ATT&CK를 활용한 GloVe 기반의 TTP 임베딩 기법
Towards Glove-Based TTP Embedding With ATT&CK Framework

신찬호* · 이인섭* · 최창희*
Chanho Shin* · Insup Lee* · Changhee Choi*

* 국방과학연구소
(shinch2018@add.re.kr)

ABSTRACT

Since MITRE published the ATT&CK matrix, many researchers have utilized the TTP information in their threat analysis reports and research, especially deep learning models. Since TTP information is a string written in natural language, embedding is essential to input it into the model. We propose the method for embedding TTP using GloVe, and the method for evaluating it, called “Tactic Match Rate”. The proposed embedding method showed a tactic match rate of 0.78 for the “Command&Control” tactic and 0.06 for the “Persistence” tactic, indicating different embedding tendencies depending on the tactic.

Key Words : TTP, Embedding, GloVe, MITRE ATT&CK

1. 서론

MITRE에서 ATT&CK®[1]을 발행한 이후로 많은 연구자가 이를 활용하여 분석 보고서를 작성하고 있다. 이는 TTP 정보만으로도 공격의 흐름을 파악하여 분석 보고서를 더욱 쉽게 이해하고, 쉽게 공유할 수 있게 만들었다. 또한 여러 분야에서 딥러닝 모델이 괄목할만한 성능을 보임에 따라서, TTP 정보를 이용하여 다음 공격[2]을 예측하는 등 다양한 연구가 수행되고 있다. 하지만 TTP 정보는 문자열 데이터이기 때문에 딥러닝 모델에 사용되기 위해서는 임베딩 과정이 필수적이다. 본 논문은 자연어처리 분야에서 널리 쓰이는 GloVe[3]를 이용하여 TTP 정보를 임베딩하고, 그 임베딩 결과를 평가하는 방법인 “전술 일치율”을 제안한다. 실험 결과, “Command&Control” 전술에 대하여 평균 0.78, “Persistence” 전술에 대하여 평균 0.06의 전술 일치율을 보여, 전술에 따라 임베딩 경향성이 다름을 보였다.

2. 서론

2.1 MITRE ATT&CK

MITRE ATT&CK은 공격자의 전술(Tactic)과 기술(Technique)에 관해 설명한 표이다. 현재 버전 13.0까지 나왔으며, 14개의 전술, 196개의 기술, 411개의 세부기술로 이루어져 있다. 전술은 기술의 목적과 이유를 나타내며, 기술은 공격자가 목적을 달성하기 위해 어떻게 행동하는지에 대해 나타낸다. 전술별로 적게는 8개부터 많게는 42개까지의 기술이 포함되어있으며, 각 기술은 세부기술을 포함할 수 있다.

2.1 GloVe 임베딩

GloVe 임베딩은 동시 등장 행렬을 기반으로 고안된 단어 임베딩 방법론이다. GloVe 임베딩은 윈도우(Window) 기반의 동시 등장 행렬을 이용하여 단어의 통계 정보를 반영하고, 동시 등장 확률 기반의 손실함수를 이용으로 단어 간 유추 능력을 학습한다.

3. 제안하는 방법

3.1 데이터셋

데이터셋은 학습에 충분한 양을 확보하기 위하여 총 두 가지를 함께 사용하였다. 첫 번째는 MITRE에서 공개한 각 그룹과 소프트웨어가 많이 쓰는 TTP 정보이다. 각 그룹과 소프트웨어가 가지는 TTP 목록을 문서로 간주하고, 이들이 사용한 TTP를 단어로 간주하였다. 두 번째는 “CyberCriminal_Campaign_Collections(CCC)”[4] 보고서 데이터셋이다. CCC 데이터셋은 연도별로 수행된 공격에 대한 보고서를 정리한 데이터셋으로, 분석 전문가들이 각 보고서 내에 기술된 TTP 정보를 수동으로 라벨링 작업을 수행하였다.

이때, 세부기술 단위로 구성하면 TTP의 등장 빈도수 대비 TTP 수가 너무 낮아지는 문제점이 있다. 따라서 논문에서는 세부기술을 기술 단위로 변환하여 데이터셋을 구성하였다. 데이터셋에서 등장한 TTP는 총 180개, 사용된 데이터 수는 2103개이다.

Table 1. Dataset description

Target	TTP
APT1	T1087, T1583, T1560, ...
Thrip	T1059, T1048, T1588, ...
Carbanak	T1071, T1547, T1059, ...
ENISA_Threat...	T1566, T1190, T1189, ...

3.2 임베딩 학습 방법

GloVe를 이용하여 표 1의 데이터셋을 학습한다. GloVe는 먼저 다음 표 2와 같이 윈도우 기반의 동시 등장 행렬과 동시 등장 확률을 구성하고, 이를 기반으로 한 손실함수를 학습한다.

Table 2. Example of co-occurrence matrix

Count	T1087	T1071	T1560	T1189
T1087	0	1	2	0
T1071	1	0	1	1
T1560	2	1	0	0
T1189	0	1	0	0

3.3 임베딩 평가 방법: 전술 일치율

TTP 정보의 경우 자연어와 달리 문법적 특성이 없으므로 의미론적 성능만을 측정하게 된다. 기술의 의미론적 정보가 학습되었다면, 같은 전술에 속한 기술은 유사한 의미를 지닌다고 볼 수 있으므로 임베딩 값이 유사해야 한다. 논문에서는 기술의 임베딩 값과 가장 유사한 Top-10 기술을 뽑아, 해당 기술들이 같은 전술에 속하는 비율을 살펴보는 "전술 일치율" 평가 방법을 제안한다.

4. 실험 결과

그림 1은 전체 TTP에 대한 임베딩 성능이다. x축은 Top-10 비율을, y축은 각 비율에 속하는 TTP의 개수를 나타낸다. 같은 전술에 속하는 기술의 비율이 0.0인 기술은 총 41개로 가장 많고, 0.7을 제외한 0.1~0.9까지는 평균 15개로 고르게 분포해있다.

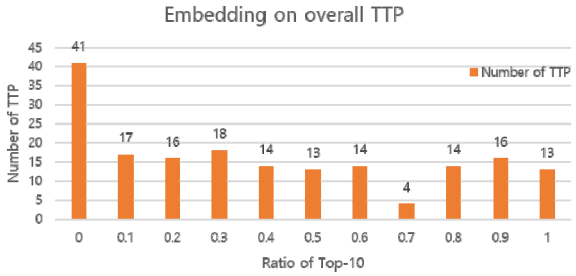

Figure 1. Embedding results of overall TTP techniques

하지만 그림 2와 그림 3에서 볼 수 있듯이 이는 전술별로 임베딩 성능이 판이하여 나타난 결과이다. 그림 2는 "Command&Control" 전술에 속하는 기술들의 Top-10 전술 일치율 결과로, 평균 0.78의 전술 일치율을 보인다. 반면 그림 3은 "Persistence" 전술에 속하는 기술들의 Top-10 전술 일치율은 평균 0.06으로 낮은 것을 볼 수 있다. 이 두 그림으로부터 전술별로 임베딩 성능이 크게 차이가 남을 알 수 있다. 이는 기술이 사용되는 목적에 따라 등장하는 빈도수와 같이 사용되는 기술의 종류에 따라 의존성이 크게 다르기 때문이다.

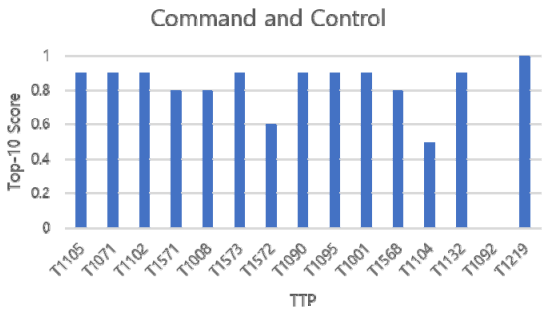

Figure 2. Embedding results of Command and Control Tactic

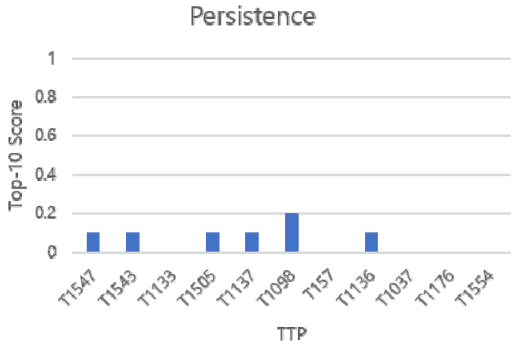

Figure 3. Embedding results of Persistence Tactic

5. 결론

논문에서는 GloVe 임베딩을 이용하여 TTP 정보를 임베딩하고, 전술 단위 유사도를 통해 그 성능을 측정하였다. 실험 결과 "Command&Control" 전술에서는 0.78, "Persistence" 전술에서는 0.06의 전술 일치율을 보여, 전술에 따라 임베딩 경향성이 다름을 알 수 있었다. 앞으로는 임베딩 성능이 저조한 전술과 기술을 분석하고, 보다 범용적이면서 성능이 높은 임베딩 방법론에 대한 연구를 진행할 예정이다.

References

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

[2] C. Choi, C. Shin, C. Shin, S. Seo, and I. Lee, "Deep learning for estimating next action of cyber attack", KIMST Annual Conference Proceedings, pp. 1075-1076, 2021.

[3] J. Pennington, R. Socher and C. Manning, "GloVe: Global Vectors for Word Representation", Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1532-1543, 2014.

[4] APT CyberCriminal Campaign collections, https://github.com/CyberMonitor/APT_CyberCriminal_Campagin_Collections