

Temporal Graph Auto-Encoders for Anomaly Detection in Industrial Internet of Things

Postdoctoral position - 1 year (renewable), starting in January 2025

LISIC Laboratory - Université du Littoral Côte d'Opale - Saint Omer, France

The Industrial Internet of Things (IIoT) is a rapidly evolving paradigm in which industrial sensors, machines, and other instruments are connected to the internet, enabling device-to-server and device-to-device communications for real-time data exchange. IIoT systems produce two types of data where detecting anomalies is crucial: (i) device communication logs and (ii) device measurements. For the former, devices are now exposed to attacks or intrusions, thus making it necessary to search for signs of these events in the communication logs. For the latter, any equipment or operational issues will lead to changes in IIoT measurements and potentially critical production shutdowns, making it vital to promptly detect these events to trigger corrective actions.

A natural model for IIoT data are attributed weighted temporal graphs: nodes can represent devices; node features can represent their measurements; and time-varying weighted edges can capture various types of information, such as their communications, their statistical dependencies, flows through them, etc. Figure 1 gives an illustration. Anomaly detection in temporal graphs is a highly active area of research. Yet, due to a lack of adapted tools, most methods do not directly analyze the temporal graph. Instead, they project it into time-series or graphs: technique that allows to leverage existing methods for time-series (time anomalies) or graphs (structural anomalies). However, such projections often involve drastic loss of information and they can make both normal and abnormal instances project equally, as Figure 1 also illustrates.

This postdoc project proposes to go beyond existing approaches and to develop adapted algorithms that directly search for anomalies in the temporal graph.

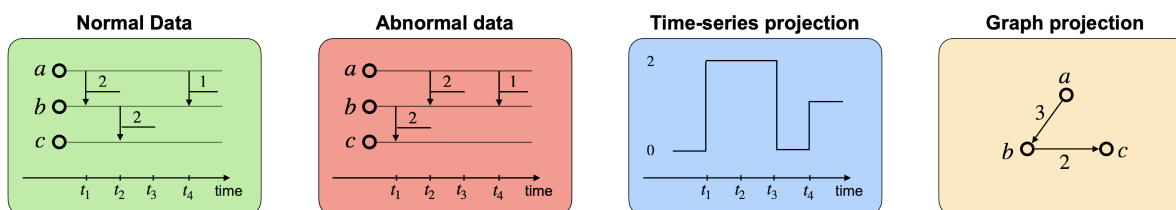


Figure 1: Illustration of temporal graphs with nodes a, b, c and weighted links on the time interval $[t_1, t_5]$. The arrow $a \rightarrow b$ at t_4 may represent a packet sent from a to b that took 1 second and had 1 byte of data. The temporal graph of the green panel may be considered normal as there is a time-respecting path from a to c . The temporal graph from the red panel may be considered abnormal as information from a cannot reach c . Anomaly detection approaches often aggregate edges: vertically to project the data into a time series (blue panel), or horizontally to project the data into a graph (yellow panel). In the example, both temporal graphs project to the same time series and graph.

The recruited postdoc will collaborate with us on the following research axes:

1. **Temporal graph-based auto-encoder architecture.** We aim to build an anomaly detector based on the auto-encoder architecture shown in Figure 2. This architecture calls to extend either time-series based encoders or static graph based encoders to temporal graphs [1]. Contrary to recent efforts that attempt to encode temporal graphs by summarizing their temporal information in individual link, node, or sub-graph encodings [2], we search for a single encoding that jointly summarizes structure and dynamics. We plan to address this challenge by building upon recent results that combine graph theory and signal processing into a unified formalism for temporal graphs [3].

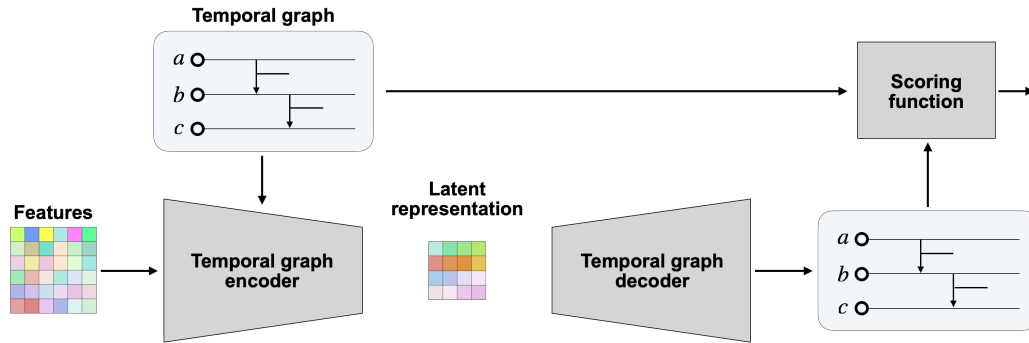


Figure 2: A temporal graph-based auto-encoder architecture for anomaly detection.

2. **Detection of network intrusions and attacks in communication logs.** We plan to evaluate the proposed architecture in real-world IIoT traffic datasets that capture various types of intrusions and attacks [4]. Such events correspond to abnormal subsets of the temporal graph, which are themselves temporal graphs. Thus, the developed architecture must produce large scores if it is run on those sub-temporal graphs. The challenge lies in the exponential number of sub-temporal graphs to test. We are interested in exploring pruning strategies, potentially by leveraging the latent representation, which has proven effective in other applications [5].
3. **Detection of faults in IIoT measurements.** It is common practice to use graphs to represent the statistical dependencies in time-series data from sensor measurements [6]. Dependencies are rarely stationary, thus a temporal graph representation can enhance the time-space localization of anomalies. We aim to explore the advantages of our approach in real-world, potentially on-line, scenarios, such as machine health monitoring, transportation network monitoring, or other use cases that may arise from collaborations with local companies.

Profile. We look for highly motivated candidates with relevant experience in anomaly detection, graph machine learning, and/or deep learning. Experience in Python programming, cybersecurity and/or streaming algorithms is a plus. Ideal candidates will have a publication record in selective AI conferences.

Application. Interested candidates are invited to send a cover letter and a detailed CV (with a publication list and the contact details of two references) to:

- Esteban Bautista: esteban.bautista@univ-littoral.fr
- Claire Guilloteau: claire.guilloteau@univ-littoral.fr

Applications will be reviewed on a rolling basis until the position is filled.

References

- [1] S. M. Kazemi et al., *Representation learning for dynamic graphs: A survey*. Journal of Machine Learning Research, vol. 21, no. 70, pp. 1–73, 2020.
- [2] E. Bautista, et al. *MAD: Multi-Scale Anomaly Detection in Link Streams*. Proceedings of the 17th ACM International Conference on Web Search and Data Mining. 2024.
- [3] E. Bautista et al., *Link Streams as a Generalization of Graphs and Time Series*. In 2023 IEEE 5th International Conference on Cognitive Machine Intelligence (CogMI), pp. 150-158, 2023.
- [4] Y. Wu, et al., *Graph Neural Networks for Anomaly Detection in Industrial Internet of Things*. IEEE Internet Things J., vol. 9, no. 12, pp. 9214–9231, Jun. 2022.
- [5] C. Guilloteau, et al., *Generative Networks for emulating synthetic sky images*. Kavli Summer Program in Astrophysics, 2019.
- [6] Z. Chen et al., *Graph neural network-based fault diagnosis: a review*, arXiv: arXiv:2111.08185. Nov. 15, 2021.