Poster Abstract: Graph Learning on Cyber-Physical Knowledge Graphs

Avia Anwar
avia_anwar@mines.edu
Colorado School of Mines
Golden, Colorado, USA

Umut Mete Saka
saka@mines.edu
Colorado School of Mines
Golden, Colorado, USA

Gabe Fierro
gtfierro@mines.edu
Colorado School of Mines
Golden, Colorado, USA

Abstract

Knowledge graphs (KGs) are machine-readable representations of cyber-physical systems (CPS) which can be incomplete due to the size and complexity of CPS. Knowledge graph completion (KGC) models can predict missing edges, but perform poorly and fail to generalize on CPS KGs due to the high heterogeneity and small size of those graphs. In this work, we propose an ontology-informed heterogeneous Graph Neural Network (GNN) architecture that integrates hierarchical parent-class layers to enhance generalization in CPS KGs. Our approach outperforms traditional heterogeneous GNNs and Node2Vec-based methods in edge prediction tasks, offering a promising solution for CPS KGs.

CCS Concepts

• **Computer systems organization** → *Embedded and cyber-physical systems*; • **Computing methodologies** → **Knowledge representation and reasoning**; **Neural networks**.

Keywords

ontology, graph neural networks, cyber-physical systems

ACM Reference Format:
Avia Anwar, Umut Mete Saka, and Gabe Fierro. 2025. Poster Abstract: Graph Learning on Cyber-Physical Knowledge Graphs. In *The 23rd ACM Conference on Embedded Networked Sensor Systems (SenSys '25), May 6–9, 2025, Irvine, CA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3715014.3724049

1 Introduction

Cyber-physical systems (CPS) use embedded sensing and controls to optimize operations and perform anomaly detection and fault diagnosis. CPS lack structured machine-readable representations despite their data-rich nature, making it difficult to develop new applications. Ongoing research into knowledge graphs (KG) aims to standardize CPS representations and simplify software configuration for key applications like anomaly detection and model-predictive control [1, 5]. KGs are directed, labeled graphs that model relationships (edges) between entities and properties (nodes). Ontologies define the structure and vocabulary of KGs to unify labels and enable interoperability between multiple KGs. Fig. 1 shows an example CPS modeled with the Brick ontology [1]. Creating KGs is challenging due to the lack of structured metadata in CPS and its inherent heterogeneity. [2] and automated methods often miss nodes or omit edges, resulting in incomplete KGs.

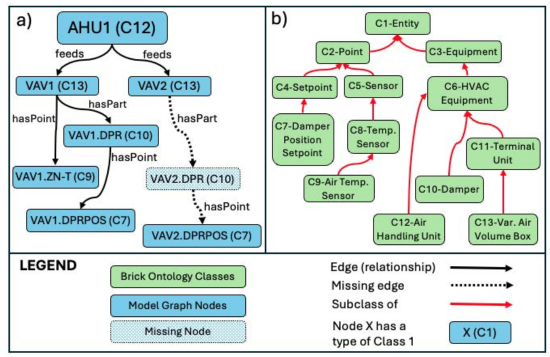

Figure 1: (a) A cyber-physical knowledge graph representing a segment of an HVAC system in a building. The nodes in the model are defined based on an ontology, as illustrated in (b). For example, AHU1 is an instance of the "Air Handling Unit" class. (b) The ontology defines the class hierarchy.

Knowledge graph completion (KGC) [3] techniques have potential to improve incomplete KGs. Graph Neural Networks (GNNs) [8] are effective for CPS tasks such as object detection and anomaly detection [4], but have not been applied to CPS KGs for completion tasks. CPS KGs face challenges during knowledge representation and model construction due to their small size and high heterogeneity (§2), meaning multiple node and edge types. These characteristics of CPS KGs necessitate specialized GNN architectures to efficiently handle ontology complexity and scalability. To address these challenges, we develop a custom heterogeneous GNN architecture that incorporates ontology-aware knowledge, as detailed in §3. Our model enhances generalization for low-density data by introducing parent-class layers, allowing nodes to inherit information from their hierarchical structure. Our evaluation in §4 shows that ontology-aware heterogeneous architecture predicts missing relationships with higher accuracy.

2 Challenges for Graph Learning in CPS

The first step in applying KGC is knowledge representation, or embedding, where the graph is represented as vectors to be used in GNNs. These embeddings are learned from individual triples or graph topology, often using random walk-based methods, such as Node2Vec [6]. While topology-based methods perform better, they rely on large and rich neighborhoods typical of general knowledge bases which contain millions of nodes and edges. CPS KGs

SenSys '25, May 6–9, 2025, Irvine, CA, USA

ACM ISBN 979-8-4007-1479-5/25/05
https://doi.org/10.1145/3715014.3724049

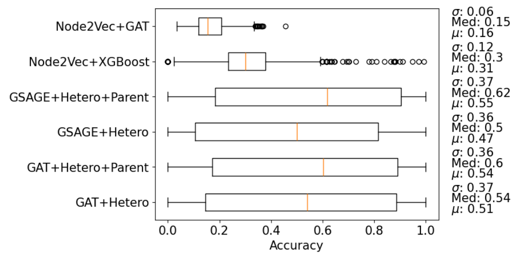

Figure 2: Comparison of accuracy across models

are significantly smaller [9] and highly site-specific with varying structures and relationships. Existing datasets do not sufficiently cover the relations needed for a GNN to generalize to unseen KGs.

Another consideration is KG heterogeneity, or the diversity of the edge and node types in the graph. GNNs for heterogeneous KGs require distinct subnetworks for each combination of edge and node class, allowing models to treat nodes and edges differently based on context. Each CPS KG has different classes and edges present, depending on what assets are deployed in the cyber-physical environment. Our analysis of CPS KGs shows that these graphs can have 30x-450x *more* relation types than existing knowledge graph datasets. Large number of classes and relation types increases the size of heterogeneous GNNs and can impact training time if each type gets its own subnetwork [7]. For instance, Brick [1] contains 1253 classes and 54 edge labels, creating 42M possible relation types.

3 Ontology-Informed Heterogeneous GNNs

The key challenge in CPS KGs is predicting unseen relation types in the test graph. To address this, we develop a custom heterogeneous GNN architecture that incorporates ontology-aware knowledge. A knowledge graph consists of set of triples (s, p, o) and a relation type (C_s, p, C_o) is defined by a subject class C_s, an edge label p, and object class C_o. Traditional heterogeneous GNNs model treat each relation type independently, failing to capture the ontology's hierarchical structure. Our architecture enhances the network by incorporating parent class information from the ontology, creating new relation based on the parents to generalize across unseen KGs (Figure 3). This allows the model to learn from the hierarchical class structure (Figure 1) of the ontology reducing dependency of the presence of a specific entity type in the training set.

We expand GNN sub-networks to include parent-class combinations, improving generalization but increasing the average network size from 347.22 layers, 1.1×10^8 parameters to 555.16 layers, 1.7×10^8 parameters with a 2x increase in the training time.

4 Experimental Evaluation

We conduct experiments using the Mortar dataset [5], consisting of 45 independent CPS KGs, to test how well our model performs in link prediction tasks. We use each graph to train an independent model, and test all models across each of the 45 graphs. We compare Node2Vec-based models with two heterogeneous GNN architectures (GraphSAGE and GAT), both with and without our ontology-aware structure.

Fig. 2 shows that GraphSAGE with ontology-aware featurization is the best-performing model with a median accuracy of 62%. Our results suggest that ontology-aware heterogeneous GNNs improve model performance over traditional methods like Node2Vec.

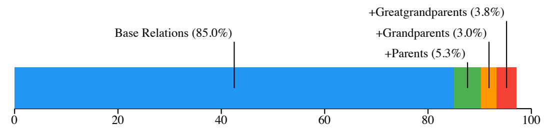

Figure 3: Adding GNN layers for super classes increases coverage on unseen graphs.

5 Conclusion

In this work, we address the challenge of knowledge graph completion in CPS KGs by developing an ontology-informed heterogeneous GNN architecture. By incorporating parent-class layers, our approach improves link prediction accuracy and enhances generalization in sparse CPS KGs. Experimental results show that our method outperforms traditional graph learning techniques, particularly in handling missing relationships. As future work, we will explore node prediction in CPS KGs to fully address the KGC problem in this domain to improve accuracy.

Acknowledgements

This material is based upon work supported by a NIST PREP fellowship and the National Alliance for Water Innovation (NAWI), funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Industrial Efficiency and Decarbonization Office, under DE-FOA-0001905.

References

[1] Bharathan Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, Mario Berges, David Culler, Rajesh Gupta, Mikkel Baun Kjærgaard, Mani Srivastava, and Kamin Whitehouse. 2016. Brick: Towards a Unified Metadata Schema For Buildings. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, Palo Alto CA USA, 41–50. https://doi.org/10.1145/2993422.2993577

[2] Arka A. Bhattacharya, Dezhi Hong, David Culler, Jorge Ortiz, Kamin Whitehouse, and Eugene Wu. 2015. Automated Metadata Construction to Support Portable Building Applications. In *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments - BuildSys '15*. ACM Press, Seoul, South Korea, 3–12. https://doi.org/10.1145/2821650.2821667

[3] Zhe Chen, Yuehan Wang, Bin Zhao, Jing Cheng, Xin Zhao, and Zongtao Duan. 2020. Knowledge Graph Completion: A Review. *IEEE Access* 8 (2020), 192435–192456. https://doi.org/10.1109/ACCESS.2020.3030076

[4] Guimin Dong, Mingyue Tang, Zhiyuan Wang, Jiechao Gao, Sikun Guo, Lihua Cai, Robert Gutierrez, Bradford Campbel, Laura E. Barnes, and Mehdi Boukhechba. 2023. Graph Neural Networks in IoT: A Survey. *ACM Transactions on Sensor Networks* 19, 2 (May 2023), 1–50. https://doi.org/10.1145/3565973

[5] Gabe Fierro, Marco Pritoni, Moustafa Abdelbaky, Daniel Lengyel, John Leyden, Anand Prakash, Pranav Gupta, Paul Raftery, Therese Peffer, Greg Thomson, and David E. Culler. 2020. Mortar: An Open Testbed for Portable Building Analytics. *ACM Transactions on Sensor Networks* 16, 1 (Feb. 2020), 1–31. https://doi.org/10.1145/3366375

[6] Aditya Grover and Jure Leskovec. 2016. Node2vec: Scalable Feature Learning for Networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, San Francisco California USA, 855–864. https://doi.org/10.1145/2939672.2939754

[7] Michael Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2017. Modeling Relational Data with Graph Convolutional Networks. arXiv:1703.06103 [stat]

[8] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2021. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 32, 1 (Jan. 2021), 4–24. https://doi.org/10.1109/TNNLS.2020.2978386

[9] Yanhui Xu and Dan Wang. 2022. Understanding the Willingness to Share Building Data by a Social Study Based on Privacy Calculus Theory. In *Proceedings of the 9th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. ACM, Boston Massachusetts, 59–68. https://doi.org/10.1145/3563357.3564059