Politecnico di Torino 1859

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

COMPUTER AND CONTROL ENGINEERING

DM630/HUAWEI - AI for Secured Networks: Language Models for Automated Security Log Analysis

Funded By	Huawei Technologies France SAS [P.iva/CF:04451063739] MINISTERO DELL'UNIVERSITA' E DELLA RICERCA [P.iva/CF:97429780584]

Supervisor	MELLIA MARCO - marco.mellia@polito.it

Contact	MELLIA MARCO - marco.mellia@polito.it VASSIO LUCA - luca.vassio@polito.it

Context of the research activity	Network security analysts are a key component of the defence infrastructure of an organization. They continuously and manually analyze security alarms and logs to make decisions against undesired intrusions. Language Models (LMs) demonstrated huge potential in processing texts. The research will evaluate the capabilities of LM agents (lightweight, large and multi-modal ones) in automating the investigations of security logs and performing zero-shot classification through generalization. Progetto finanziato dal PNRR a valere sul DM 630/2024 - CUP E14D24002440004

	Research objectives: Investigate and evaluate the capabilities of LLM agents in automating the manual investigations of the security analyst. This would assist them in analysis and incident reporting. The candidate will perform research to determine whether, and to what extent, the recent advances in language models could be used to automate and assist security analysts in the process (i) of learning the security-device rules by example and (ii) autonomously investigating the challenging cases currently analyzed by humans. In the second phase, the candidate will investigate how and if lightweight and generalizable language models can extract insights from raw data, as today large language models can do. The goal is to investigate whether models with limited supervision and a minimal number of trusted labels can attain comparable performance to generic large language models (LLMs) when

Objectives

applied to specific tasks such as code understanding, classification, anomaly detection, bug detection, or identifying security breaches.

The research will consider multi-modal embeddings to conceptually constrain the embeddings towards the right task.
By forcing the model to create multi-modal embeddings conceptually constrained to the right task, the model will possess the ability to generalize and autonomously reason about novel and previously unencountered tasks. For instance, test joint learning of (i) natural language label explanation of the security threat and (ii) the packet payload, using, e.g., contrastive learning techniques.

The project will involve a collaboration with Huawei Technologies France and Politecnico di Torino.

Outline of the research work plan:

1st year- Study of the state-of-the-art for security log analysis and state-of-the-art language models in ML.- Data collection and analysis of raw and structured data on security devices such as Firewall/Intrusion Prevention Systems (IPS), Endpoint Detection and Response (EDR) and Cloud security services.

2nd year- Adaptation and extension solutions to learn the security-device rules by example and autonomously investigate complex cases.- Propose and develop innovative solutions to the problems of cyber threats analysis with Language models.- Propose multi-modal embeddings for network raw data and security logs.

3rd year - Tune the developed techniques and highlight possible strategies to counteract the various threats.- Application of the strategies to new data for validation and testing.

References:
- Boffa, M., Valentim, R. V., Vassio, L., Giordano, D., Drago, I., Mellia, M., & Houidi, Z. B. (2023). LogPr\'ecis: Unleashing Language Models for Automated Shell Log Analysis. arXiv preprint arXiv:2307.08309- Boffa, M., Milan, G., Vassio, L., Drago, I., Mellia, M., & Houidi, Z. B. (2022, June). Towards nlp-based processing of honeypot logs. In 2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW) (pp. 314-321). IEEE.- Boffa, M., Vassio, L., Mellia, M., Drago, I., Milan, G., Houidi, Z. B., & Rossi, D. (2022, December). On using pretext tasks to learn representations from network logs. In Proceedings of the 1st International Workshop on Native Network Intelligence (pp. 21-26).

List of possible venues for publications:
- Security venues: IEEE Symposium on Security and Privacy, IEEE Transactions on Information Forensics and Security, ACM Symposium on Computer and Communications Security (CCS), USENIX Security Symposium, IEEE Security & Privacy;
- AI venues: Neural Information Processing Systems (NeurIPS), International Conference on Learning Representations (ICLR), International Conference on Machine Learning (ICML), AAAI Conference on Artificial Intelligence, ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD), European Conference on Machine Learning and Knowledge Discovery in Databases (ECML/PKDD);

	- Computer networks venues: Distributed System Security Symposium (NDSS), Privacy Enhancing Technologies Symposium, The Web Conference (formerly International World Wide Web Conference WWW), ACM International Conference on Emerging Networking EXperiments and Technologies (CoNEXT), USENIX Symposium on Networked Systems Design and Implementation (NSDI).

Skills and competencies for the development of the activity	- Good programming skills (such as Python, Torch, Spark) - Excellent Machine Learning knowledge - Knowledge NLP and LM - Basics of Networking and security