J. Internet Comput. Serv.
ISSN 1598-0170 (Print) / ISSN 2287-1136 (Online)
http://www.jics.or.kr

MITRE ATT&CK 모델을 이용한 사이버 공격 그룹 분류☆

Cyber attack group classification based on MITRE ATT&CK model

최 창 희1* 신 찬 호1 신 성 욱1
Chang-hee Choi Chan-ho Shin Sung-uk Shin

요 약

정보통신 환경의 발전으로 인하여 군사 시설의 환경 또한 많은 발전이 이루어지고 있다. 이에 비례하여 사이버 위협도 증가하고 있으며, 특히 기존 시그니처 기반 사이버 방어체계로는 막는 것이 어려운 APT 공격들이 군사 시설 및 국가 기반 시설을 대상으로 빈번하게 이루어지고 있다. 적절한 대응을 위해 공격그룹을 알아내는 것은 중요한 일이지만, 안티 포렌식 등의 방법을 이용해 은밀하게 이루어지는 사이버 공격의 특성상 공격 그룹을 식별하는 것은 매우 어려운 일이다. 과거에는 공격이 탐지된 후, 수집된 다량의 증거들을 바탕으로 보안 전문가가 긴 시간 동안 고도의 분석을 수행해야 공격그룹에 대한 실마리를 겨우 잡을 수 있었다. 본 논문에서는 이러한 문제를 해결하기 위해 탐지 후 짧은 시간 내에 공격그룹을 분류해낼 수 있는 자동화 기법을 제안하였다. APT 공격의 경우 일반적인 사이버 공격 대비 공격 횟수가 적고 알려진 데이터도 많지 않으며, 시그니처 기반의 사이버 방어 기법을 우회하도록 설계가 되어있으므로, 우회가 어려운 공격 모델 기반의 탐지 기법을 기반으로 알고리즘을 개발하였다. 공격 모델로는 사이버 공격의 많은 부분을 모델링한 MITRE ATT&CKⓇ을 사용하였다. 공격 기술의 범용성을 고려하여 영향성 점수를 설계하고 이를 바탕으로 그룹 유사도 점수를 제안하였다. 실험 결과 제안하는 방법이 Top-5 정확도 기준 72.62%의 확률로 공격 그룹을 분류함을 알 수 있었다.

☞ 주제어 : 사이버 공격, 공격 그룹 유사도, 공격 그룹 분류, APT, MITRE ATT&CK

ABSTRACT

As the information and communication environment develops, the environment of military facilities is also development remarkably. In proportion to this, cyber threats are also increasing, and in particular, APT attacks, which are difficult to prevent with existing signature-based cyber defense systems, are frequently targeting military and national infrastructure. It is important to identify attack groups for appropriate response, but it is very difficult to identify them due to the nature of cyber attacks conducted in secret using methods such as anti-forensics. In the past, after an attack was detected, a security expert had to perform high-level analysis for a long time based on the large amount of evidence collected to get a clue about the attack group. To solve this problem, in this paper, we proposed an automation technique that can classify an attack group within a short time after detection. In case of APT attacks, compared to general cyber attacks, the number of attacks is small, there is not much known data, and it is designed to bypass signature-based cyber defense techniques. As an attack model, we used MITRE ATT&CKⓇ which modeled many parts of cyber attacks. We design an impact score considering the versatility of the attack techniques and proposed a group similarity score based on this. Experimental results show that the proposed method classified the attack group with a 72.62% probability based on Top-5 accuracy.

☞ keyword : Cyber attack, attack group similarity, attack group classification, APT, MITRE ATT&CK

1. 서 론

정보통신 기술의 가파른 발전으로 군사 시설과 국가 기반 시설의 많은 부분이 전산화되고 있다. 하지만, 이에 비례하여 보안 위협도 가파르게 증가하고 있다. 다양한 종류의 사이버 보안 위협 중 특히 APT(Advanced Persistence Threat) 공격은 방어가 매우 어렵다[1-4]. 이런 공격들은 대부분 국가가 지원하는 것으로 추정되고 있으며, 풍부한 자금 및 조직력으로 기존의 시그니처 기반 사이버 방어 기법을 우회하도록 설계되어 있다. 이에 대응하기 위해 많은 연구자가 시그니처가 변해도 본질적으로 변하지 않는 공격 행위에 집중하기 시작했고, 이를 모델링하고 분석하는 연구가 진행되었다[5].

미국의 군수 업체인 Lockheed Martin은 기존 군의 킬체

1. Defense Cyber Technology Center, Agency for Defense Development, Seoul, 05661, Korea.
* Corresponding author (changhee84@add.re.kr)
[Received 24 August 2022, Reviewed 17 September 2022(R2 14 October 2022), Accepted 25 October 2022]
☆ 본 논문은 2022년 한국 인터넷 정보학회 춘계학술발표대회에서 발표한 논문인 “공격 기술 정보 기반 사이버 공격 그룹 분류 기술”를 확장한 것이다.
☆ 이 논문은 2022년 정부(방위사업청)의 재원으로 국방과학연구소의 지원을 받아 수행된 연구임(912880601)

인 개념을 도입한 사이버 킬체인(Cyber Kill Chain)을 모델링 하였다[6-7]. 2018년에는 미국 정부의 지원을 받는 비영리 연구단체인 MITRE가 사이버 공격자의 관점에서 모델링한 MITRE ATT&CK을 출시하였고, 이는 빠른 속도로 발전하고 있다[8]. 현재 v11.3까지 출시되었으며, 사이버 보안 업계에 표준처럼 사용되고 있다. MITRE ATT&CK에서는 사이버 공격 그룹에 대한 정보도 제공하고 있는데, 이는 보안 전문가가 작성한 보고서, 블로그 글 등을 기반으로 작성한 것이다.

국가가 지원하는 공격으로 추정되는 APT 공격의 경우 피해가 발생하기 전에 공격 그룹을 알아내면 향후 예상되는 최종 목표, 사용될 공격 기술, 해킹 툴 등을 추정할 수 있어 사이버 공격 대응에 큰 도움이 될 수 있다. 과거에는 사이버 공격이 탐지된 후, 다수의 증거(evidence)를 수집하고 많은 시간을 들여 보안 전문가가 분석하여야 공격 그룹에 대한 실체를 밝혀낼 수 있었다. 하지만 이는 많은 시간이 지난 사후 분석으로 이미 피해를 본 후에야 공격 그룹을 분류할 수 있는 큰 단점이 존재한다. 또한 안티포렌식(Anti-Forensic) 기법[9]이 발달하면서 증거를 수집하기가 점점 어려워짐에 따라 증거기반 공격 그룹 분류도 어려워지고 있다. 더욱이 탐지를 회피하기 위해 적대적 공격 기법을 활용한 악성코드 변종 생성기술이 연구되고, 효과를 보임에 따라 국가 지원 추정 APT 공격의 실체를 밝혀내는 것이 점점 어려워지고 있다[10].

이에 대한 대책으로 증거, 시그니처 등의 정보가 아닌 호스트 내 공격 행위 정보를 기반으로 사이버 공격 탐지 및 대응(EDR: Endpoint Detection and Response) 하는 연구가 진행되고 있다[11]. EDR에서는 호스트 내에서 수행되는 행위에 대한 로그를 실시간으로 수집하고 각 행위를 연결하여 그래프를 생성하고 이를 이용하여 탐지하는 방식을 주로 사용한다[12]. 본 논문에서는 그래프 기반의 EDR 시스템이 공격 행위를 탐지하고 MITRE ATT&CK에서 정의한 공격 기술로 라벨링 하여 근 실시간으로 제공한다고 가정하고 연구를 진행하였다. EDR 시스템에서는 호스트를 감시하고 있다가 공격 행위가 일어나면 이를 탐지하여 MITRE ATT&CK의 공격 기술로 라벨링을 하고 이를 시간 혹은 논리 순서에 따라 공격 체인으로 엮어서 데이터를 제공한다. 통상적으로 EDR 시스템에서 공격 체인의 악성 여부를 판별하여 제공하므로, 본 연구에서는 악성으로 판별된 공격 체인만을 대상으로 공격 그룹을 판별하는 알고리즘을 연구하였다.

국가 지원 추정 APT 공격은 그 수가 매우 적고, 데이터도 완전하지 않은 경우가 많다. MITRE ATT&CK에서는 v10.0 기준 127개의 공격 그룹이 사용한 공격 기술을 분석하였는데, 본 논문에서는 이것을 학습 데이터로 사용하였다. 각 공격 기술마다 공격 그룹에 미치는 영향을 고려하여 영향성 점수를 정의하였고, 이를 이용하여 그룹 유사도 점수를 설계하였다. 국가지원 추정 APT 공격을 분석한 보고서를 테스트 데이터로 이용한 실험을 통해 제안하는 방법이 자동화가 가능하며, 또한 우수한 성능을 보이는 것을 입증하였다. 이후 논문의 구성은 다음과 같다. 2장에서는 공격 그룹 분류에 관한 과거 연구를 소개하고, 3장에서는 본 연구에서 활용하고 있는 MITRE ATT&CK에 대한 전반적인 내용을 기술하였다. 4장에서는 사이버 공격 그룹 분류를 위한 데이터셋 구축 방법을 설명하였다. 5장에서는 그룹 유사도 점수를 이용한 그룹 분류 방법에 대해서 제시하고, 6장에서는 검증을 위한 실험 설계, 결과 및 분석을 기술하였다. 마지막으로 7장에서는 결론 및 향후 연구 방향에 대해서 제시하였다.

2. 관련 연구

과거에는 공격 그룹에 대한 양질의 정보를 얻기가 어려웠고 정보의 정확도가 떨어졌기 때문에, 공격 그룹을 프로파일링하거나 예측하는 제한적인 연구가 주로 진행되었다. 2012년 Watters 연구진은 다양한 종속 변수와 독립 변수를 이용하여 사이버 공격자 모델을 정의한 연구를 진행하였다[13]. 이들은 공간적 스케일과 시간적 스케일을 고려하여 변수를 선택하였으며, 상관관계 분석을 통해 공격 그룹을 분석하였다. 2014년에는 Kapetanakis 연구진이 사건의 추론(Case-based reasoning)을 바탕으로 공격 그룹을 추론하는 연구를 진행하였다[14]. 공격 기술 레벨, 위험 회피, 교육 수준, 성별, 속도, 실수, 안티 포렌식 행위 등을 사용하여 공격 그룹을 평가하였다. 이들은 MyCBR [15]이라는 툴을 이용하여 공격 사건을 프로파일링하였다. Cho 연구팀은 도메인의 유사도를 비교하여 동일 공격 그룹을 예측하는 연구를 진행하였다[16]. 이들은 도메인 이름 시스템(DNS)에 질의를 할 수 있는 시스템 툴인 nslookup 및 악성코드 행위 분석을 통해서 도메인을 알아내고 유사한 도메인 이름을 가지면 동일 그룹으로 추정하였다. Han 연구팀은 웹 해킹을 기반으로 한 사이버 공격에서 추출할 수 있는 여러 가지 특징점에 대해서 중요도를 정의하고, 이를 특징점과 곱하여 합하는 방식으로 유사도 점수를 산출하였다[17]. 이들은 특징점들의 유사도를 규칙 기반으로 검출하여 사용하였다. 이러한 연구들은

특정 사건을 분석하거나 기존 공격 그룹들과의 유사성 산출에는 효과적이지만, 많은 공격 그룹을 대상으로 하는 식별이나, 분류에는 한계를 보인다.

2014년에 사이버 방어 작전 프레임워크 기반의 공격 그룹을 분류하는 방법이 제안되었다[18]. 공격 그룹이 공격을 진행할 때 생성되는 디지털 단서 및 공개 출처정보(OSINT: Open Source Intelligence)를 이용하여 그래프를 생성하고 원형 연결 기법 및 강제 직접 연결 기법을 활용하여 공격 그룹을 분류하였다. 이 논문은 전자메일, 첨부파일, 악성코드, OSINT에서 추출하는 시그니처를 기반으로 하고 있어서, 최신 APT 공격 대응이 어려운 단점이 존재한다. 2017년에는 전자메일 헤더의 통계 분석을 통해 사이버 공격자를 추적하는 연구가 진행되었다[19]. 이 연구에서는 이메일 헤더에서 IP, 언어셋, 발송 시간 등의 특징점을 추출하고 통계를 내서 사이버 공격자와 연관 지을 수 있는 단서를 제공하였다. 추가로 이러한 특징점을 이용하여 클러스터링하는 기법도 제안하였다[20]. 앞선 두 연구 또한 시그니처 기반의 연구로 최신 APT 공격에는 대응이 어렵다. APT 공격을 수행할 때 문서의 취약점을 이용하여 초기 진입(Initial access)을 하는 경우가 있는데, 이러한 악성 문서 분석을 통해 공격자 혹은 공격자의 시스템을 특정 지을 수 있는 특징점을 추출하는 연구가 진행되었다[21-22]. 해당 연구들 또한 시그니처 기반으로, 과거 다른 연구와 같은 문제점을 가지고 있다. 2020년에는 사이버 공격 정보의 연결 분석을 통해서 공격 그룹을 분석하는 연구가 진행되었다[23]. 이들은 악성코드 기반의 연관성을 분석하거나, 전파 기반의 연관성을 분석하여 군집화하는 방식으로 공격 그룹을 분석하였다. 해당 연구는 악성코드를 취득할 수 있는 공격 중반부 이후에 적용할 수 있고 수동으로 분석해야 하는 부분이 많아 조기 공격 탐지 및 예방은 어려운 단점이 있다. 위에 기술한 시그니처 혹은 증거기반으로 공격 그룹을 분류하는 기술들은 공통으로 안티포렌식이나 기만 기법에 취약한 단점이 존재한다.

공격 기술 모델 기반의 연구는 거의 이루어지지 않고 있다가 최근에 연구되기 시작하였다. 시그니처 기반 그룹 식별 혹은 분류 기술이 가지고 있던 단점을 극복하기 위해, 최근에는 MITRE ATT&CK를 활용한 연구가 진행되고 있다. 2021년 Shin 연구팀은 MITRE ATT&CK의 전술(Tactic)을 벡터화하고 이를 병합하여 APT 공격 그룹을 벡터로 표현하는 방법을 제안하였다[24]. 이들은 제안한 벡터끼리 코사인 유사도(Cosine similarity)를 이용하여 공격 그룹 유사도를 산출하였고, 각 전술에 대한 영향도를 고려하여 합계를 구할 때 이를 반영하는 방법을 사용하였다. 이들은 3개의 유명한 공격 그룹에 대해서 유사도를 분석하였다. MITRE ATT&CK의 기술 정보를 기반으로 구축한 데이터베이스를 활용하여 샴 네트워크(Siamese network)를 기반으로 공격 그룹을 분류하는 연구도 진행되었다[25]. 이들은 사이버 공격에 대해서 분석한 보고서를 수집하는 저장소에서 선별한 보고서에 대해 공격 기술을 자동으로 추출하여 데이터 셋을 구축하였다. 이를 전술 순서대로 정렬하여 시퀀스를 만들고, 양방향을 고려하여 샴 네트워크 기반의 딥러닝을 설계하여 그룹을 분류하였다. 이 방법은 공격 기술의 종류뿐만 아니라 공격이 일어난 순서가 그룹 분류에 많은 영향을 끼친다는 가정하에 진행된 연구이다. 또한 데이터의 수가 적을 수밖에 없는 환경에서 샴 네트워크를 활용하여 공격 그룹을 분류하였다. 해당 연구는 높은 정확도를 달성하였지만, 분류 가능한 조건을 만족하는 그룹 수가 적다는 단점이 존재하였다.

본 논문에서는 과거 기술들의 단점을 보완하기 위해 MITRE ATT&CK 기반으로 공격 기술에 영향성 점수를 부여하고, 이를 바탕으로 공격 그룹 유사도 점수를 설계하였다. MITRE ATT&CK에서 제공하는 공격 그룹이 사용했었던 공격 기술을 이용하여 학습하였고, APT 공격 분석 보고서를 데이터로 테스트하여 성능을 검증하였다.

3. MITRE ATT&CK

MITRE ATT&CK은 사이버 공격자의 전술과 공격 기술을 모델링한 것이다[8]. 2018년 v1이 공개된 이후로 여러 번 업데이트를 거쳐 2022년 4월 25일에는 v11이 공개되어 있다. 본 논문에서는 2021년 10월에 발표된 v10.0을 기준으로 연구가 진행되었다. 이 버전에서는 표 1와 같이 총 14개의 전술과 188개의 공격 기술 및 378개의 하위 공격 기술로 이루어져 있다. 하나의 공격 기술이 다수의 전술에 포함되는 경우가 표 1에서는 중복되어 집계되었다.

전술은 사이버 공격자가 세운 전략적 목적을 달성하려는 구체적인 실행 수단이나 방법 등을 의미한다. 예를 들면 “TA0001:Initial access”의 경우에는 공격자가 피해자의 시스템에 처음으로 접속하는 방법을 의미한다. 공격 기술은 전술보다 하위 개념으로써 전술을 달성하기 위한 세부적인 방법을 의미한다. 예를 들면 “TA0001:Initial access”를 달성하기 위해서는 “T1566:Phishing”을 방법을 사용할 수 있다. 하위 공격 기술은 이보다 더 세부적인 방법을

기술한 것이다. “T1566:Phishing” 방법 중 “T1566.001:Spearphishing attachment”는 이메일이나 문자메시지 등에 악성 코드나 스크립트를 첨부하여 감염을 시키는 하위 공격 기술이다. MITRE ATT&CK 이러한 공격 기술을 기존 로그 등을 활용하여 탐지하는 “Detection”이라는 항목도 설계해 두었다. 예를 들면 “T1566.001:Spearphishing attachment”을 탐지하기 위해서 표 2에 기술되어 있는 데이터 소스에서 로그를 추출하면 된다.

(표 1) 전술에 따른 공격 기술 개수
(Table 1) Number of attack techniques according to tactics.

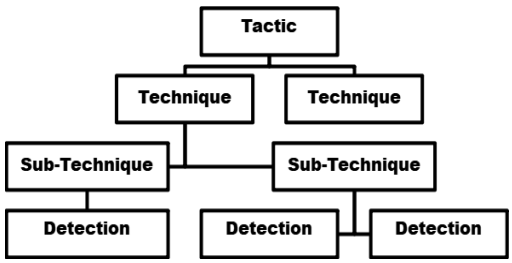
ID	이름	공격 기술 개수
TA0043	Reconnaissance	41
TA0042	Resource development	38
TA0001	Initial access	19
TA0002	Execution	33
TA0003	Persistence	106
TA0004	Privilege escalation	95
TA0005	Defense evasion	164
TA0006	Credential access	55
TA0007	Discovery	42
TA0008	Lateral movement	21
TA0009	Collection	36
TA0011	Command and control	38
TA0010	Exfiltration	17
TA0040	Impact	26

(표 2) T1566.001의 탐지를 위한 데이터 소스
(Table 2) Data source for detecting T1566.001

ID	데이터 소스	데이터 컴포넌트
DS0015	Application log	Application log content
DS0022	File	File creation
DS0029	Network traffic	Network traffic content
		Network traffic flow

MITRE ATT&CK은 그림 1과 같이 계층적으로 볼 수 있다. 본 논문에서는 그래프 기반의 EDR 시스템이 정보를 제공해준다고 가정하고 있는데, 이때 호스트에서 자료를 수집할 수 있는 소스가 필요하다. 즉 EDR시스템에서는 MITRE ATT&CK에서 정의한 데이터 소스에서 로그를 추출하고, 이를 하위 공격 기술, 공격 기술, 전술로 추상화하는 것이 가능하다. 과거에는 데이터 소스 레벨에서 시그니처를 추출하고 사이버 공격 그룹을 식별 또는 분류하는 데 사용했었지만, 본 논문에서는 이보다 2단계 추상화된 레벨인 공격 기술을 이용하여 공격 그룹을 식별하였다.

MITRE ATT&CK에서는 공격 그룹이 사용한 공격 기술들을 분석하였다. 버전 10에서는 127개의 공격 그룹이 분석되어 있으며, 이는 다양한 공격 분석 보고서 및 문헌을 토대로 조사한 것이다. 공격 그룹이 사용한 공격 기술과 사용한 소프트웨어가 표 3과 같이 기술되어 있다.



(그림 1) MITRE ATT&CK 계층적 구조
(Figure 1) Hierarchical structure of MITRE ATT&CK

(표 3) 공격 그룹이 사용한 공격 기술 및 소프트웨어
(Table 3) Attack techniques and software used by attack group

공격 그룹	공격 기술	소프트웨어
APT29	T1548.002, T1583.001 T1098.001, T1087, ...	S0552, S0635, S0054, S0154, ...
FIN6	T1134, T1087.002, T1560.003, T1119, ...	S0552, S0154, S0381, S0503, ...
menuPass	T1087.002, T1583.001, T1560.001, T1119, ...	S0552, S0160, S0144, S0106, ...

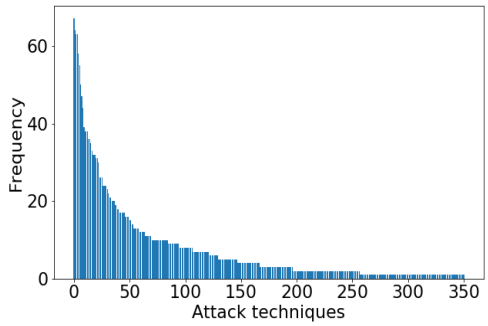
4. 데이터셋 구축

국가가 지원하는 것으로 추정되는 APT 공격의 경우에는 공격을 은폐하기 위한 기만 기술과 안티포렌식 기술이 적용된 경우가 많아 관련 데이터를 모으는 것이 어렵다. 또한 공격 횟수가 많지 않기 때문에, 학습에 필요한 충분한 데이터양을 모으기도 어렵다. 우리는 공격 그룹의 학습에 이용될 데이터의 가장 중요한 요건으로 신뢰성을 선정하였다. 데이터양이 적기 때문에 작은 에러도 큰 영향을 미칠 수 있기 때문이다. 조사해본 결과 공개된 데이터 중 가장 신뢰도가 높다고 판단되는 MITRE ATT&CK의 공격 그룹 데이터를 학습 데이터로 사용하였다. 테스트에 이용될 데이터는 학습에 이용될 데이터보다 신뢰도가 낮아도 양이 많은 것을 중시하였다. 우리는 APT 공격

분석 보고서를 자동으로 라벨링 해주는 기술인 rcATT [31]를 사용하여 테스트 데이터를 구축하였다.

4.1 MITRE ATT&CK Group

MITRE ATT&CK에서는 사이버 공격 그룹에 대한 설명, 사용하였던 공격 기술, 사용한 소프트웨어 등을 조사하여 공개하고 있다. 버전 10 기준으로 총 127개의 그룹이 조사되었으며, 등장하는 공격 기술의 종류는 351개, 총 개수는 2,743개이다. 그룹당 평균적으로 기술된 공격 기술의 개수는 21.60개이고 사용한 공격 기술이 가장 많이 기록된 그룹은 "APT29"로 83개의 공격 기술을 사용하는 것으로 조사되었다.



(그림 2) 공격 그룹들이 사용하는 공격 기술 빈도
(Figure 2) Frequency of attack techniques used by attack groups

그림 2는 127개의 공격 그룹들이 사용하는 공격 기술의 빈도를 나타낸 것이다. 가장 많이 사용되는 공격 기술은 "T1204.002: User Execution-Malicious File"로 67개의 그룹에서 사용하였다. 절반 이상의 그룹이 사용된 것으로 조사된 이 공격 기술은 ".exe", ".pdf", ".doc"등의 확장자를 가진 악성코드를 실행하는 행위이며, 피싱 메일에 첨부된 것을 실행하는 경우가 대부분이다. 두 번째로 많이 사용한 것은 "T1566.001: Phishing: Spearphishing Attachment"로 피싱 메일의 첨부파일을 붙이는 기술이다. 이는 64개의 공격 그룹에서 사용된 것으로 확인되었다. 반면 "T1615: Group Policy Discovery"는 "Turla"에서 피해 시스템에서 정보를 모을 때 사용하던 공격 기술로, 타 공격 그룹에서는 발견되지 않은 것이다.

사용된 대부분의 공격 기술은 기술 보고서나 기술 블로그 글 등을 인용하여, 신뢰도가 높은 편이다. 또한 검색해본 바로 공개된 데이터 중에서는 가장 조사된 그룹 수가 많았다. 따라서 본 논문에서는 이 데이터를 정답(ground-truth)으로 간주하고 공격 기술 영향성 점수 및 그룹 유사도 점수를 산출할 때 이용하였다. 표 4는 본 논문에서 사용한 데이터의 샘플을 나타낸 것이다. 여기에는 하위 공격 기술도 포함되어 있지만, 본 논문에서는 하위 공격 기술을 공격 기술로 추상화하여 사용하였다.

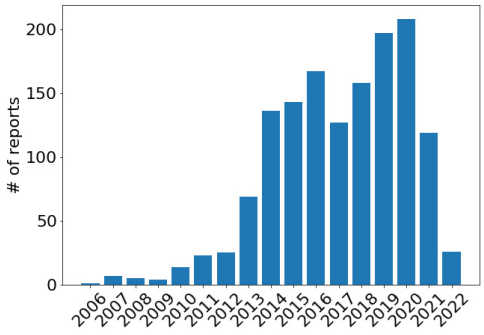
(표 4) 논문에서 사용한 각 공격 그룹의 공격 기술 샘플
(Table 4) A sample of attack techniques for each attack group in this paper

공격 그룹	공격 기술
APT29	T1547.001, T1203, T1546.008, T1548.002, …
Turla	T1110, T1566.002, T1012, T1059.001, …
Patchwork	T1083, T1082, T1055.012, T1559.002, T1560,…

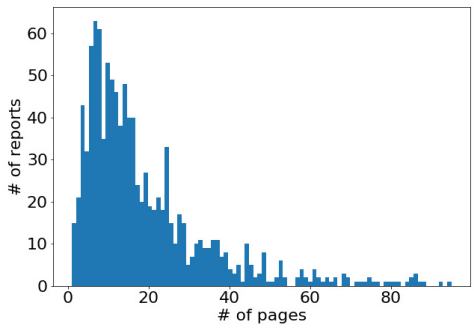
4.2 Cyber Criminals Campaign Collections

사이버 공격이 탐지되면, 여러 보안 업체들이 위협 분석 보고서를 발간하거나, 블로그 등에 분석 결과를 공유한다. 발행처가 다양하므로, 이를 한데 모은 저장소들이 운영되고 있다[26-28]. 이 중에서 가장 체계적이고, 양이 많은 저장소는 Cyber Criminal Campaign Collection[26] 이다. 현재까지 약 1,400여 개의 보고서 및 블로그 글이 저장되어 있으며 2006년도 보고서부터 존재하고, 매달 새로운 보고서가 저장된다. 그림 3은 수집된 연도별 보고서의 개수를 나타낸 것이다.

모든 문서가 사이버 공격 캠페인을 묘사한 것은 아니며, 그룹이나, 국가, 악성코드 등 다양한 주제로 작성한 보고서가 수집되고 있다. 정확하게 일치하는 것은 아니지만 페이지 수가 적으면 블로그 글이나 단편적인 공격 기술에 대한 설명일 확률이 높고, 페이지 수가 많으면 본 논문에서 테스트에 이용할 수 있는 사이버 공격 캠페인에 대한 것일 확률이 높다. 그림 4는 문서의 페이지 수 분포를 나타낸 것이다. 100페이지보다 많은 수의 보고서 8건은 가독성을 위해 그림 4에서는 제외하였다. 우리는 이 중에서 캠페인으로 판별된 보고서만을 추려 테스트에 사용하였다.



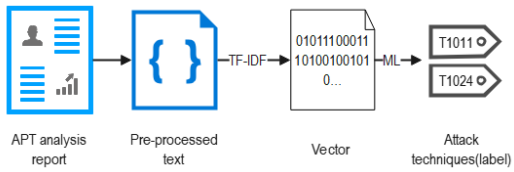
(그림 3) 연도별 보고서 개수
(Figure 3) Number of reports by year



(그림 4) 페이지 수 별 보고서 개수
(Figure 4) Number of reports by number of pages

4.3 공격 기술 추출

표 4에서 학습에 사용한 MITRE ATT&CK의 공격 기술을 활용하기 위해서는 테스트에 사용할 데이터에도 마찬가지로 MITRE ATT&CK의 공격 기술이 포함되어야 한다. 가장 좋은 방안은 사이버 공격에 정통한 보안 전문가가 보고서를 읽고 MITRE ATT&CK의 공격 기술을 라벨링하는 것이다. 하지만 1,400여개의 보고서를 전부 수동으로 읽고 라벨링하는 것은 큰 비용과 시간이 들어간다. 이러한 수고를 줄이기 위해서 보고서를 대상으로 자동으로 공격 기술을 라벨링하는 연구들이 진행됐다[29-31]. 우리는 전체과정의 자동화가 가능한 Legoy 연구진의 방법 rcATT(Report Classification by Adversarial Tactics and Techniques)[31]을 사용하였다. 그림 5와 같이 이들은 보고서나 블로그 글을 텍스트로 변환하고, 불필요한 문자를 제거하여 전처리를 진행하였다. 그 후 언어처리에서 사용되



(그림 5) rcATT 동작 과정
(Figure 5) Overall process of rcATT

는 TF-IDF(Term frequency - inverse document frequency) 기술[32]을 이용하여 문서를 벡터화하였다.

이들은 Scikit-learn 라이브러리[33]에서 제공하는 멀티 라벨 분류기를 이용하여 공격 기술과 전술을 라벨링 하였다. 해당 논문은 2019년도에 발표된 것으로, 이들은 실험에 MITRE ATT&CK 버전 5를 사용한 것으로 추정된다. 본 논문에서는 하위 공격 기술이 존재하는 버전 10을 기준으로 재학습하여 공격 기술을 추출하였다.

예로 아래 문장은 공격자가 호스트에 에이전트를 설치하고 관리 서버를 이용한다는 내용이다. rcATT 알고리즘은 "T1059:Command-Line Interface"로 알맞게 라벨링 하였다.

> ... To bypass this process, the attacker pretends to be the management server and sends a commend to the agent ...

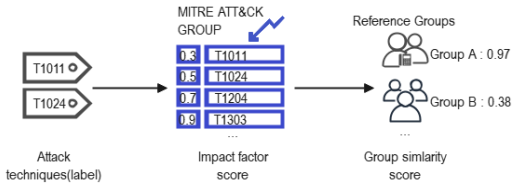
표 5는 추출한 공격 기술과 전술의 예제를 나타낸 것이다. MITRE ATT&CK의 공격 그룹에서는 전술 정보를 제공하지 않아서, 본 논문에서도 rcATT로부터 추출한 정보의 전술 정보 및 하위 기술을 제외하였다.

(표 5) rcATT를 이용하여 추출한 공격 기술, 전술
(Table 5) Attack techniques and tactics extracted using rcATT

보고서	전술
	공격 기술
TinyPos	TA0005, TA0009
	T1560, T1059, T1140, T1070, T1070
Fin8 Sardonic	TA0002, TA0003, TA0004, TA0005, ...
	T1134, T1071, T1560, T1573, ...
Sharp Panda	TA0001, TA0002, TA0003, ...
	T1071, T1560, T1547, T1059, T1543, ...

5. 제안하는 방법

제안하는 방법의 전체적인 흐름은 그림 6과 같다. 임의의 사이버 공격 캠페인이 입력으로 주어지면, 라벨링된 MITRE ATT&CK의 공격 기술들에 대하여 영향성 점수를 적용한다. 그 후 영향성이 반영된 공격 기술들을 제시하는 유사도 점수를 이용하여 산출하고, 이 중 가장 점수가 높은 공격 그룹을 선정한다. 5.1 절에서는 제안하는 방법의 근거를 얻기 위한 분석을 서술하고, 5.2절에서는 영향성 점수, 5.3절에서는 공격 그룹 유사도 점수를 산출하였다.



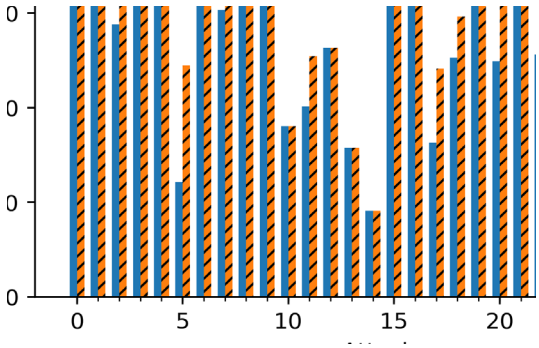
(그림 6) 제안하는 방법의 흐름도
(Figure 6) Overall process of the proposed method

5.1 사이버 공격 그룹의 공격 기술 분석

본 논문에서는 국가가 지원하는 것으로 추정되는 사이버 공격 그룹의 경우에는 그들이 사용하는 공격 기술의 종류가 추후 크게 변화하지 않을 것이라 가정하고 연구를 진행하였다. 이와 같은 가정을 검증하기 위해서, 4.1장에서 구축한 그룹 데이터와 4.2장과 4.3장에서 구축한 공격 캠페인 데이터의 연관 관계를 분석할 필요가 있다.

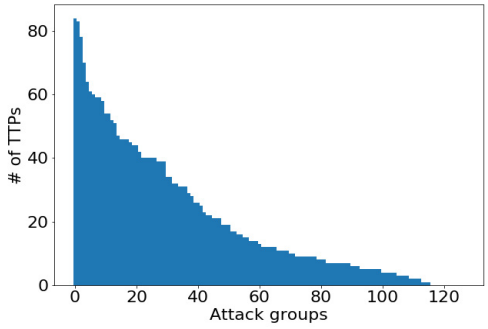
공격 그룹이 사용하는 공격 기술의 종류가 연도 별로 크게 변화하지 않음을 검증하기 위해, 우선 각 사이버 공격 그룹이 과거에 사용하였던 공격 기술을 재사용하였는지 분석하였다. 신뢰도를 위해 최소 3개의 공격 기술 및 3개 이상의 분석 보고서가 발행된 데이터 총 315개에 대하여 분석하였다.

각 그룹에 대하여, 공격 연도마다 직전 연도의 공격 기술을 재사용하였는지 분석하였다. 또한 이전 모든 연도에 사용되었던 공격 기술을 재사용하였는지도 조사하였다. 그림 7은 이를 각 그룹에 대해서 그래프로 나타낸 것이다. 직전년도 재사용률 그룹 평균은 53.51%이고, 과거 모든 연도(2006년~2022년) 그룹 재사용률 그룹 평균은 65.49%로 분석되었다. 이 중 가장 최근인 2022년만 살펴보면 그룹 평균은 75.21%로 많은 공격 기술 방법을 재사용하는 것을 알 수 있었다. 실제로 공격 모듈을 다시 만드는 것은 비용과 시간이 많이 드는 작업으로 악성코드나 스크립트를 재사용하는 사례가 종종 있다.



(그림 7) 각 공격 그룹의 과거 공격 기술 재사용률
(Figure 7) Reuse rate of attack techniques for each attack group

MITRE ATT&CK에서는 3장에서 설명한 것과 같이 127개의 공격 그룹을 식별하고 이들이 과거 사용했던 공격 기술들을 명시하였다. 그림 8는 각 공격 그룹이 사용하였던 공격 기술의 개수를 나타낸 것이다. 정보 부족, 혹은 활동 저조 등의 이유로 공격 기술의 개수가 적은 그룹이 존재하는데, 10개 미만의 공격 기술이 명시된 그룹 49개는 정보 부족으로 인하여 분석이 어렵다고 판단되어 본 논문에서는 제외하였다.



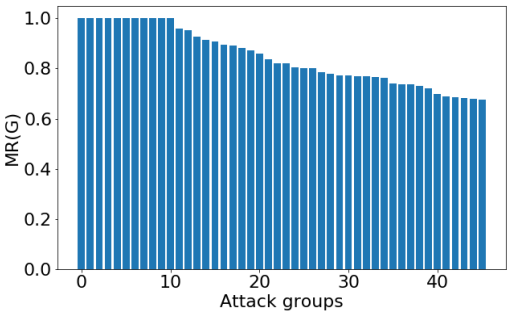
(그림 8) 공격 그룹들이 사용하는 공격 기술의 개수
(Figure 8) Number of attack techniques used by attack groups

MITRE ATT&CK이 분석한 공격 그룹의 공격 기술이 4.2장에서 언급한 분석 보고서들의 공격 기술에 포함되어

있는지 확인하기 위해서 수식 1과 같이 그룹 G에 대해서 일치율 MR(G)를 정의하였다. y는 해당 공격 그룹이 수행한 캠페인에 대한 공격 분석 보고서이다. A(G)는 MITRE ATT&CK에서 그룹 G에 대해서 분석한 공격 기술이며, R(y)는 보고서 y에 대해서 rcATT 알고리즘을 이용하여 추출한 공격 기술이다.

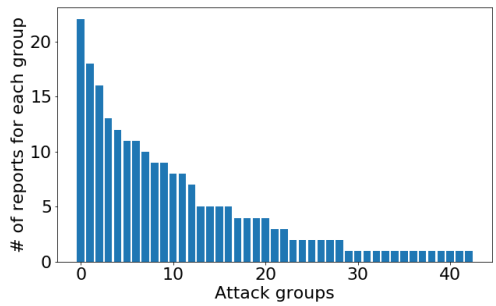
$$MR(G) = \sum_{y \in G} \frac{n(A(G) \cap R(y))}{n(R(y))} / n(G) \quad \text{(수식 1)}$$

그림 9는 각 그룹의 공격 기술과 해당하는 보고서의 공격 기술의 일치율 MR(G)를 나타낸 것이다. 일치율이 평균 73.2%로 MITRE에서도 침해 분석 보고서를 상당 부분 참고하였을 것으로 추측된다. 다만 머신러닝 기반의 rcATT 알고리즘이 완벽하게 공격 기술을 추출하는 것은 불가능하다. 따라서, 정확한 분석을 위해 65% 미만의 일치율을 보이는 그룹은 공격 기술 추출에 실패하였다고 간주하여 제외하였다.

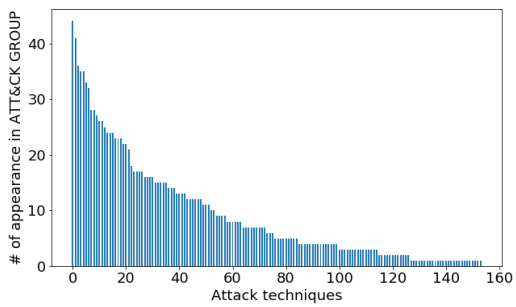


(그림 9) 각 공격 그룹 별 일치율 MR(G)
(Figure 9) Match rate for each attack group

본 논문에서는 rcATT를 이용하여 공격 기술을 라벨링하고 있어, 하위 공격 기술 대신에 공격 기술을 사용하였다. 예를 들면 "T1548.001"은 "T1548"으로 변경하여 사용하였다. 데이터 셋에는 공격 분석 보고서 뿐만 아니라 짤막한 블로그 글이나 게시글의 형태도 있으며 이런 경우 공격 기술에 대한 언급이 없거나 rcATT 알고리즘이 공격 기술을 검출하지 못하는 경우가 있었다. 이러한 분석 보고서를 제외하기 위해 공격 기술이 10개 초과 라벨링 된 경우만 고려하였다. 최종적으로 분류에 유효한 공격 그룹 총 46개, 공격 분석 보고서를 222편 식별하였다. 그림 10은 최종적으로 분석 및 선별을 마친 후 각 그룹에 대해 발행된 보고서의 편수를 나타낸 것이다. 평균적으로는 하나의 그룹당 4.83편의 보고서가 발행되었다. 가장 많은 수의 보고서가 발행된 공격 그룹은 Lazarus Group이며, 총 22편이 발행되었다.



(그림 10) 공격 그룹당 보고서 발행 편수
(Figure 10) Number of reports for each attack group



(그림 11) 각 공격 기술의 공격 그룹 등장 빈도수
(Figure 11) Number of attack group appearances for each attack technique

보고서의 공격 기술이 72.1%의 높은 확률로 MITRE ATT&CK Group의 공격 기술에 포함되기는 하지만, 이 사실이 공격 그룹 분류 성능을 보장하지는 않는다. 공격 그룹들이 즐겨 사용하는 공통적인 공격 기술이 다수 존재하기 때문이다. 그림 11은 각 공격 기술이 MITRE ATT&CK Group에 포함된 빈도수를 나타낸 것이다. 하나의 공격 기술은 평균 9.30개의 공격 그룹에 등장하였으며, 이것은 하나의 그룹에 20.21%의 확률로 같은 공격 기술이 등장한다는 뜻이다. 더욱이 가장 많이 등장하는 공격 기술은 분석 대상 그룹 46개 중 44개에서 등장하여 해당 공격 기술로는 그룹을 분류할 수 없는 문제가 발생하였

다. 이 문제를 해결하기 위해서 우리는 5.2장에서 공격 기술 영향성 점수를 제안하였다.

5.2 공격 기술 영향성 점수

5.1장에서 분석한 것과 같이 사이버 공격 그룹들이 공통적으로 수행해야 하는 기술들이 존재한다. 표 6은 공격 그룹이 자주 사용한 기술 상위 5개를 나타낸 것이다. 가장 자주 등장하는 "T1059 : Command and Scripting Interpreter"는 최근 대부분의 사이버 공격에서 기본적으로 사용되고 있는 "powershell", "windows command shell", "python", "java script", "visual basic"등을 사용한 스크립트 기술이다. 분석 대상 그룹 46개 중의 44개 그룹에 등장하고 있어, 분류에 도움을 주지 못한다.

(표 6) 공격 그룹들이 자주 사용한 기술(Top-5)
(Table 6) Techniques frequently used by attack groups(Top-5)

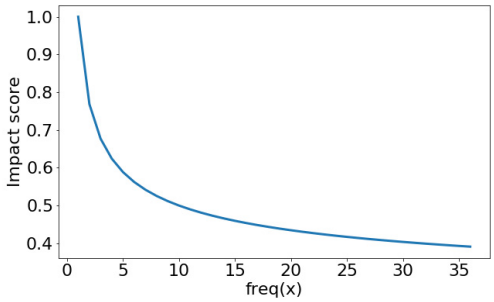
등장 기술	T1059	T1027	T1105	T1204	T1566
등장 횟수	44	41	36	35	35

이러한 특성을 고려해봤을 때 자주 등장하는 공격 기술의 경우에는 영향성을 낮추고, 드물게 등장하는 공격 기술은 영향성을 높이는 것이 합리적이다. 이러한 특성을 반영할 수 있도록 수식 2와 같이 영향성 점수(impact score) IM(x)를 설계하였다. x는 공격 기술, *freq(x)*는 공격 그룹에서 사용한 공격 기술의 빈도수이다. α값을 이용하여 무시할 수 있는 공격 기술의 빈도수를 조절할 수 있도록 설계하였다. 그림 12는 공격 기술 등장 빈도에 따른 영향성 점수를 그래프로 나타낸 것이다.

$$IM(x)=\begin{cases}\dfrac{1}{1+\log 10(freq(x))} & ,freq(x)\le\alpha\\ 0 & ,freq(x)>\alpha\end{cases} \quad \text{(수식 2)}$$

5.3 그룹 유사도 점수

공격 기술의 영향성 점수를 기반으로 보고서 y와 그룹 G에 대한 공격 그룹 유사도 점수를 수식 3과 같이 설계하였다. MITRE ATT&CK에서 공격 그룹에 라벨링한 공격 기술의 개수는 그림 8에서 보는 것과 공격 그룹마다 다르다. 따라서 수식 3에서와 같이 공격 그룹의 공격 기술 개수를 이용하여 값을 조정해주었다. 기존 연구에서는 유사



(그림 12) freq(x)에 따른 공격 기술 영향성 점수
(Figure 12) Impact score of attack technique according to freq(x)

도 점수의 동점이 많이 나오는 단점이 있었으나, 본 연구에서는 공격 그룹의 개수로 나눔으로써 동점을 방지할 수 있었다.

$$SIM(y,G)=\sum_{x} IM(x)/\log(length(G)),\quad \forall x\in A(G)\cap R(y) \quad \text{(수식 3)}$$

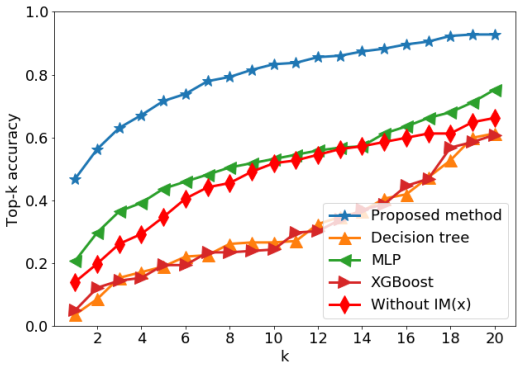
제안하는 그룹 유사도 점수는 값이 클수록 공격 분석 보고서의 그룹이 MITRE ATT&CK의 그룹과 유사하다는 것을 의미한다.

6. 실험 결과

우리가 실험을 수행한 환경은 인텔 i9-7900X, 128GB, 엔비디아 GeForce RTX 3090, 윈도우 10이며, 개발 언어는 python 3.8을 이용하였다.

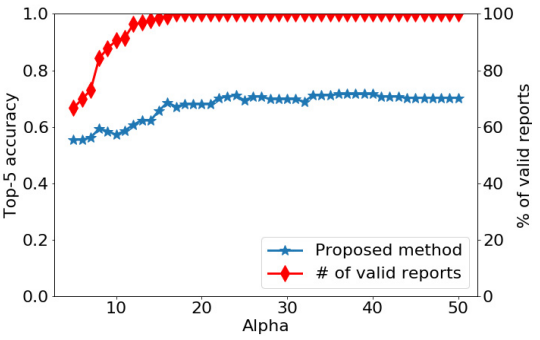
통상적인 머신러닝 및 딥러닝 알고리즘과의 성능 차이를 비교하기 위해 우리는 4-gram을 이용하여 공격 기술의 특징점을 추출하고, 머신러닝의 일종인 DT(Decision Tree), MLP(Multi Layer Perceptron), XGBoost[34]를 이용하여 분류를 진행하였다. 또한 최저 한계치를 분석하기 위해 영향성 점수 IM(x)를 사용하지 않고 그룹 유사도 점수를 산출하였다. 그림 13은 제안하는 방법과 비교군들의 top-k 정확도를 나타낸 것이다. 판별 가능한 총 그룹의 개수가 46개이므로 top-1일 때 랜덤 기댓값은 1/46 = 2.17%이다. 영향성 점수를 고려하지 않은 경우의 top-5 정확도는 34.68%이고, DT는 18.92%, XGBoost 는 19.37%, MLP는 43.69%의 낮은 성능을 보였다. 이러한 낮은 성능은 학

습 데이터가 하나의 라벨당 하나밖에 존재하지 않는 극단적인 학습 환경에 기인한다. 이에 비해 제안하는 방법의 top-5 정확도는 72.62%으로 타 알고리즘과 비교해 월등히 우수한 성능을 보여주었다.



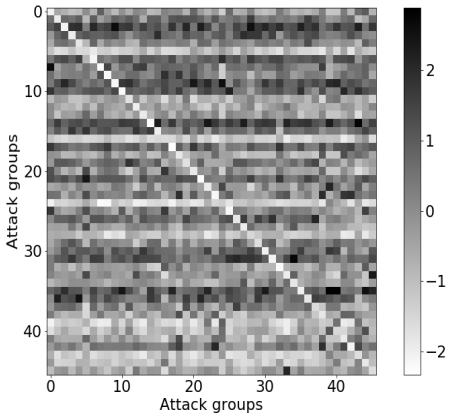
(그림 13) k값에 따른 top-k 정확도
(Figure 13) Top-k accuracy according to the k

그림 14은 α값에 따른 유효보고서의 수를 나타낸 것이다. α값이 너무 적으면, 대부분의 값이 0 이 되어 결과가 유효하지 않다. 본 논문에서는 모든 그룹에 대해서 유사도 값이 0 이 나올 때는 해당 테스트 보고서가 유효하지 않다고 판단하였다. 실험 결과 α값이 20 이상이면 대부분의 보고서가 유효하여 그룹 분류가 가능하였다. 정확도의 경우 근소하게 차이가 있는데, α값은 40 부근이 최적값으로 보여 이 값을 사용하였다.



(그림 14) α값에 따른 유효 보고서 개수 및 정확도
(Figure 14) Number of valid reports and accuracy according to the α

그림 15는 MITRE ATT&CK 공격 그룹 간 유사도를 혼동 행렬(confusion matrix)로 시각화 한 것이다. 대각선의 경우 같은 그룹이므로 값이 가장 크지만, 가시화를 위해서 가장 작은 수로 설정하고 각 행에 대해서 정규화를 진행하였다. 이를 통하여 공격 그룹 간의 유사도를 분석하는 것이 가능하다.



(그림 15) 각 그룹 간 유사도
(Figure 15) Similarity for each groups

7. 결 론

APT 공격을 수행하는 공격 그룹에 대한 정보는 매우 한정적이고, 대부분 수동으로 전문가들이 분석하는 경우가 많다. 최근에는 이를 자동으로 분석하려는 연구들이 진행되었는데, 정확도가 낮거나 적은 수의 공격 그룹만을 분석하는 단점이 있었다. 본 논문에서는 MITRE ATT&CK에서 분석한 공격 그룹의 공격 기술을 토대로 전처리 및 영향성 점수, 그룹 유사도 점수를 제안하였다. 실험 결과 46개의 공격 그룹에 대해 top-5 기준 72.62%의 높은 정확도로 분류할 수 있음을 알 수 있었다.

앞으로 정확도를 높이기 위해, 소량의 데이터에 적합한 머신러닝 기법을 도입하고, 양질의 데이터를 확보하기 위해 사이버 공격 분석 보고서 외의 자료를 탐색하여 데이터를 보강하는 연구를 진행할 예정이다. 또한 보고서로부터 rcATT를 이용하여 추출한 라벨의 정확도가 본 연구의 정확도에 영향을 끼치는 만큼, 추후 라벨링 정확도를 높이기 위한 알고리즘 연구도 필요하다.

향후 본 논문에서 제안하는 모델 기반 그룹 분류 방법

을 회피하려고 일부러 공격 기술 방법을 달리하거나 타 공격 그룹의 공격 기술을 모사하는 APT 공격 그룹이 나타날 가능성이 크다. 이러한 문제를 해결하기 위해서 시그니처 기반 기술, 포렌식 기반 아티팩트 분석 등 단위 기술로부터 추출 가능한 정보들을 통합하여 공격 그룹을 분류하는 기술이 연구된다면 국방 측면에서 사이버 공격 그룹 식별을 하는 데 큰 도움이 될 수 있을 것으로 기대된다.

참고문헌(Reference)

[1] Liu, D., Zhang, H., Yu, H., Liu, X., Zhao, Y., Lv, G., "Research and application of APT attack defense and detection technology based on big data technology", Proceedings of IEEE 9th International Conference on Electronics Information and Emergency Communication, pp. 1-4, 2019. https://doi.org/10.1109/ICEIEC.2019.8784483

[2] Choi, C. H., Shin, C. H., Shin, S. U., Seo, S. Y., Lee, I. S., "Deep learning for estimating next action of cyber attack", Proceedings of Korea Institute of Military Science and Technology annual conference, pp. 1075-1076, 2021.

[3] Choi, C. H., Shin, S. U., Shin, C. H., "Performance evaluation method of cyber attack behaviour forecasting based on mitigation", Proceedings of International Conference on information and communication Technology Convergence, pp. 13-15, 2021. https://doi.org/10.1109/ICTC52510.2021.9620951

[4] Choi, C. H., Shin, C. H., Shin, S. U., "Cyber attack group classification based on TTP information", Proceedings of Internet Computing and Service spring conference, vol. 23, no. 1, pp. 7-8, 2021.

[5] Al-Mohannadi, H., Mirza, Q., Namanya, A., Awan, I., Cullen, A., Disso, J., "Cyber-attack modeling analysis techniques:An overview", Proceedings of IEEE 4th international conference on future internet of things and cloud workshops, pp. 69-76, 2016. https://doi.org/10.1109/W-FiCloud.2016.29

[6] Hutchins, E. M. Cloppert, M. J., and Amin, R., M. "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chain", Journal of Leading Issues in Information Warfare & Security Research, vol. 1 no. 1, pp. 80, 2011. https://lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf

[7] Kim, H., Kwon, H. J., and Kim, K. K., "Modified cyber kill chain model for multimedia service environments", Journal of Multimedia Tools and Application, vol .78 no. 3, pp. 3153-3170, 2019. https://doi.org/10.1007/s11042-018-5897-5

[8] MITRE ATT&CK, https://attack.mitre.org

[9] M. Gül and E. Kugu, "A Survey on anti-forensics techniques,", Proceedings of International Artificial Intelligence and Data Processing Symposium, pp. 1-6, 2017. https://doi.org/10.1109/IDAP.2017.8090341

[10] Kawai, M., Ota, K., and Dong, M., "Improved malgan: Avoiding malware detector by leaning cleanware features", Proceedings of IEEE International Conference on Artificial Intelligence in Information and Communication, pp. 40-45, 2019. https://doi.org/10.1109/ICAIIC.2019.8669079

[11] Hwang, C. W., Kim, D. Y., and Lee, T. J., "Semi-supervised based unknown attack detection in EDR environment", Transactions on Internet and Information Systems. vol. 14, no. 12, pp. 4909-4926, 2020. https://doi.org/10.3837/tiis.2020.12.016

[12] Milajerdi, S. M., Gjomemo, R., Eshete, B., Sekar, R., and Venkatakrishnan, V. N., "Holmes: real-time apt detection through correlation of suspicious information flows.", Proceedings of IEEE Symposium on Security and Privacy, pp. 1137-115. 2019. https://doi.org/10.1109/SP.2019.00026

[13] Watters, P., McCombie, S., Layton, R., and Pieprzyk J., "Characterising and predicting cyber attacks using the cyber attacker model profile(CAMP)", Journal of Money Laundering Control, vol. 15, pp. 430-441, 2012. https://doi.org/10.1108/13685201211266015

[14] Kapetanakis, S., Filippoupolitis, A., Loukas, G., and Murayziq, T., "Profiling cyber attackers using case-based reasoning", Proceedings of 19th UK workshop on case-based reasoning, pp. 39-48, 2014. https://researchgate.net/publication/301221761_Profiling_cyber_attackers_using_Case-based_Reasoning

[15] Stahl, A., and Roth-Berghofer, T., “Rapid prototyping of CBR Applications with the Open Source Tool my CBR”, Proceedings of the 9th European Conference on Advances in Case-Based Reasoning, pp. 615-629, 2008. https://doi.org/10.1007/978-3-540-85502-6_42

[16] Cho, H. S., Lee, S. G., Kim, B. I., Shin, Y. S., and Lee, T. J., “The study of prediction of same attack group by comparing similarity of domain”, Proceedings of International conference on information and communication technology convergence, pp. 1220-1222, 2015. https://doi.org/10.1109/ICTC.2015.7354779

[17] Han, M. L., Han, H. Ch., Kang, A. R., Kwak, B. I., Mohaisen, A., and Kim H. K., “WHAP: Web-hacking profiling using case-based reasoning”, Proceedings of IEEE Conference on Communication and Network Security pp., 344-345, 2016. https://doi.org/10.1109/CNS.2016.7860503

[18] Kim, W. J., Park, C. W., Lee, S. J., and Lim J. S., “Methods for Classification and Attack Prediction of Attack Groups based on Framework of Cyber Defense Operations”, Journal of KIISE:Computing Practices and Letters. vol. 20, no.6, pp.317-328, 2014. http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02432562

[19] Choi, C. H., Lee, H. S., Jung, I. H., Yoo, C. G., and Yoon, H. S., “Statistical Analysis of EML Header for Cyber Attacker Tracing”, Proceedings of Korea Institute of Military Science and Technology annual conference, pp.1141-1142, 2017.

[20] Choi, C. H., Lee, H. S., Jung, I. H., Park, J. H., and Yoon, H. S.,“E-mail Clustering for Cyber Attack Attribution”, Proceedings of Korea Institute of Military Science and Technology annual conference, pp.1289-1290, 2018.

[21] Jung, I. H., Lee, H. S, Choi, C. H., Yoo, C. G., and Yoon, H. S., “A Study for Specific information identification of attackers through document type malware analysis”, Proceedings of Korea Institute of Military Science and Technology annual conference, pp.1185-1186, 2017.

[22] Jung, I. H., Lee, H. S, Choi, C. H., and Yoon, H. S., “A Study for Creator System Information Identification Based on Document Type Malware”, Proceedings of Korea Institute of Military Science and Technology annual conference, pp.1504-1505, 2018

[23] Son, K. H., Kim, B. I., and Lee, T. J., “Cyber-attack group analysis method based on association of cyber-attack information”, Transaction on Internet and Information Systems, vol. 14, no. 1, pp.260-280, 2020. https://doi.org/10.3837/tiis.2020.01.015

[24] Shin, Y. S., Kim, K. M., Lee, J., Lee, K. H., “ART: Automated reclassification for threat actors based on ATT&CK matrix similarity”, Proceedings of World Automation Congress, pp.15-20, 2021. https://doi.org/10.23919/WAC50355.2021.9559514

[25] Choi, C. H., Shin, C. H., Shin, S. U., Seo, S. Y., Lee, I. S., “Cyber Attack Group Classification using Siamese LSTM”, Proceedings of Korea Institute of Military Science and Technology annual conference, pp. 1425-1426, 2022.

[26] APT & CyberCriminal Campaign Collections, https://github.com/CyberMonitor/APT_CyberCriminal_Campagin_Collections

[27] APTNotes, https://github.com/aptnotes/data,

[28] APT report collected blackorbird, https://github.com/blackorbird/APT_REPORT

[29] Husari, G., Al-Shaer, E., Ahmed, M., Chu, B., and Niu, X., “TTPDrill: Automatic and accurate extraction of threat actions from unstructured text of CTI Sources”, 33rd annual computer security applications conference, pp. 103-115, 2017. https://doi.org/10.1145/3134600.3134646

[30] Threat Report ATT&CK Mapping(TRAM), https://github.com/center-for-threat-informed-defense/tram/

[31] Legoy, V., Caselli, M., Seifert, C., and Peter, A, “Automated retrieval of ATT&CK tactics and techniques for cyber threat reports.”, arXiv preprint arXiv:2004.14322, 2020. https://doi.org/10.48550/arXiv.2004.14322

[32] Mikolov, T., Chen, K., Corrado, G., and Dean, J., “Efficient estimation of word representations in vector space”, arXiv preprint arXiv:1301.3781, 2013. https://doi.org/10.48550/arXiv.1301.3781

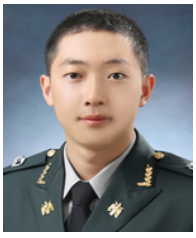
[33] Scikit-learn, https://scikit-learn.org

[34] XGBoost, https://github.com/dmlc/xgboost

◐ 저 자 소 개 ◑

최 창 희(Changhee Choi)
2008년 연세대학교 컴퓨터과학과(공학사)
2010년 한국과학기술원 대학원 전산학과(공학석사)
2013년 한국과학기술원 대학원 전산학과(공학박사)
2013년~현재 국방과학연구소 연구원
관심분야 : 머신러닝 기반 사이버 보안, AI, GAN, 디지털 포렌식
E-mail : changhee84@add.re.kr

신 찬 호(ChanHo Shin)
2018년 고려대학교 사이버국방학과(학사)
2018년~현재 국방과학연구소 현역연구원
관심분야 : 정보보호, 인공지능
E-mail : shinch2018@add.re.kr



신 성 욱(Sunguk Shin)
2017년 고려대학교 사이버국방학과(학사)
2017년~현재 국방과학연구소 현역연구원
관심분야 : 정보보호, 딥러닝, 강화학습
E-mail : ssw1419@add.re.kr