Notes: Transformer in CSec : A Survey

Presented by:
Abhisek Keshari

Research Supervisor:
Dr. Virendra Singh

Outline

- Proposed Problem Statements / Ideas:
  - Detect Adversarial Behaviour by Applying NLP Techniques to Command Lines
  - NLP methods in HIDS
  - Use of Transformer Arch. over Network Flow Data

Proposed Problem Statements / Ideas

- Detect Adversarial Behaviour by Applying NLP Techniques to Command Lines.
- NLP methods in HIDS
- Use of Transformer Arch. over Network Flow Data

Detect Adversarial Behaviour (1/2)

Operations on computer systems frequently use the command line. Applications perform simple tasks (such as software updates) via tools or scripts, sysadmins deploy jobs that run on multiple machines, and technical users spend a great deal of their time in console windows. Adversaries also heavily utilize the command line. If a system has been compromised, an attacker will perform a majority of their actions using built-in system tools, at the command line. This strategy benefits the attacker, since their actions will look very similar to other normal tasks performed on the system. Methodology designed to automatically detect whether a system has been compromised needs to be able to tell the difference between benign and malicious command line operations. In order to build mechanisms capable of classifying command lines in this way, we first need to understand what they do – in other words, we need to be able to parse them in a similar way to how we parse natural languages. This article describes the process we've been using to develop methodology capable of parsing and categorizing command lines at F-Secure.

Source

Detect Adversarial Behaviour (2/2)

Using NLP techniques to perform NER and POS tagging as shown in the source.

Effort is required for data annotation but use of NLP techniques over CLI Commands are something new.

NLP methods in HIDS (1/4)

- This paper discuss several NLP methods for Host-based Intrusion Detection System.

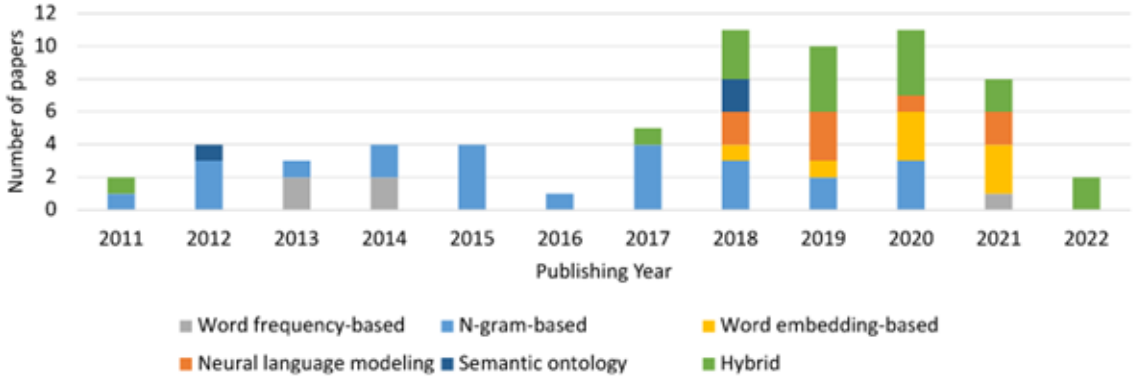
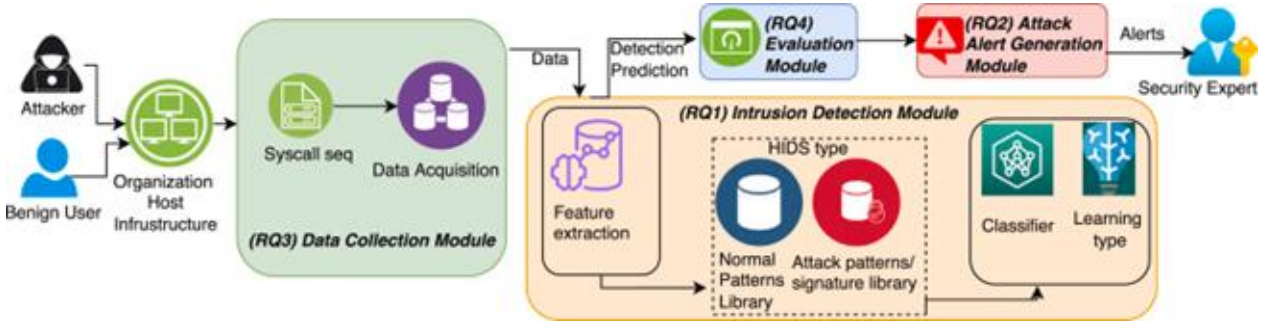


Fig. 4. Number of papers in NLP method categories used in HIDS which are published in each year from 2011 to 2022.



NLP methods in HIDS (2/4)

- Data Representation

Table 1. Example syscall sequences for reading two files (Grimmer et al., 2021).

Seq1	"open, read, write, open, read, write"
Seq2	"open, open, read, write, read, write"
Seq3	"open, open, read, read, write, write"

NLP method	Strengths	Weaknesses
Neural lang model	• Predicts a future syscall sequence possibly to be executed during an attack. • Combining the known invoked syscall traces with predicted future syscall sequences helps to improve the intrusion detection performance • Modeling syscalls helps to capture interword relationships.	• Modeling the system behavior requires a huge amount of data
Hybrid	• Makes HIDS more reliable and resilient against evasion and adversarial attacks by combining decisions from heterogeneous detectors • Gains the advantages of multiple features or models to lower FAR	• Requires high computation overhead compared to single methods.

NLP methods in HIDS (3/4)

Neural language modeling methods:

- RNN-VED-based language model
- LSTM-based language model
- GRU-based language model

Hybrid:

- Combination of heterogeneous classifiers using different NLP methods.
- Combination of n-gram and TF-IDF.
- Combination of n-gram and statistical approaches.
- Combination of n-gram and data augmentation methods.

NLP methods in HIDS (4/4)

As previous slide suggest there are several hybrid methods being used for HIDS, but there is no method that uses new STOA.

As there are sequence of ***open, read and write***, one can create a sequential model with attention capabilities for greater results.

Use of Transformer Arch. over Network Flow Data

There are limited paper which uses transformer arch. over network flow data.

In depth analysis shows that one can work on the embeddings of transformer to feed new embeddings in Transformer Arch.

Systematic Literature Review: One can work on SLA about transformer + Network Flow / Packet Flow data.

Todo - Finding good dataset with little class imbalance and have structured information

References

Transformer :

- Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017). (**Main**)
- Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." *Advances in neural information processing systems* 27 (2014). (**Seq-to-Seq**)
- Illustrated Guide to Transformers- Step by Step Explanation
- The Illustrated Transformer
- Transformers from Scratch
- Understanding LSTM Networks
- Word Vectors

Problem statements :

- Sworna, Zarrin Tasnim, Zahra Mousavi, and Muhammad Ali Babar. "NLP methods in host-based intrusion detection Systems: A systematic review and future directions." *Journal of Network and Computer Applications* (2023): 103761. (**NLP-HIDS**)
- Detecting adversarial behaviour by applying NLP techniques to command lines (**NLP-CIi**)
- Hendler, Danny, Shay Kels, and Amir Rubin. "Detecting malicious powershell commands using deep neural networks." *Proceedings of the 2018 on Asia conference on computer and communications security*. 2018. (**NLP-CIi**)

References

Embedding Methods:

- Babaria, Rushi, et al. "FlowFormers: Transformer-based Models for Real-time Network Flow Classification." *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*. IEEE, 2021. (**Emb. Method - 1**)
- Fang, Cheng, et al. "A method of network traffic anomaly detection based on Packet Window Transformer." *2022 7th IEEE International Conference on Data Science in Cyberspace (DSC)*. IEEE, 2022. (**Emb. Method - 2**)
- Wu, Zihan, et al. "RTIDS: A robust transformer-based approach for intrusion detection system." *IEEE Access* 10 (2022): 64375-64387. (**Emb. Method - 3**)
- Marino, Daniel L., et al. "Self-supervised and interpretable anomaly detection using network transformers." *arXiv preprint arXiv:2202.12997* (2022). (**Emb. Method - 4**)

References

- Wen, Qingsong, et al. "Transformers in time series: A survey." *arXiv preprint arXiv:2202.07125* (2022). (**Survey-1**)
- Tuli, Shreshth, Giuliano Casale, and Nicholas R. Jennings. "Tranad: Deep transformer networks for anomaly detection in multivariate time series data." *arXiv preprint arXiv:2201.07284* (2022).
- Wang, Xixuan, et al. "Variational transformer-based anomaly detection approach for multivariate time series." *Measurement* 191 (2022): 110791.
- Zhang, Hongwei, et al. "Unsupervised anomaly detection in multivariate time series through transformer-based variational autoencoder." *2021 33rd Chinese Control and Decision Conference (CCDC)*. IEEE, 2021.
- Xu, Jiehui, et al. "Anomaly transformer: Time series anomaly detection with association discrepancy." *arXiv preprint arXiv:2110.02642* (2021).