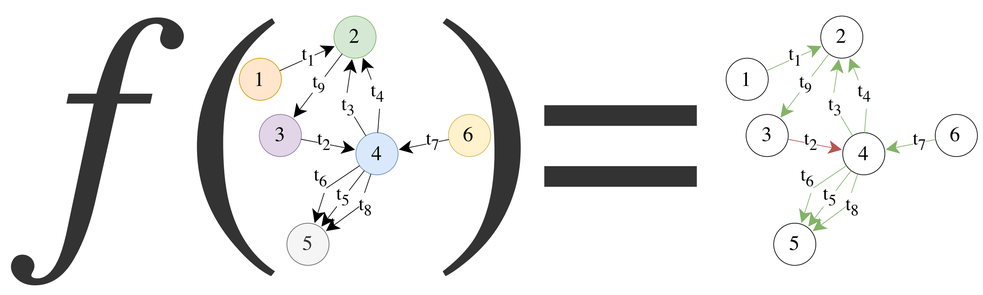
One-Class Learning on Temporal Graphs for Attack Detection in Cyber-Physical Systems

Robin Buchta*, Tobias Fritz§, Carsten Kleiner*, Felix Heine*, Gabi Dreo Rodosek§,
*Institute for Applied Data Science Hannover (Data|H), Hochschule Hannover - University of Applied Sciences and Arts, Hanover, Germany
§Research Institute CODE, Universität der Bundeswehr München, Munich, Germany
{robin.buchta, carsten.kleiner, felix.heine}@hs-hannover.de
{tobias.fritz, gabi.dreo}@unibw.de

Motivation & Objective

Various domains, including critical infrastructures, industry, and the private sector, deploy cyber-physical systems (CPS). These systems integrate IT and OT components and interact with the environment. CPS often operate as black boxes, hindering effective attack detection. Our research addresses the challenge of detecting attacks in CPS relying on network data and learning on normal behavior.



How can we detect attacks in CPS by relying on network data and without specific attack instances?

Contributions

- We explore novel one-class learning for attack detection on temporal graph structures in CPS based on network communication.
- We create a benchmark for graph-based attack detection for CPS on two datasets and provide the first baselines.
- We analyze the strengths and weaknesses of Edgebank and TGN in this environment and identify the potential for improvement in pre-processing, method application, and datasets.

Methodology

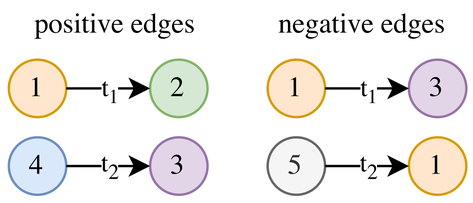
Datasets → Loaders → Models → Evaluation → Results

Datasets & Loaders

Datasets: TON_IoT [1] & Edge_IIoT [2]. We treat each IP address as a node and each flow as an event.

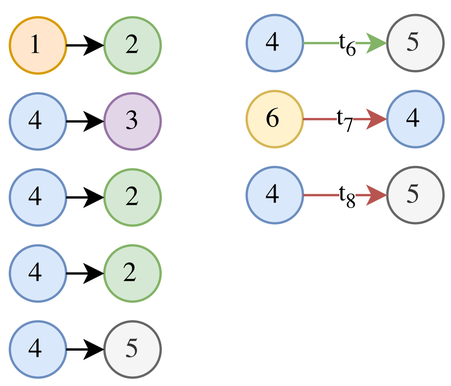
	TON_IoT	Edge_IIoT
Nodes	16.745	165.160
Edges	461.043	2.096.419
Train Edges	100.000	100.000
Benign Edges	161.043	603.558

Negative sampling for training: We randomly permute the source and destination node for every event to generate a negative sample for every positive sample.



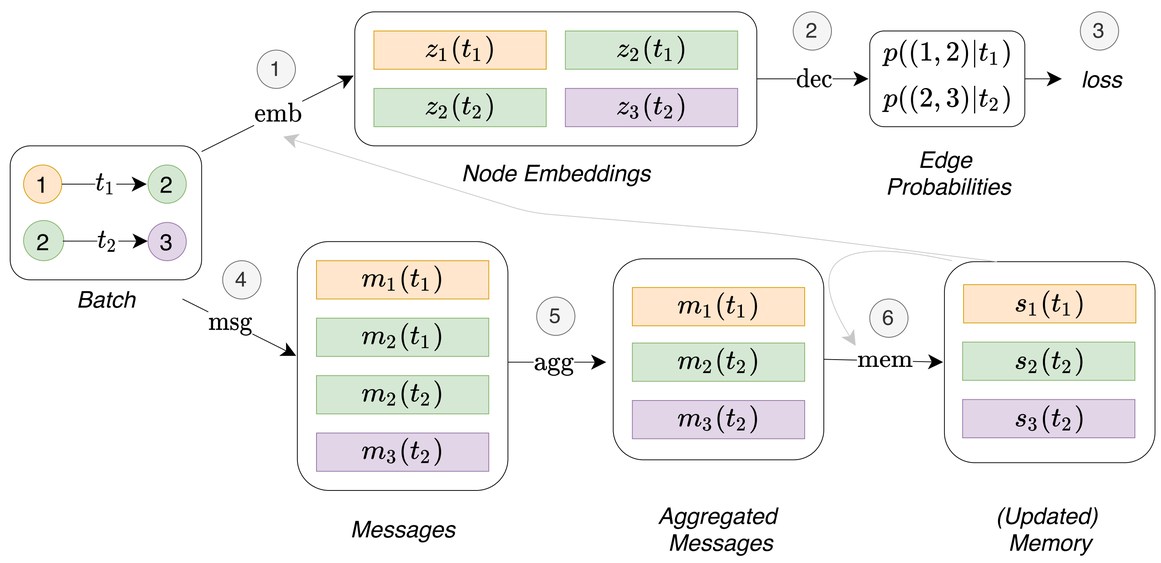
Baseline: EdgeBank

- EdgeBank [3], is memorization-based.
- The memory holds known connections without considering attributes.
- Recognition is based on a comparison the batch with the current state of the memory.
- There are two variants: The unlimited ∞ and time window-based tw version.
- As an example, we take the graph from above and infer the last three.



Temporal Graph Networks

- TGN [4] comprises a memory module, message function, aggregator, updater, and final embedding. The decoder uses a multilayer perceptron to output probabilities.
- We used supervised learning with negative sampling.
- Our objective function is to minimize the binary cross entropy.



Computations performed by TGN on a batch of time-stamped interactions [4].

Evaluation

Update Strategies: In the predicted knowledge strategy, we only use the nodes of benign events. Attack knowledge is a fictitious scenario to see how the methods perform when only benign nodes are used for updating. Blind is the conventional method, where all nodes are used for the update, and no is used to make no updates.

Metrics: The confusion matrix is used for detailed analysis and the F1 score as a comparative value.

Results

Strategy	EdgeBank$_\infty$		EdgeBank$_{tw}$		TGN$_{avg(e50)}$	
Predicted Knowledge	**0.76**	0.99	0.69	0.99	0.64	0.60
Attack Knowledge	**0.90**	0.99	0.89	0.99	0.72	0.63
Blind	0.09	0.52	0.13	0.52	**0.39**	0.52
None	0.77	0.99	**0.77**	0.99	0.73	0.77

Conclusions

- Our experiments revealed that EdgeBank, a heuristic, suffices for simple attacks.
- Although EdgeBank outperformed TGN in this experiments.
- However, for more complex attacks, EdgeBank fails because it recognizes every incoming connection as benign if it is in memory.
- TGN demonstrated potential but requires further exploration.

References

[1] Nour Moustafa. "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets". In: *Sustainable Cities and Society* 72 (2021), p. 102994. issn: 2210-6707.

[2] Mohamed Amine Ferrag et al. "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning". In: *IEEE Access* 10 (2022), pp. 40281–40306.

[3] Farimah Poursafaei et al. "Towards Better Evaluation for Dynamic Link Prediction". In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo et al. Vol. 35. Curran Associates, Inc., 2022, pp. 32928–32941.

[4] Emanuele Rossi et al. *Temporal Graph Networks for Deep Learning on Dynamic Graphs*. 2020. arXiv: 2006.10637 [cs.LG].