# Cyber Threat Detection using Machine Learning on Graphs

JAKUB REHA
jakubr@kth.se

January 24, 2023

## 1 Thesis title

Cyber Threat Detection using Machine Learning on Graphs

## 2 Background

Cyber attacks are ubiquitous and increasingly prevalent in the industry, society, and governmental departments and agencies. They affect economy, politics, and individuals. Ever-increasingly skilled, organized, and funded threat actors combined with ever-increasing volumes and modalities of monitorable data require increasingly automated, sophisticated, and innovative cyber defense solutions. State-of-the-art enterprise and endpoint security systems conduct threat detection on dynamic graph representations of computer systems and enterprise communication networks. Comprehensive modelling followed by autonomous, un- and self-supervised learning of the normal behavior of network entities allow for the detection of malicious actions as graph anomalies.

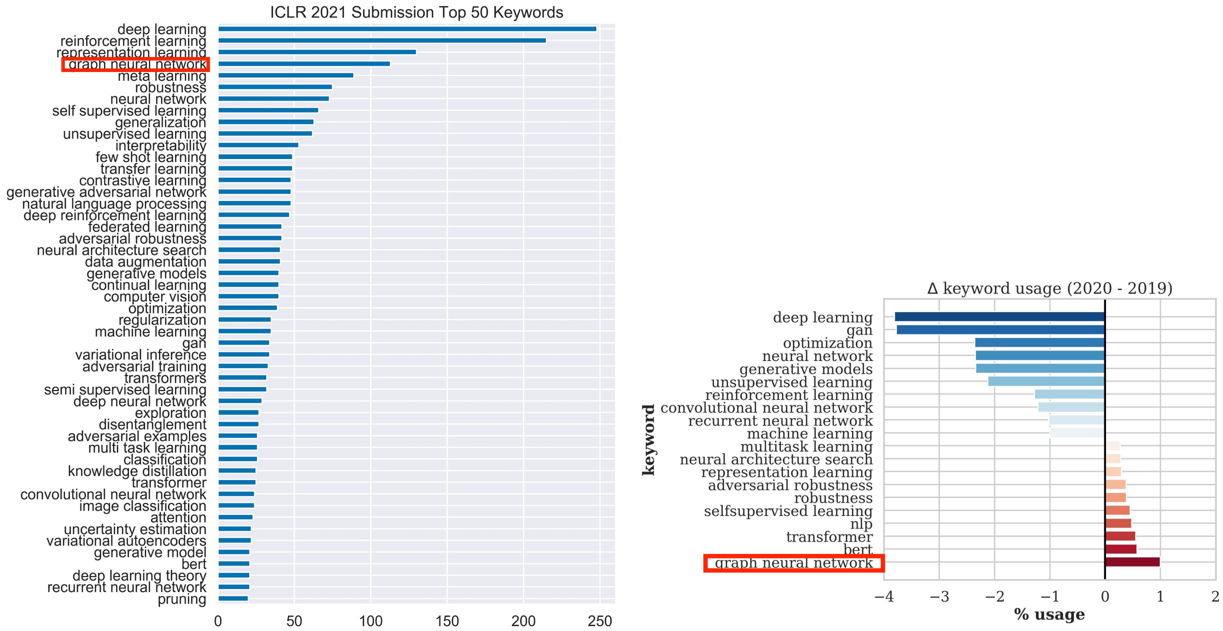

Figure 1: Recent popularity of graph neural networks at machine learning conferences.
Source: https://petar-v.com/talks/MLPL-GNN.pdf

Graph Neural Networks (GNNs), such as Graph Convolutional Network (GCN), Graph Attention Network (GAT) and Graph Isomorphism Networks (GIN), are becoming the de facto default types of architectures to solve prediction tasks on graph-structured data. At the same time, there is an increasing awareness among researchers and in the industry about the flexibility and benefits of an explicit exploitation and modelling of relations between entities of data, processes, systems, networks, and knowledge as graphs. GNNs generate latent representations for nodes based on structural, semantic, and possibly temporal information aggregated over the nodes' neighborhoods. Formulation of node-, edge-, and graph-level prediction tasks allow for a broad range of applications at un-, self-, semi-, and full levels of supervision. Figure 1 shows the recent popularity of GNNs based on keyword matching in ICLR conference proceedings.

The Artificial Intelligence for Cyber Security (AI4Sec) Research Team of Huawei Technologies, located at Huawei's Munich Research Center, is specialized in graph machine learning, anomaly detection and threat detection in computer and enterprise networks, among other things. The value of this degree project to the team and company is to extend the team's research scope from anomaly detection in enterprise network graphs to anomaly detection in host-level provenance graphs constructed from endpoint logs. This research direction is in its infancy as there are only three papers on GNN-based anomaly detection in provenance graphs available thus far, all published at top conferences and journals in 2022 [1, 2, 3]. The availability of publicly available datasets, extensive relevant expertise and experience of the supervising AI4Sec Team members, and the comprehensive training in machine learning subjects, algorithms and data analytics by the student, Jakub Reha, makes this research-oriented degree project well-scoped.

# 3 Research question

**What are effective self-supervised approaches to detecting novel cyber threats as anomalies in provenance graphs constructed from endpoint logs?**

This primary research question entails the following secondary research questions, not all of which will need to be answered in the scope of this degree project:

1. What are suitable GNN architectures, prediction tasks, loss functions, and graph anomaly detection tasks for cyber threat detection, where anomalies correspond to cyber threats?

2. How to best construct graph representations of user and system activities from endpoint/audit logs suitable for graph ML based threat detection?

3. What are the characteristics and properties of the constructed graphs and how can those be leveraged to make graph ML models perform well on the downstream task of cyber threat detection?

4. What are the categories of cyber threats that are suitable for detection via graph ML and graph anomaly detection?

5. How to cope with the inherent problem of anomaly based approaches to threat detection resulting in high false positive rates caused by the ubiquitous presence of benign anomalies?

6. How to utilize the temporal information provide as timestamps in the underlying endpoint logs for threat detection?

7. How to address attack strategies, where attackers use benign data to evade anomaly detection models (adversarial attacks)?

# 4 Hypothesis

The expected outcome of this degree project is a developed proof of concept of a self-supervised graph machine learning-based system capable of detecting cyber threats as anomalies in suitable graph representations constructed from endpoint/audit logs.

As this research field is in its infancy, there are only very few recently published – but high-quality – papers available addressing this particular or closely related problems [1, 2, 3]. This degree project is expected to yield and document a collective understanding of these works and formulate, develop, and test a new approach to this yet-to-be concretized threat detection problem – possibly addressing a slightly different detection goal and/or using different algorithmic approaches than those proposed in [1, 2, 3].

# 5 Research method

1. Conduction of a structured and documented literature study during which the student familiarizes themselves with the following areas:
   (a) Provenance based cyber threat detection (priority:[1, 2, 3], supportive, e.g. [4], [5])
   (b) Graph summarization on system audit logs (e.g. [6])
   (c) Self supervised ML on graphs (e.g. [7])
   (d) Graph ML based anomaly detection (e.g. [8, 9, 10, 11])
2. Concretization of the cyber threat detection goal: Specification of categories of cyber attacks/threats to target (guided by the industrial supervisor). The MITRE ATT&CK knowledge base will be analyzed to help answer this research question.
3. Drafting and formulation of preliminary algorithmic ideas inspired by the studied literature including but not limited to [1, 2, 3] and based on brainstorming sessions with the industrial supervisor and the team.
4. Identification of suitable assessment metrics and validation criteria based on common practices used in the related research literature.
5. Identification of baseline methods to compare the performance of the cyber threat detection system to.
6. Identification of suitable public datasets based on related literature. Candidate datasets are listed in Section 10.3.
7. Familiarization with the data via explorative data analyses (using, e.g., Python).
8. Application and testing of different approaches to construct provenance graphs from endpoint/audit logs.
9. In order to be able to effectively detect cyber attacks/threats as anomalies, the student will analyze statistical properties of the cyber attacks/threats of interest and statistically characterize deviations of data corresponding to cyber attacks/threats from data corresponding to normal ”benign” activities/behavior.
10. Possible generation of one or multiple complementary synthetic datasets to support the research.
11. Update and concretization of the preliminary algorithmic design idea(s).
12. Conduction of feature engineering, i.e., identification and extraction of candidate features.

13. Creation and maintenance a well-structured, documented, and largely self-contained git repository for degree project.
14. Implementation of the previously designed algorithm(s) (PyTorch, PyG).
15. Identification of suitable experiments and experimental setups.
16. Conduction of experiments including but not limited to hyperparameter tuning, an ablation study, and runtime analysis. Performance will be evaluated by the initially idetified metrics.

# 6 Background of the student

The student has experience in Machine Learning both academically and professionally. Below is provided a list of relevant courses that the student has taken during their Bachelor and Master Degree studies:
DTU (Bachelor studies):

- Introduction to Cyber Systems (02135)

KTH (Master studies):

- Machine Learning (DD2421)
- Machine Learning, Advanced Course (DD2434)
- Artificial Intelligence (DD2380)
- Probabilistic Graphical Models (DD2420)
- Deep Learning in Data Science (DD2424)

HKUST (Exchange studies):

- Machine Learning with Structured Data (COMP4222) (Graph Machine Learning)
- Operating Systems (COMP3511)
- Design and Analysis of Algorithms (COMP3711)

The student is familiar with the following relevant programming languages:
Python, PyTorch, SQL, C++, TensorFlow, Java, Git

# 7 Supervisor at the company/external organization

1. Dr. Claas Grohnfeldt
   (a) Mentor, advisor, and technical supervisor of the student (Jakub Reha)
   (b) Defined the topic and technical scope of this degree project in alignment with his team and the student
   (c) Served as the hiring manager and technical interviewer for the Master Thesis Student position corresponding to this degree project that has been filled by the student
   (d) Point of contact for the student for questions and issues related to:
      i. administration, company culture, and processes
      ii. personal and professional development
      iii. hardware/assets/equipment
      iv. software and tools

v. practices in research and development

vi. vi. algorithms, modeling, data science, machine learning

(e) Will hold regular meetings with the student to discuss and support the progress of this degree project

(f) Qualification and role in the company:

i. Principal Research Engineer in the AI4Sec Team of Huawei Munich Research Center

ii. Co-initiator, project lead, and manager of the long-term project "Graph-Learning-based Network Security Analytics and ThreAt Detection (GLAAD)"

iii. Specialized in algorithms, machine learning (especially on graphs), network science, anomaly detection, threat detection, and network security

2. Michele Russo, Sr. Research Engineer

   (a) Advisor and technical co-supervisor of the student (Jakub Reha)

   (b) Contributed to the identification of the topic and technical scope of this degree

   (c) Point of contact for the student for questions and issues related to:

   i. personal and professional development

   ii. hardware/assets/equipment

   iii. software and tools

   iv. practices in research and development

   v. programming, cyber security, machine learning

# 8 Suggested examiner at KTH

Professor Sarunas Girdzijauskas confirmed that he would be the examiner.

# 9 Suggested supervisor at KTH

Ahmed Emad Samy Yossef Ahmed confirmed that he would be the supervisor.

# 10 Resources

## 10.1 Mother project and expertise

This degree project will be an integrable part of an ongoing long-term Huawei-internal project, which started in 2019 and will end no sooner than 2024, called "Graph Learning-based network security Analytics and threAt Detection (GLAAD)". Both supervisors listed in Section 7 are leading researchers and developers in this project with expertise and comprehensive experience in research, algorithm design, machine learning, graph anomaly detection, threat detection, and cyber security in general.

## 10.2 Hardware and Software

The student will be provided with any hardware and software necessary to complete this degree project including but not limited to the following: Hardware:

1. Electrically adjustable desk, work laptop, monitor, keyboard, mouse, docking hub, etc.

2. Access to CPU Clusters and an Apache Spark Cluster

3. Access to a GPU

Software:

1. Reference Management System: Zotero
2. Programming language: Python version 3.9 or later
3. Libraries: PyTorch, PyG, PySpark, etc.
4. Code editors: Visual Studio Code, PyCharm, and JupyterLab
5. Documentation: LaTeX (Overleaf, VS Code, VIM, TeXstudio, or similar) and, possibly, Markdown (VS Code, Obsidian)
6. GitLab

### 10.3 Datasets

In order to allow for publication of the output of the degree project, the student will work with public datasets (possibly complemented by self-synthesized datasets for smaller experiments).
Below is a list of candidate public datasets related to threat detection on provenance graphs constructed from endpoint/audit logs:

- DARPA TC E3 (Transparent Computing Engagement 3 Data) Link
- DARPA OpTC (Darpa Operationally Transparent Cyber (OpTC) Dataset) Link
- CERT's Dataset (Insider Threat Test Dataset) Link
- LANL's Dataset (Los Alamos National Laboratory Dataset) Link

In addition to these four datasets, the survey paper on provenance-based intrusion detection systems [5] references a total of 12 related datasets that might qualify for development, tests, and validation conducted in the scope of this degree project.

## 11 Eligibility

Passed courses at second cycle level of at least 60 credits, including:

1. courses that are relevant for the degree project (see Section 6)
2. a course in theory of knowledge and research methodology (Introduction to the Philosophy of Science and Research Methodology (DA2205))

## 12 Study Planning

After the completion of:

1. DD2301 Program Integrating Course in Machine Learning on 16.01.2023

the thesis is the last element of the student's education.

# References

[1] J. Zengy, X. Wang, J. Liu, Y. Chen, Z. Liang, T.-S. Chua, and Z. L. Chua, "Shadewatcher: Recommendation-guided cyber threat analysis using system audit records," in *2022 IEEE Symposium on Security and Privacy (SP)*, 2022.

[2] Z. Huang, Y. Gu, and Q. Zhao, "One-class directed heterogeneous graph neural network for intrusion detection," in *2022 the 6th International Conference on Innovation in Artificial Intelligence (ICIAI)*. Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3529466.3529480

[3] S. Wang, Z. Wang, T. Zhou, X. Yin, D. Han, H. Zhang, H. Sun, X. Shi, and J. Yang, "threatrace: Detecting and tracing host-based threats in node level through provenance graph learning," *IEEE Transactions on Information Forensics and Security*, 2022.

[4] X. Han, T. F. J. Pasquier, A. Bates, J. Mickens, and M. I. Seltzer, "UNICORN: runtime provenance-based detector for advanced persistent threats," *CoRR*, 2020. [Online]. Available: http://arxiv.org/abs/2001.01525

[5] M. Zipperle, F. Gottwalt, E. Chang, and T. Dillon, "Provenance-based intrusion detection systems: A survey," *ACM Comput. Surv.*, 2022. [Online]. Available: https://doi.org/10.1145/3539605

[6] Z. Xu, P. Fang, C. Liu, X. Xiao, Y. Wen, and D. Meng, "Depcomm: Graph summarization on system audit logs for attack investigation," in *2022 IEEE Symposium on Security and Privacy (SP)*, 2022.

[7] Y. Liu, M. Jin, S. Pan, C. Zhou, Y. Zheng, F. Xia, and P. Yu, "Graph self-supervised learning: A survey," *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[8] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu, "A comprehensive survey on graph anomaly detection with deep learning," *IEEE Transactions on Knowledge and Data Engineering*, 2021.

[9] X. Luo, J. Wu, A. Beheshti, J. Yang, X. Zhang, Y. Wang, and S. Xue, "Comga: Community-aware attributed graph anomaly detection." Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3488560.3498389

[10] K. Liu, Y. Dou, Y. Zhao, X. Ding, X. Hu, R. Zhang, K. Ding, C. Chen, H. Peng, K. Shu, L. Sun, J. Li, G. H. Chen, Z. Jia, and P. S. Yu, "Bond: Benchmarking unsupervised outlier node detection on static attributed graphs," in *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2022. [Online]. Available: https://openreview.net/forum?id=YXvGXEmtZ5N

[11] Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis, "Anomaly detection on attributed networks via contrastive self-supervised learning," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

# Acronyms

**DTU** Technical University of Denmark

**GNN** Graph Neural Network

**HKUST** Hong Kong University of Science and Technology

**KTH** KTH Royal Institute of Technology

**ML** Machine Learning

**TDR** Teacher Directed Reading

**TRL** Technology Readiness Level