The enhancement of the human-machine and process-machine interactions in machine learning-based malicious activity detection systems

Summary

TalTech School of Information Technologies, Department of Software Sciences offers a 4- year PhD position in the field of ICT.

Research field:	Information and Communication Technology
Supervisors:	Sven Nõmm Hayretdin Bahsi
Availability:	This position is available.
Offered by:	School of Information Technologies Department of Software Science
Application deadline:	Applications are accepted between September 01, 2020 00:00 and October 02, 2020 23:59 (Europe/Zurich)

Description

Background and motivation

Cyber-attack tools and techniques have evolved rapidly as the cyber domain is a very lucrative target for different threat actors with varying sophistication levels (e.g., cybercriminals, state-sponsored hacker groups, opportunistic attackers). It is widely accepted that human, technology and process aspects should be covered for comprehensive cyber protection. The interaction between these aspects has not been properly investigated in the literature, and machine learning (ML) solutions utilized in the cybersecurity domain are not different in this respect. ML is perceived as a significant solution instrument for the complex problem domain of cyber threats. Despite the fact that the literature includes many promising research results, the incorporation of ML methods into the current technological solutions is limited. We conjecture that the main reason for this discrepancy between the academic achievement and scarce real-world implementation is due to the myopic view that does not reflect the human and process considerations in the learning models. The operations in current security operating centres take place in complex processes with the involvement of various human analyst roles [1]. Human analysts are interested in the high interpretability of the results besides the optimized accuracy [2]. However, interpretability within the cybersecurity field has drawn very little attention in the research communities. Finding labelled data is very difficult due to the lack of enough human resources. The response of the current technology to this problem has been just the utilization of unsupervised learning methods. Unfortunately, it suffers from low accuracy values and inadaptability of learning models to different organizations. Active learning methods, which enable the acquisition of human know-how into the models, could be adapted to the cybersecurity processes. Transfer learning could lead to broader usage of learning models created for one organization to other organizations.

Main objective and research methods

In this research study, we aim to develop learning methods that adapt transfer and active learning into the cyber-security problems and investigate the interpretability in this problem domain. However, we consider that the other application areas such as health, manufacturing, smart systems (e.g., vehicles, cities) can benefit from our research outcomes. Quantitative and qualitative interpretability metrics will be developed and evaluated for the cybersecurity field within this research. Interpretability will be investigated within model-specific and model-agnostic contexts. Transfer and active-learning methods will be explored and adapted to the problem domains. The publicly available datasets about IoT botnet, mobile malware, cyber-attacks to autonomous vehicles, SCADA systems (i.e., including the datasets generated by our research group) will be utilized.

Expected impact

The ability to explain results achieved by the application of machine learning and AI techniques would allow more synergetic human-machine cooperation. On the one hand, it will provide valuable information to increase the goodness of machine learning techniques. On the other hand, it will help to improve the performance of the human agents participating in the cooperation.

Candidate's Background and Knowledge

The candidate is expected to have a solid knowledge of mathematics and statistics and concepts of the cybersecurity and completed master level courses in machine learning and data mining. Experience in the application of deep learning is preferable. Possess skills in software development. Good knowledge of Python is preferable.

References

[1] Guerra-Manzanares, Alejandro; Nõmm, Sven; Bahsi, Hayretdin (2019). Towards the integration of a post-hoc interpretation step into the machine learning workflow for IoT botnet detection. Proceedings 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019: 16-19 December, Boca Raton, Florida, USA

[2] Carvalho, Diogo V., Eduardo M. Pereira, and Jaime S. Cardoso. "Machine Learning Interpretability: A Survey on Methods and Metrics." Electronics 8.8 (2019): 832.

[3] Guerra-Manzanares, Alejandro; Medina-Galindo, Jorge; Bahsi, Hayretdin; Nõmm, Sven (2020). MedBIoT: Generation of an IoT Botnet Dataset in a Medium-sized IoT Network. Proceedings of the 6th International Conference on Information Systems Security and Privacy, ICISSP 2020, February 25-27, 2020

To get more information or to apply online, visit https://taltech.glowbase.com/positions/141 or scan the the code on the left with your smartphone.