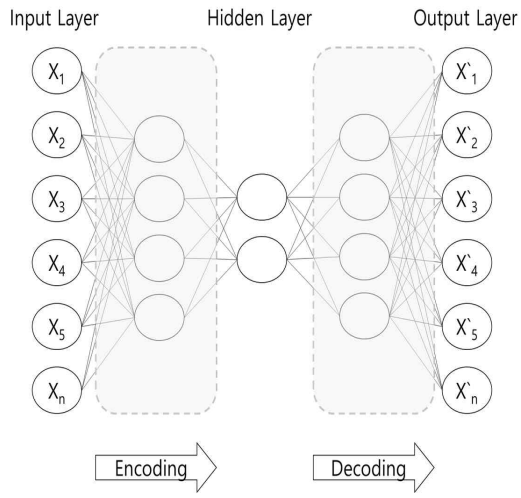
<표 1> 비정상적인 동작을 판별하기 위한 Feature Engineering

분류	세부 사항
비정상적인 네트워크 접속 시도	과거 단 한 번도 접속 기록이 없는 PC가 rare한 destination으로 네트워크 접속 발생
	동일 그룹 내 접속 이력이 없는 ip에 네트워크 접속 발생
	비정상적인 시간에 rare한 destination으로 네트워크 접속 발생 (주말, 오후 10시 ~ 오전 6시)
	비정상적인 주기로 rare한 destination으로 네트워크 접속 발생
	특정 process를 통해 network 접속 발생
비정상적인 데이터 다운로드/업로드	비정상적인 시간, 경로 및 주기로 PE/zip/script/dll 파일을 생성한 경우
	평소 다운로드하지 않은 크기의 PE/zip/script/dll 파일을 생성한 경우
	특정 process가 PE/zip/script/dll 파일 생성한 경우
프로세스의 비정상적인 데이터 전송	내부 사용자가 비정상적인 시간에 문서나 파일 액세스 (읽기, 삭제, 이동)
	특정 사용자가 과거에 사용하지 않던 process를 실행(powershell, WMI 등)
	Script를 다운로드 받고, 이를 실행하기 위한 process를 실행(wscript.exe, cscript.exe 등)

벤트 유형, 프로세스 유형, 파일 유형, 이벤트 시간이며 총 62개의 feature를 생성한다. 프로세스명은 2-gram feature hashing으로 50개의 feature를 추출한다. feature hashing 기법의 장점은 프로세스 경로와 목적지 IP에서 존재하는 모든 string 값을 feature로 활용할 수 있으므로 특정 문자가 탈락하면서 발생하는 데이터의 손실이 없다는 것이다. 목적지 ip도 마찬가지로 2-gram feature hashing으로 4개의 feature를 추출한다. 또한, 목적지 IP의 의심스러운 로컬 IP를 추가하기 위해 A 및 B 클래스 영역을 분리하여 Min Max Scaling을 적용한다. 목적지 IP feature는 네트워크 행위만 해당하며, 시스템 행위는 목적지 IP 필드가 존재하지 않기 때문에 0을 부여한다. 프로세스 유형, 이벤트 유형, 파일 유형은 해당하는 종류에 따라 feature를 각각 다른 값으로 매핑(mapping)한다. 프로세스 유형은 크게 shell 계열, script 계열, word 계열, 일반 프로세스로 분류한다. 파일 유형은 크게 PE(Portable Executable), script, zip, doc, image, media, 기타 파일로 분류한다. 이벤트 유형은 네트워크 행위, 파일 행위, 모듈 행위, 프로세스 행위, 레지스트리 행위로 분류한다. 이벤트 시간은 요일과 평일 및 주말로 구분하였고, 3시간마다 feature를 각각 다른 값으로 mapping 한다.

3.2.2 AutoEncoder based Anomaly Score

AutoEncoder는 비지도(unsupervised) 방식에서 주로 사용하며 encoder와 decoder 과정이 서로 대칭을 이루는 구조로 그림 2와 같다. 학습하는 과정 중 encoding에서 입력 데이터에 대한 차원을 축소하고, decoding에서 입력 데이터를 재구성한다. 즉, 신경망(neural network)에서 입력과 출력을 최대한 비슷하게 만들어 내는 가중치를 찾아내는 것이 AutoEncoder의 특징이다. 이러한 특징을 이용해 이상 공격징후를 탐지할 수 있다. AutoEncoder를 통해 일정 기간의 정상으로 가정한 데이터를 이용하여 학습 모델을 생성하고, 신규 이벤트 로그가 입력되었을 때 학습 모델과의 예측값과 신규 이벤트 로그인 실제값에 대한 오차를 손실(loss)값으로 사용한다. 손실값의 계산 함수는



(그림 2) AutoEncoder 구조

수식 1과 같이 평균 제곱 오차(Mean Squared Error, MSE)를 사용한다.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{Y}_i - Y_i\right)^2 \quad (1)$$

산출된 loss 값은 anomaly score를 사용할 수 있지만, 누적분포함수(Cumulative Distribution Function, CDF)를 기반으로 통계해석을 진행하였다. 누적분포함수는 누적확률분포라고도 하며, 확률 변수 X에 대한 누적확률분포 F(x)의 수학적 정의는 수식(2)과 같다. 몇 가지 누적확률분포 표시의 예를 들어 확률 변수가 -∞이상, -1 미만의 구간 내에 존재할 확률은 수식(3), 확률 변수가 -∞이상, 0 미만의 구간 내에 존재할 확률은 수식(4), 확률 변수가 -∞이상, 1 미만의 구간 내에 존재할 확률은 수식(5)과 같다.

$$F(x) = P(\{X < x\}) = P(X < x) \quad (2)$$
$$F(-1):\ P(\{-\infty \le X \leftarrow 1\}) \quad (3)$$
$$F(0):\ P(\{-\infty \le X < 0\}) \quad (4)$$
$$F(1):\ P(\{-\infty \le X < 1\}) \quad (5)$$

학습데이터셋의 loss값을 이용하여 데이터의 통계 분석을 위해 평균과 표준편차, z-score를 계산하고, 이를 이용하여 테스트데이터셋의 CDF값을 계산하여 anomaly score로 사용한다. 따라서, anomaly score는 0과 1사이의 수로 표현되며, 일

정 수준 이상일 경우 이상 공격징후로 판단한다.

3.2.3 Complex Analysis

AutoEncoder를 통해 이벤트 로그 하나의 이상 공격징후를 탐지했다면, complex analysis는 2개 이상의 이벤트 로그를 분석하여 이상 공격징후를 탐지하는 방법이다. 최근 대부분의 공격 시나리오는 사회공학적 기법으로 악성파일을 다운받고 실행하면, 외부 C&C(Command and Control) 서버와 통신하여 악성행위를 한다. 따라서, 앞서 설명한 AutoEncoder에서 일정 수준 이상의 anomaly score를 갖는 네트워크 이벤트 로그를 대상으로 한다. 이러한 네트워크 행위가 발생 후, 5분 이내에 동일한 IP/Process에서 PE 또는 ZIP 형식의 파일생성 행위가 발생했을 때 이상 공격징후로 판단한다. 또한, PE/ZIP 형식의 파일생성 행위를 한 후, 5분 이내에 동일 IP/Process에서 일정 수준 이상의 anomaly score를 부여받은 네트워크 행위가 발생했을 때 이상 공격징후로 판단한다.

3.3 MITRE ATT&CK Anlaysis

MITRE ATT&CK analysis는 EDR 이벤트 로그의 Field 정보를 고려하여 MITRE ATT&CK에 제공하는 Tactics별 Techniques에 해당하는 이벤트 로그를 탐지한다. 이상 공격징후를 탐지하기 위해서 Tactics별 Techniques에 해당하는 MITRE Rule을 생성한다. MITRE Rule을 사용하여 이벤트 로그의 이벤트 유형, 프로세스 명, 프로세스 경로, 커맨드 라인(command line), 레지스트리 경로 Field 정보에 해당하는 이벤트 로그를 탐지한다. MITRE Rule은 크게 4개의 의심행위와 20개의 탐지명이 있다. 20개의 탐지명에는 65개의 세부적인 탐지명으로 분류한다. MITRE Rule 기반 탐지 방법은 EDR 이벤트 로그의 Field 정보를 사용하여 정의한 MITRE Rule에 해당이 될 때 Alert를 발생한다. 예를 들어, 그림 3은 모든 프로세스의 공통의심 행위에서 정상 프로세스로 위장한 의심 행위를 탐지하는 의사 코드(pseudo code)이다. Windows 보조프로그램에서 네트워크 연결을 시도하는 행위가 여기에 포함되며, 실제

Algorithm 1 – MITRE Rule 기반의 탐지 방법

Detection Name: 정상 프로세스로 위장한 의심 행위
Detailed Detection Name : Windows 보조 프로그램에서 네트워크 연결 시도

ProcName = 프로세스 명
EventType = 이벤트 유형
AuxProgram = ['calculator.exe', 'mspaint.exe', 'notepad.exe', 'snippingtool.exe']
Output : Alert 번호, Alert 명, Tactics, Techniques

```
1.  For event in EDR Event log:
2.      If ProcName in AuxProgram:
3.          If EventType == 'NetworkConnect':
4.              Alert Number = '1.1'
5.              Alert Name = 'Network Connection in Windows Auxiliary Program'
6.              Tactics = 'Initial Access'
7.              Techniques = 'Drive-by Compromis'
```

(그림 3) MITRE Rule 기반의 탐지 방법

Windows PC에서 자주 사용하는 계산기, 그림판, 메모장, 캡처도구 등이 Windows 보조프로그램에 해당한다. 이러한 정상적인 Windows 보조프로그램에서는 외부에 네트워크 연결을 시도하지 않으며, 이로 위장하여 네트워크 연결을 시도하면 모두 탐지하며 이상 공격징후로 판단한다. 이외에도 Shell, Script, Powershell에서의 의심행위와 의심스러운 시스템 정보 수집 행위 및 기타 공격의심 행위가 있다.

3.4 하나의 분석파일 기반 종합판단 정책

Anomaly, Complex, MITRE Rule에서 탐지된 이벤트 로그를 그대로 Alert를 발생시키면, 제안 모델의 성능 저하의 원인이 되며, 많은 양을 이벤트 로그를 보안 관리자가 분석해야 하기 때문에 정작 중요한 실제 공격상황을 놓칠 수 있다. 그러나, 실제 공격상황은 정상과 비교해 극히 드물며, 분석에 정말 필요한 이벤트 로그만 보안 관리자에게 Alert를 발생시켜야 하므로 전체적인 결과의 양을 줄이는 과정이 필요하다. 따라서, Anomaly, Complex, MITRE Rule 중 단 하나라도 이상 공격징후로 판단한 것을 하나의 분석 파일로 구성하고, 통합 분석을 위해 항목별 위험 정도에 따른 점수를 부여 후 합산 결과를 통해 종합적으로 판단한다. 합산 점수가 일정 수준 이상의 이벤트 로그만 보안 관리자에게 Alert를 발생시킨다. AutoEncoder 기반 Anomaly Score가 높은 경우 이상 공격징후로 판단하지만, 판단 근거에 대한 신

뢰성이 부족하므로 제일 낮은 가중치 점수를 부여한다. Complex analysis는 판단 근거가 명확하고, 공격의 초기 단계를 탐지하기 때문에 위험도가 제일 높은 가중치 점수를 부여한다. MITRE Rule은 20개의 탐지명에 따라 점수를 부여한다. 예를 들어, 의심스러운 정보 수집 행위는 공격에 직접적인 영향이 없는 정보 수집 행위로 위험도가 제일 낮은 가중치 점수를 부여한다. 하지만, cmd.exe와 같은 Shell 계열 프로세스에서 스크립트(script) 실행을 위한 Wscirpt.exe와 같은 Script 계열 프로세스를 실행될 때 공격에 직접적인 영향이 있으므로 위험도가 제일 높은 가중치 점수를 부여한다. 또한, 이상 공격징후로 판단한 이벤트 로그가 연속적으로 3번 발생했다면 해당 점수에 낮은 가중치 점수를 증가시키고, 5번 이상 발생했다면 해당 점수에 또 낮은 가중치 점수를 증가시킨다. 따라서, Anomaly, Complex, MITRE Rule 중 단 하나라도 이상 공격징후로 판단한 하나의 분석 파일에서 위험도에 따라 가중치 점수를 부여하고, 이것이 연속적으로 발생했을 경우 낮은 가중치 점수를 증가시킨다. 총 합산 점수가 일정 수준 이상일 경우 보안 관리자에게 Alert를 발생시킨다. 분석이 필수적인 위험도가 가장 높은 이벤트 로그는 일정 수준만큼 부여했기 때문에 총 합산 점수에 상관없이 보안 관리자에게 Alert가 발생하게 된다.

3.5 보안 관리자 관점에서의 운영

보안 관리자 관점에서는 분석 결과를 통해 탐지된 Alert에 대한 Whitelist를 관리하며 효율적으로 운영할 수 있다. 따라서, 보안 관리자가 Whitelist를 관리하면서, Alert 정보에 Whitelist를 적용하여 나온 최종 분석해야 할 IP/Process에 대한 정보를 Alarm 형태로 받는다. 또한, Alarm이 발생하면 보안 관리자는 전체 EDR 이벤트 로그를 분석할 필요 없이 Anomaly, Complex, MITRE Rule 중 단 하나라도 이상 공격징후로 판단한 하나의 분석 파일을 먼저 분석하여 시간 단축할 수 있다.

4. 실험결과

실험을 위한 개발 환경은 다음과 같다. 프로그래밍 언어로 Python 3.7을 사용하며, 신경망 프레임워크에는 TensorFlow 2.3 및 Keras 2.4.3과 해당 버전에 호환되는 CUDA를 사용했다. 어플리케이션으로 jupyter notebook 6.2.0 그리고 학습에는 NVIDIA의 Tesla T4 GPU가 사용되었다.

4.1 Dataset

실험에 사용된 이벤트 로그는 실제 상용 endpoint 환경에서 사용되고 있는 G사 제품의 로그이며, 총 89개의 field로 구성되어 있다. 실제 상용 환경에서 수집한 데이터이기 때문에 파일 이름

_index	IP	EventTim	ProcNam	ProcPath	EventTyp	EventSub	RemoteIP	LocalIP	ParentPro	FileType	CmdLine	RegNewK	RegKeyPa	RegValue	SHA256
endpoint2	10.58.*.*	1.59E+12	Kaka******	C:₩Progra	network	NetworkC	14.0.*.*	10.58.*.*							
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	network	NetworkC	23.45.*.*	10.58.*.*							
endpoint2	10.58.*.*	1.59E+12	extr*******	C:₩WINDC	process	ProcessStart			UCAI******	PE	"C:₩WINDOWS₩system32₩extrac32.exe" "(				bd050055f
endpoint2	10.58.*.*	1.59E+12	back*******	C:₩WINDC	process	ProcessStart			svch*******	PE	"C:₩WINDOWS₩system32₩backgroundTas				74b332340
endpoint2	10.58.*.*	1.59E+12	Disk*******	C:₩Progra	process	ProcessStart			UCAI******	PE	"C:₩Program Files (x86)₩LG Software₩LG				0b6aca1a1
endpoint2	10.58.*.*	1.59E+12	svch*******	C:₩WINDC	file	FileSetAttr									
endpoint2	10.58.*.*	1.59E+12	svch*******	C:₩WINDC	network	NetworkC	23.41.*.*	10.58.*.*							
endpoint2	10.58.*.*	1.59E+12	IsTh*******	C:₩Progra	process	ProcessStart			UCAI******	PE	"C:₩Program Files (x86)₩LG Software₩LG				ee655b29a
endpoint2	10.58.*.*	1.59E+12	Conh*******	C:₩WINDC	process	ProcessStart			IsTh*******	PE	₩??₩C:₩WINDOWS₩system32₩conhost.ex				baf97b2a6
endpoint2	10.58.*.*	1.59E+12	svch*******	C:₩WINDC	file	FileMove									
endpoint2	10.58.*.*	1.59E+12	svch*******	C:₩WINDC	file	FileMove									
endpoint2	10.58.*.*	1.59E+12	svch*******	C:₩WINDC	file	FileCreate									
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	process	ChildProcessCreate				PE	"C:₩WINDOWS₩system32₩cmd.exe" /c sch				4b2f2b322
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	module	ModuleLoad				PE					138d9f0ed
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	module	ModuleLoad				PE					6b6275eb6
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	module	ModuleLoad				PE					6647e5a26
endpoint2	10.58.*.*	1.59E+12	UCAI******	C:₩Progra	network	NetworkC	23.45.50.64	10.58.56.64							

(그림 4) 실험에 사용한 데이터셋 예시

<표 2> Dataset의 구성

데이터 분류	수집 기간	이벤트 수
학습 데이터	2020. 04. 01 ~ 2020. 04. 15	11,366,967
테스트 데이터	2020. 04. 16	2,355,271
	2020. 04. 17	1,015,683
	2020. 04. 18	417,066
	2020. 04. 19	824,853
	2020. 04. 20	1,427,097

과 파일 경로는 암호화되어있어 사용할 수 없다. 따라서, 우리는 로그 식별자 field를 제외하고, 실험에 필요한 15개 field를 시간순으로 추출하여 실험 데이터셋으로 사용했다. 그림 4는 사용한 데이터셋의 샘플을 보여주며, 민감한 정보는 비식별 조치했다. 데이터셋 구성은 2020년 4월 01일부터 15일에 사용된 11,366,967개 로그는 학습 데이터로 사용하며, 2020년 4월 16일부터 20일까지 5일에 걸쳐 하루 단위로 수집된 데이터를 테스트 데이터로 사용했다. 표 2는 본 논문에서 사용한 데이터셋을 정리한 표이다.

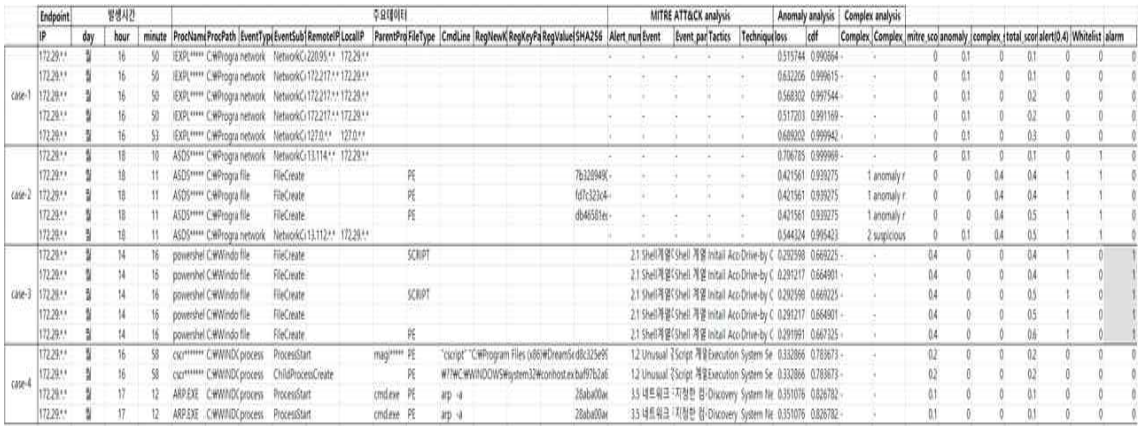
4.2 이상 공격징후 종합판단 결과

5일간의 테스트 데이터를 대상으로 Anomaly, Complex, MITRE Rule을 이용해 이상 공격징후 종합판단 결과 주말과 비교하여 평일에 많은 의심 이벤트 로그가 탐지되었다. 이는 각각의 endpoint에서 발생한 이벤트 로그이기 때문에 보안 관리자는 이벤트 로그 단위가 아닌 endpoint 단위로 분석할 수 있으며, 이상 공격징후로 최종판단된 결과는 표 3과 같다. 각각의 endpoint를 나타내는 IP를 기준으로 이벤트 로그의 주요 정보와 분석 결과를 통해 관리자는 신속하게 위험상황을 인지하고 대응할 수 있다. 그림 5는 최종 분석 결과의 일부를 보여준다. Anomaly, Complex, MITRE Rule에 의해 하나라도 탐지된 이벤트 로그는 유닉스 타임스탬프를 기준으로 표현된 이벤트 발생 시간을 한국 시간으로 변경하고, 원본 데이터의 정보와 Anomaly, Complex, MITRE Rule 각각의 분석 정보를 보여주며, 합산 점수 및 Whitelist 여부를 보여준다. 보안 관리자는 마지막 alarm에 해당하는 이벤트 로그 정보들을 이용하여 해석할 수 있다.

<표 3> Anomaly/Complex/MITRE 탐지 결과

수집 기간	탐지된 이벤트 수	Endpoint의 수
2020. 04. 16	2,449	13
2020. 04. 17	191	9
2020. 04. 18	31	4
2020. 04. 19	1,161	12
2020. 04. 20	4,669	27

(case-1) 탐지패턴은 월요일 16시 50분에 발생했으며, 일상적이지 않은 곳으로 네트워크 접속을 시도했다. 이는 네트워크 접속 행위이면서 anomaly score가 특정 임계치(threshold)인 0.99 이상이므로 rare한 네트워크 접속 이벤트 로그로 판단하



(그림 5) 이상 공격징후 종합판단 결과

며 anomaly 분석에 일차적으로 탐지되었다. 하지만 anomaly만으로는 이상 공격 여부를 판단하는데 있어 충분한 근거가 되지 않기 때문에 가장 낮은 0.1 점수를 부여했다. 또한 이상 공격징후로 판단한 이벤트 로그와 동일한 IP/Process에서 같은 의심 행위를 3번 이상 반복했기 때문에 3, 4번째 로그에 대해 0.1의 가중치 점수를 부여하였고, 5번 이상 반복한 5번 로그에 대해서는 0.2의 가중치 점수를 부여했다.

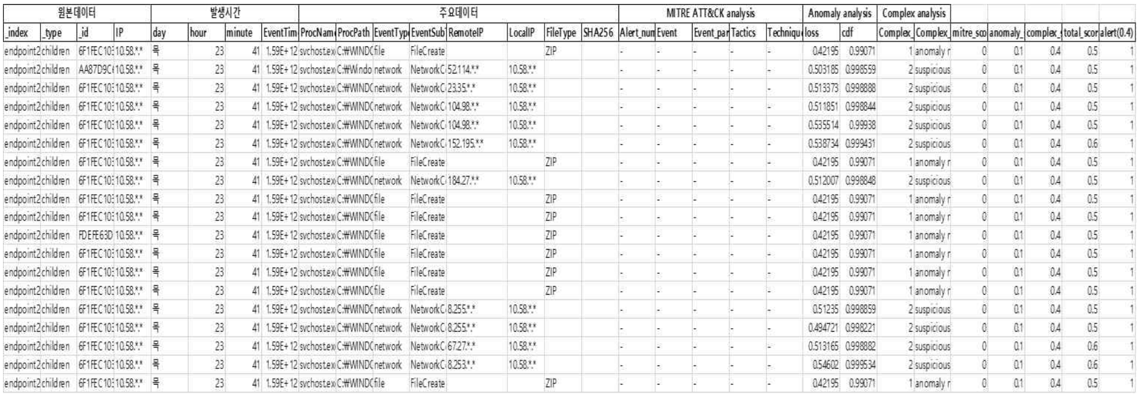
(case-2) 탐지패턴은 월요일 18시 10분에 발생했다. 해당 Process는 rare한 곳으로 네트워크 접속을 시도하여 anomaly에 탐지되었다. 이후 5분 이내에 동일한 IP/Process에서 파일 생성(file create) 행위가 발생했다. 이는 Complex의 1번 공격패턴에 해당하는 행위로 이상 공격징후로 탐지되었고, 같은 행위가 연속적으로 3번 발생하였기 때문에 0.1의 가중치 점수를 부여했다. 또한, 동일한 IP/Process에서 rare 한 곳으로 네트워크 접속을 시도가 발생하여 Complex 2번 패턴으로 탐지되었다. Complex score는 판단 근거가 명확하기 때문에 2가지 패턴 모두 0.4를 부여했다. Complex는 위험도가 높은 이벤트 로그이기 때문에 기본적으로 0.4 점수를 부여해 바로 보안 관리자에게 Alert를 발생시키도록 설계했다. 하지만 6~10번 로그 모두 사전에 Whitelist로 등록되어있기 때문에 정상로그로 판단하여 Alert를 발생시키지 않는다.

(case-3) 탐지패턴은 MITRE Rule에 의해 탐지된 이벤트 로그이다. Shell 계열인 Powershell 프로세스에서 script 파일을 생성한 행위로 MITRE Rule에 탐지되었다. 이러한 행위는 위험도가 높기 때문에 0.4 점수를 부여하여 바로 관리자에게 Alert를 발생시킨다. MITRE Rule에서도 마찬가지로 동일한 IP/Process에서 같은 의심 행위를 3번 이상 반복 시 0.1의 가중치 점수를, 5번 이상 반복 시 0.2의 가중치 점수를 부여했다.

(case-4) 탐지패턴은 프로세스가 비정상적 경로에서의 실행되거나 대상 PC의 IP/MAC 주소 테이블을 수집한 것으로 탐지되었다. 이는 각각 위험도에 따라 0.2, 0.1의 점수를 부여했지만, 보안 관리자에게 Alarm은 가지 않는다.

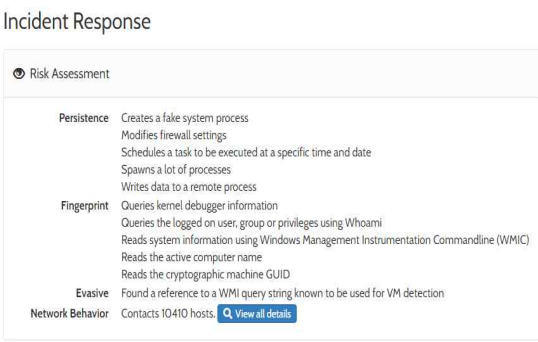
4.3 이상 공격징후 검증 결과

제안 모델은 Anomaly, Complex, MITRE Rule을 통해 종합적으로 이상 공격징후를 판단한다. 그러나, label이 없는 EDR 환경에서 실험을 진행했기 때문에 제안 모델을 평가할 수 없다. 따라서, 탐지된 이상 공격징후 데이터가 실제 악성 행위를 하는 프로세스인지 Hybrid Analysis[14] 검색을 통해 검증한다. Hybrid Analysis 사이트는 무료 악성 소프트웨어 분석 서비스로 검색 기능을 통해 관련 정보를 제공한다. 검색 결과가 없는 프로세스는 실제 전문가 분석이 필요한 것으로 생각된다. 탐지된 이상 공격징후 검증은 Alert로 탐지된 로그를 대상으로 프로세스 명을 검색하여 위험도 점수(threat score)에 따른 위협 수준을 기준으로 검증한다. 예를 들어, 2020년 4월 16일 목요일 데이터셋에서 가장 많은 Alert를 발생시킨 svchost.exe 프로세스는 윈도우즈 서비스를 백그라운드로 구동



원본데이터				발생시간				주요데이터								MITRE ATT&CK analysis					Anomaly analysis		Complex analysis						
_index	_type	_id	IP	day	hour	minute	EventTim	ProcNam	ProcPath	EventTyp	EventSub	RemoteIP	LocalIP	FileType	SHA256	Alert_num	Event	Event_par	Tactics	Techniqu	loss	cdf	Complex_	Complex_	mitre_sco	anomaly_	complex_	total_scor	alert(0.4)
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	AA87D9C	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩Windo	network	NetworkC	52.114.*.*	10.58.*.*			-	-	-	-	-	0.503185	0.998559	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	23.35.*.*	10.58.*.*			-	-	-	-	-	0.513373	0.998888	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	104.98.*.*	10.58.*.*			-	-	-	-	-	0.511851	0.998844	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	104.98.*.*	10.58.*.*			-	-	-	-	-	0.535514	0.99938	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	152.195.*.*	10.58.*.*			-	-	-	-	-	0.538734	0.999431	2	suspicious	0	0.1	0.4	0.6	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	184.27.*.*	10.58.*.*			-	-	-	-	-	0.512007	0.998848	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	FDEFE63D	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	8.255.*.*	10.58.*.*			-	-	-	-	-	0.51235	0.998859	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	8.255.*.*	10.58.*.*			-	-	-	-	-	0.494721	0.998221	2	suspicious	0	0.1	0.4	0.5	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	67.27.*.*	10.58.*.*			-	-	-	-	-	0.513165	0.998882	2	suspicious	0	0.1	0.4	0.6	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	network	NetworkC	8.253.*.*	10.58.*.*			-	-	-	-	-	0.54602	0.999534	2	suspicious	0	0.1	0.4	0.6	1
endpoint2	children	6F1FEC10	10.58.*.*	목	23	41	1.59E+12	svchost.ex	C:₩WIND	file	FileCreate			ZIP		-	-	-	-	-	0.42195	0.99071	1	anomaly r	0	0.1	0.4	0.5	1

(그림 6) 특정 Endpoint 대상 이상 공격징후 탐지 예시



(그림 7) Hybrid Analysis 기반 탐지된 프로세스 검증 예시

하는 프로세스이며, 윈도우즈 운영체제에서 필수적으로 정상 프로세스이다. 그러나, 그림 6처럼 특정 endpoint에서 15분간 rare한 곳으로 network 접속 및 신규 의심 ZIP/PE 파일이 생성되고, 이어서 평소 접속하지 않았던 곳으로의 network 접속을 반복했다. 또한, Hybrid Analysis 검색 결과 총 111개 행위 분석에 대해 63개의 malicious로 판단했으며, 이는 동일 process로 다양한 행위 중에 악성 행위로 판단한 것이 많다는 의미이다. 데이터셋 기준 가장 최근 검색된 정보에 따르면, 그림 7과 같이 해당 프로세스에서 발생한 위협 행위를 발견했다. 따라서, 과거에 악성으로 많이 사용된 프로세스이기 때문에 전문가 분석이 필요하며, 제안 모델은 이러한 의심 행위들을 탐지하는 데 이바지한다.

5. 결 론

단일 이벤트에 대한 AI 기반 이상 탐지 결과는 공격으로 확정할 수 없고, 공격의심으로 판단할 수 있다. 또한, MITRE 방식은 공격 순서가 아니라 공격 요소만을 언급한다. 따라서, 본 논문에서는 endpoint 환경에서 사전지식 없이 동작하는 MITRE ATT&CK 및 Anomaly Detection 기반 이상 공격징후 탐지기술을 제안하였다. 제안 모델은 특정 endpoint 대상으로 최종결과를 산출하였으며, 알고 있는 공격 시나리오에 대한 Tactics/Techniques 판단과 anomaly 혹은 정상이지만 알고 있지 않은 공격 의심 판단에 대한 종합적인 이상 공격징후를 탐지하고 검증했다. 실험에 사용한 데이터셋은 실제 endpoint에서 수집된 5일 치 이벤트 로그이며, 관리자는 하루 단위로 탐지된 의심 IP/Process를 분석하면 된다. 또한, Whitelist 운영을 연계하여 분석 대상을 큰 폭으로 줄일 수 있으며, 학습 데이터가 많으면 정교한 분석이 가능할 것으로 보인다.

참고문헌

[1] 한국인터넷진흥원, “사이버 위협 동향 보고서(2020년 1분기)”, pp. 1-104, 2020년 4월..

[2] 한국인터넷진흥원, “사이버 위협 동향 보고서(2020년 2분기)”, pp. 1-124, 2020년 7월..

[3] 하우리, “악성코드 분류별 통계”, https://www.hauri.co.kr/security/malicious_pop01.html, 2021년 3월..

[4] MITRE, “MITRE ATT&CK,” https://www.attack.mitre.org Mar. 2021.

[5] Chandola, Varun, Arindam Banerjee, and Vipin Kumar, "Anomaly detection: A survey." ACM computing surveys (CSUR), Vol. 41, No. 3, pp. 1-58, 2009.

[6] Ahmed, Mohiuddin, Abdun Naser Mahmood, and Jiankun Hu, "A survey of network anomaly detection techniques." Journal of Network and Computer Applications, Vol. 60, pp. 19-31, 2016.

[7] Abdallah, Aisha, Mohd Aizaini Maarof, and Anazida Zainal, "Fraud detection system: A survey." Journal of Network and Computer Applications, Vol. 68, pp. 90-113, 2016.

[8] Chalapathy, Raghavendra, and Sanjay Chawla, "Deep learning for anomaly detection: A survey", arXiv preprint, arXiv:1901.03407, 2019.

[9] Jabez, Ja, and B. Muthukumar, "Intrusion detection system (IDS): anomaly detection using outlier detection approach", Procedia Computer Science, Vol. 48, pp. 338-346, 2015.

[10] Bontemps, Loic, James McDermott, and Nhien-An

Le-Khac, "Collective anomaly detection based on long short-term memory recurrent neural networks", International Conference on Future Data and Security Engineering. Springer, Cham, 2016.

[11] Aljawarneh, Shadi, Monther Aldwairi, and Muneer Bani Yassein, "Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model", Journal of Computational Science, Vol. 25, pp. 152-160, 2018.

[12] Cook, Andrew A., Göksel Mısırlı, and Zhong Fan, "Anomaly detection for IoT time-series data: A survey", IEEE Internet of Things Journal, Vol. 7, No. 7, pp. 6481-6494, 2019.

[13] DARKTRACE, "Enterprise Immune System," https://www.darktrace.com/ ko/ Mar. 2021

[14] HYBRID ANALYSIS, "Hybrid-nalysis," https://www.hybrid-analysis.com/?lang=ko Mar. 2021.

[저 자 소 개]

황 찬 웅 (Chan-Woong Hwang)
2020년 2월: 호서대학교 정보보호학과 졸업
2020년 3월~현재: 호서대학교 정보보호학과 석사과정
<관심분야> 인공지능, 시스템 보안, 이상 탐지
email : hcw85123@gmail.com

배 성 호 (Sung-Ho Bae)
2016년 3월~현재: 호서대학교 정보보호학과
<관심분야> 정보보호, 인공지능, 이상 탐지
email : baesungho21@naver.com

이 태 진 (Tae-jin Lee)
2003년 1월~2017년 2월: 한국인터넷진흥원 R&D 팀장
2017년 3월~현재: 호서대학교 컴퓨터공학부 교수
<관심분야> 인공지능, 악성코드 분석, 시스템 보안
email : kinjecs0@gmail.com