RIT | Global Cybersecurity Institute

# Translating Cybersecurity Descriptions into Interpretable MITRE Tactics using Transfer Learning

**Reza Fayyazi, Yiyuan Steve Wufeng ,Pradumna Gautam, Shanchieh Jay Yang**

## Introduction

- **Intrusion logs and threat intelligence reports have been developed to assist security analysts**
- **Description in these logs and reports, however, can be cryptic and not easy to interpret. Thus:**

**We ask:**

Given a description of cyberattack techniques, how to interpret the intended effects (MITRE Tactics [1])?

- **E.g.,1,** **Initialization scripts can be used to perform administrative functions, which may often execute other programs or send information to an internal logging server.**
- **E.g.,2,** **Custom Outlook forms can be created that will execute code when a specifically crafted email is sent.**

**Privilege Escalation? Persistence? Both?**

## Related Works

**PATRL (Pseudo-Active Transfer Learning) [2]**

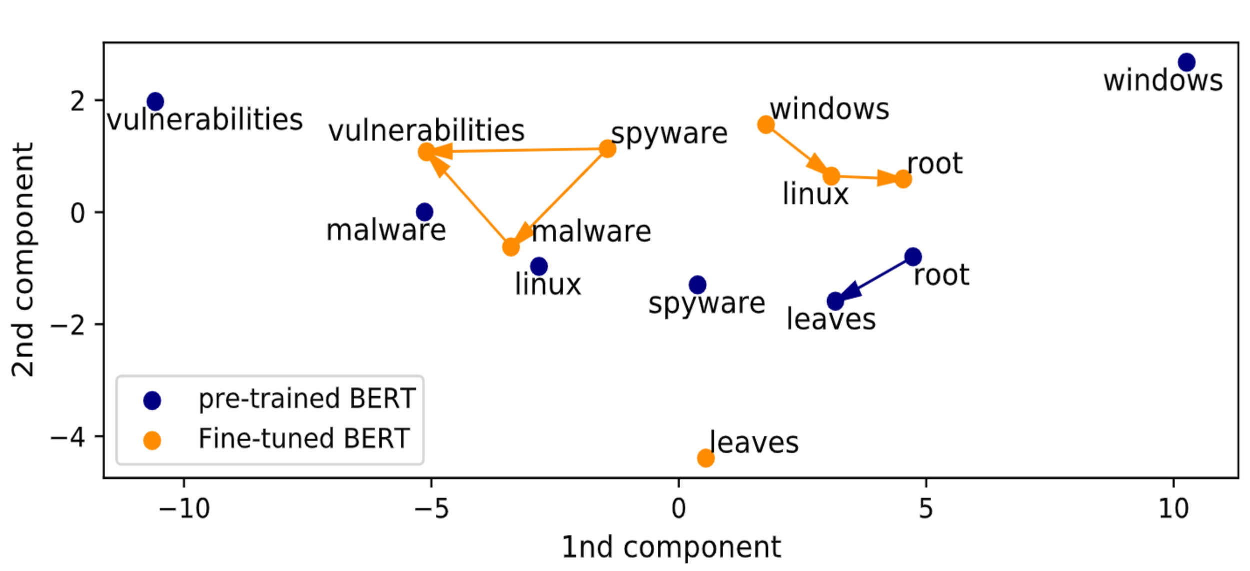
- **A semi-supervised process leveraging ULMFiT [3] to determine the attack stage of IDS alert signatures**

**BERT [4]**

- **A Transfer Learning technique to uncover the semantic information conveyed in a sentence**

**ExBERT [5]**

- **A framework that applies Transfer Learning to BERT to predict exploitability**
- **Word embedding for the pre-trained and fine-tuned BERT with cybersecurity words**



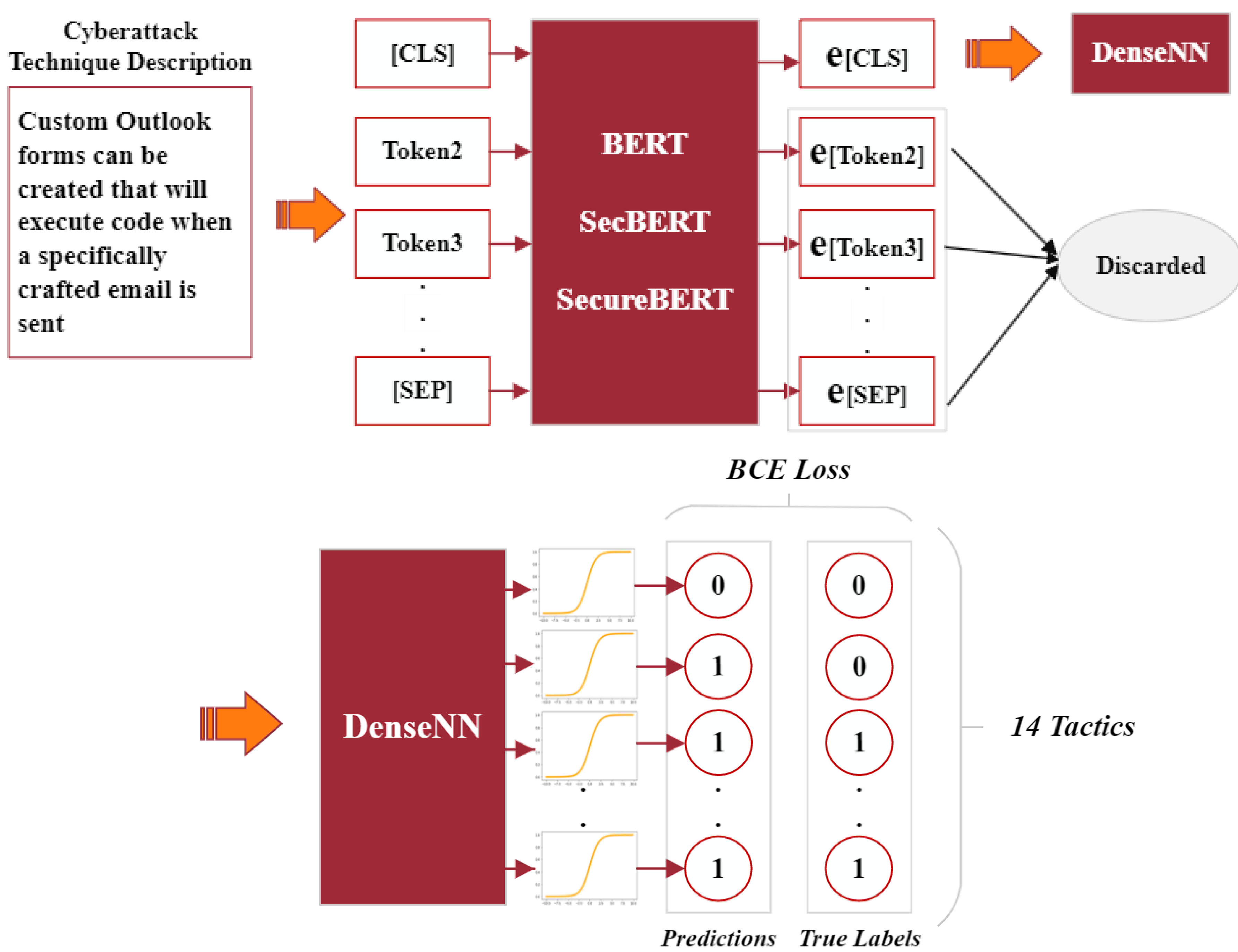

**SecBERT [6]**

- **A BERT model trained on cybersecurity texts**

**SecureBERT [7]**

- **A language model based on RoBERTa [8] that is trained on cybersecurity texts**

## Methodology




- **Multi-Label Classification for the total of 14 MITRE Tactics**
- **Total of 4500+ Descriptions with their corresponding tactic(s)**

| | PERSISTENCE | PRIV_ESC | DEF_EVA | COLLECTION | CRED_ACC | EXECUTION | LT_MOV | INI_ACC | DISCOVERY | C2 | IMPACT | RECON | EXFILTRATION | RES_DEV | TOTAL #TACTIC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PERSISTENCE | 235 | 356 | 185 | 0 | 17 | 57 | 13 | 39 | 0 | 0 | 0 | 0 | 0 | 0 | 847 |
| PRIV_ESC | 356 | 33 | 243 | 0 | 0 | 13 | 0 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 620 |
| DEF_EVA | 185 | 243 | 760 | 0 | 57 | 24 | 56 | 39 | 39 | 19 | 2 | 0 | 0 | 0 | 1377 |
| COLLECTION | 0 | 0 | 0 | 237 | 116 | 0 | 0 | 0 | 15 | 6 | 0 | 0 | 0 | 0 | 374 |
| CRED_ACC | 17 | 0 | 57 | 116 | 297 | 0 | 0 | 0 | 23 | 0 | 0 | 0 | 0 | 0 | 493 |
| EXECUTION | 57 | 13 | 24 | 0 | 0 | 234 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 332 |
| LT_MOV | 13 | 0 | 56 | 0 | 0 | 12 | 134 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 240 |
| INI_ACC | 39 | 13 | 39 | 0 | 0 | 0 | 25 | 179 | 0 | 0 | 0 | 0 | 0 | 0 | 269 |
| DISCOVERY | 0 | 0 | 39 | 15 | 23 | 0 | 0 | 0 | 243 | 0 | 0 | 0 | 0 | 0 | 320 |
| C2 | 0 | 0 | 19 | 6 | 0 | 0 | 0 | 0 | 0 | 241 | 0 | 0 | 0 | 0 | 266 |
| IMPACT | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 276 | 0 | 0 | 0 | 278 |
| RECON | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 196 | 0 | 0 | 196 |
| EXFILTRATION | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 95 |
| RES_DEV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 264 | 264 |
| | | No Overlaps | | | | With Overlaps | | | | | | | TOTAL: | | 5971 |

- **Pair-wise overlap for MITRE tactic descriptions**
- **Diagonal values correspond to the single-tactic descriptions**
- **Some descriptions match to two or more tactics. Hence, the total of 5971 instances are more than the curated descriptions**

## Results

| Model: | BERT | SecBERT | SecureBERT |
|---|---|---|---|
| Avg. Training Loss | 0.01 | 0.01 | 0.01 |
| Avg. Test Loss | 0.15 | 0.16 | 0.15 |
| Abs. Accuracy | 0.63 | 0.59 | 0.65 |
| Micro Avg. F1 Score | 0.75 | 0.72 | 0.76 |

**Table.1.** Results for running the three BERT models with 30 epochs using 5-fold cross-validation.

| *Per Tactic F1 Score:* | BERT | SecBERT | SecureBERT |
|---|---|---|---|
| COLLECTION | 0.68 | 0.63 | 0.68 |
| C2 | 0.75 | 0.73 | 0.75 |
| CRED_ACC | 0.74 | 0.72 | 0.76 |
| DEF_EVA | 0.78 | 0.77 | 0.79 |
| DISCOVERY | 0.73 | 0.66 | 0.75 |
| EXECUTION | 0.72 | 0.64 | 0.71 |
| EXFILTRATION | 0.59 | 0.57 | 0.57 |
| IMPACT | 0.77 | 0.75 | 0.82 |
| INI_ACC | 0.64 | 0.62 | 0.65 |
| LAT_MOV | 0.57 | 0.56 | 0.59 |
| PERSISTENCE | 0.78 | 0.73 | 0.78 |
| PRIV_ESC | 0.72 | 0.72 | 0.74 |
| RECON | 0.89 | 0.83 | 0.88 |
| RES_DEV | 0.85 | 0.85 | 0.85 |

**Table.2.** Results for per-tactic F1 score for the three models to measure the differences in values for single-label and multi-label descriptions.

## Observations & Future Works

**Based on the Results:**

- The 0.76 Micro F1 score in SecureBERT is promising in capturing semantic features of cybersecurity descriptions and dealing with multi-label data.
- The models could reasonably capture overlapping MITRE tactic descriptions

**Future Works:**

How to 1) better reflect the model's performance, 2) treat limited labeled data, 3) leverage label semantics, and 4) use a novel NSP-tuning approach to predict the intended consequences.

## References


[1] "MITRE ATT&CK®." https://attack.mitre.org/ (accessed Nov. 06, 2022).

[2] S. Moskal, S. Y.-2021 I. C. on, and undefined 2021, "Translating Intrusion Alerts to Cyberattack Stages using Pseudo-Active Transfer Learning (PATRL)," *ieeexplore.ieee.org,* Accessed: Sep. 13, 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9705037/

[3] J. Howard and S. Ruder, "Universal Language Model Fine-tuning for Text Classification," *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)*, vol. 1, pp. 328–339, Jan. 2018, doi: 10.48550/arxiv.1801.06146.

[4] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, vol. 1, pp. 4171–4186, Oct. 2018, doi: 10.48550/arxiv.1810.04805.

[5] J. Yin, M. J. Tang, J. Cao, and H. Wang, "Apply transfer learning to cybersecurity: Predicting exploitability of vulnerabilities by description," *Knowledge-based systems,* vol. 210. Elsevier B.V., Dec. 27, 2020. doi: 10.1016%2Fj.knosys.2020.106529.

[6] "SecBERT: pretrained BERT model for cyber security text, learned Cybersecurity Knowledge." https://github.com/jackaduma/SecBERT (accessed Oct. 02, 2022).

[7] E. Aghaei, E. Al-Shaer, X. Niu, and W. Shadid, "SecureBERT: A Domain-Specific Language Model for Cybersecurity Malware Deception View project Covert Communication using Network Behavioral Patterns", Accessed: Nov. 01, 2022. [Online]. Available: https://github.com/ehsanaghaei/SecureBERT

[8] Y. Liu *et al.,* "RoBERTa: A Robustly Optimized BERT Pretraining Approach," Jul. 2019, doi: 10.48550/arxiv.1907.11692.