CCC 2024 Proceedings of the Creative Construction Conference (2024)
Edited by: Miroslaw J. Skibniewski & Miklós Hajdu
https://doi.org/10.3311/CCC2024-056

LEVERAGING KNOWLEDGE GRAPHS TO ENHANCE FAULT DETECTION IN FACILITY MANAGEMENT

Arsenii Kirillov[1]
Alessandro Carbonari[1], Žiga Turk[2], Alberto Giretti[1]
1 Polytechnic University of Marche, Department of Civil and Building Engineering and Architecture (DICEA), 60131 Ancona, Italy
2 University of Ljubljana, Faculty of Civil and Geodetic Engineering, Jamova cesta 2, Ljubljana, 1000, Slovenia

Abstract
Digital twins are the most commonly used tool for improving efficiency in facilities management. However, existing digital twins lack semantics, leaving the facility maintenance team responsible for interpreting and responding to faults. To enable semantics in a digital twin, it must rely not only on the data produced by sensors, but also on a deeper knowledge of the system and the processes taking place within it. The paper proposes a framework for the automated generation of Bayesian Networks (BNs) from a single data source - a knowledge graph - which should store information from different sources, such as topology, documents originally written in natural language, and domain-specific ontologies based on RDF (Resource Description Framework). BNs will be used to infer failure symptoms and causes, while automated BN generation is expected to solve a scalability problem. These coupled tools will be investigated in terms of supporting the facility manager in decision making.

Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2024.

Keywords: Bayesian network generation, decision-making, facility management, fault detection, knowledge graph.

1. Introduction

Digital twins in the facility management are supposed to help facility managers in optimization of use and operation of buildings. However modern building management systems (BMS) are managing vast amounts of data, but interpretation of that data and decision making is under the responsibility of the facility management team.

Existing digital twin maintenance approaches mostly rely on reactive maintenance. The state of the art of the digital twin implementation made by Hakimi et. al [4] points out that prediction of faults from data obtained from sensors with algorithms or AI is still an unaddressed challenge due to lack of semantics of data. In the research conducted by Hosamo et. al [5] it's also noticed the lack of semantic part of data and missing of human knowledge in the digital twins. Addressing the gap of semantics in a digital twin will be a step to achieve a recommended strategy of Reliability Centered Maintenance which relies on the concept of equipment operation requirements definition based on its operation context [6].

Solution proposed by this paper contributes to a digital twin of a facility development which will be capable of semantic sensor data representation unified with the facility manager's knowledge. More specifically the knowledge graph composed from different data sources is proposed as a core entity of the framework.

The scope of the proposed digital twin in this paper will be limited to HVAC systems. Modern buildings use HVAC (Heating Ventilation and Air Conditioning) systems to maintain the indoor environment at a comfortable level for the occupants. In order to increase the efficiency of the system, various sensors are usually installed to assist the maintenance team. If some of the parameters are outside the desired range, the facility management team will react according to established protocols to reduce the impact of the fault and return the system to normal. Failures and inefficiencies can lead to various consequences from energy inefficiency to harm for occupants' wellness. Therefore failures need to be diagnosed in order to carrying on an effective restoration and prevent them in the future.

In the next sections we will present most relevant state of the art and conceptual framework and role and technical features of the knowledge graph

2. State of the art

Fault detection and diagnosis (FDD) methods can be divided into 3 categories: Model-based, data-driven, and rule-based or knowledge-based [7]. Model-based is highly dependent on mathematical equations which are hard to change according to the state of a system. Data-driven usually lacks semantics and relations between parts of the system for fault interpretation. Knowledge-based approaches require efforts for establishment and in the construction industry such method remains heuristic. Researchers are trying to achieve a framework which will be suitable for construction industry realities such as lack of good quality sensor data and lack of instructions on fault diagnosis.

FDD process can be divided into the following stages: Fault detection, Fault isolation, Fault identification, and Fault classification, as shown in Figure 1 . According to the research of Sahu et. al [1] on data-driven fault diagnosis methods for industrial equipment the final step of fault evaluation is the most significant because based on the evaluation result the decision on system repairings or improvements will be made.

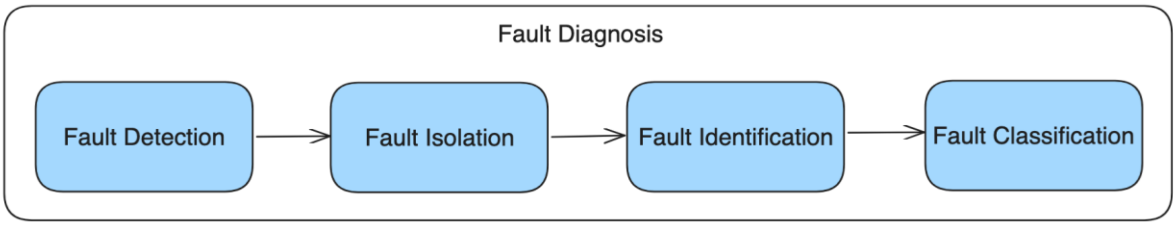

Fig. 1. Fault Diagnosis Steps;

Since data-driven FDD lacks semantics in next sections of the paper a framework to fulfil the FDD with semantics will be proposed.

2.1. Fault detection and diagnosis case studies in the construction industry

In the research conducted by Sanzana et. al [2] on application of deep learning techniques for HVAC systems optimization it’s pointed out a great potential of this approach which is unfortunately limited by lack of good quality labelled data and large amount of data input required which is usually hard to obtain in the construction industry. To address the problem of lack of semantics in FDD Mallak et. al developed an Ontology for a sensor [7].

A novel approach of detecting faults using semantic model of a building was performed by Roa et. al [3]. The semantic model of a building encoded in Brick ontology was used to achieve scalability of data-driven fault detection methods by replacing vendor specific sensors namings with a SPARQL query which is supposed to be executed upon the Brick model of any building. As a result data-driven fault detection can be scaled across many buildings that have a brick representation. Besides that, the semantic Brick model of a building can be used further to achieve next steps of FDD illustrated in Figure

1 since it captures the key missing component of existing digital twins which is the semantic relation between HVAC system components.

The gap that this paper addresses is to develop a framework for fault diagnosis which is not limited to fault detection, rather it goes beyond and includes reasoning about the fault symptoms and causes based on fusion of data-driven and knowledge-based approaches.

2.2. Fault detection and diagnosis case studies in other sectors

The lack of semantics problem has been addressed in other industries. In the case study implemented by Qu et. al [8] on decision making for power communication equipment, a graph-driven approach was proposed consisting of a knowledge extraction step and a learning recommendation layer based on a convolutional neural network. The research on using a knowledge graph for aircraft fault diagnosis, consisting of a deep learning method for extracting knowledge from heterogeneous sources and a question answering system, was proposed by Tang et. al. [9]. In this research, the simple fault ontology was constructed and various LLMs were used for knowledge extraction. For fault diagnosis in the industrial process of steel production, the reinforcement learning framework was introduced by Han et. al. [10] to complete the fault knowledge graph. In the case study of fault knowledge graph construction of robot transmission system fault diagnosis implemented by Deng et. al. [11], the fault ontology was constructed using a top-down approach and knowledge extraction was performed using various machine learning techniques. The final knowledge graph was stored in the Neo4j graph database. However, the use cases of the obtained knowledge graph weren't specified. The research conducted by Li et. al [12] proposes an integration of LLM with KG to improve the accuracy of answers to user queries in order to improve fault diagnosis. Examples from related industries show that the ability of knowledge graphs to store data from different data sources is quite suitable for the fault diagnosis paradigm.

3. Conceptual framework

The lack of explicit description of faults in the construction industry should be considered when developing a knowledge graph. Besides that the solution in case of uncertainty of connection between fault symptom and fault cause must be addressed. Usage of knowledge graph as a digital twin enables conversion to different structures that allow us to perform reasoning. Bayesian Networks (BN) is the most commonly used tool for inference in data-driven fault detection [1]. Therefore digital twin must be developed in a way that supports automated BN generation.

3.1. Overall schema

The knowledge usage schema is illustrated in Figure 2. Knowledge graph is supposed to be stored in a Neo4j graph database. The fault knowledge will be modelled as an extension of the Brick ontology that was specifically designed for BEMS and buildings’ technical systems. The numeric data from sensors is supposed to be stored separately and to be used for BN calibration. The BN which is supposed to be generated from KG must rely on both sensor and evidence data. As for the future research, integration of KG with Large Language Models (LLMs) will be investigated.

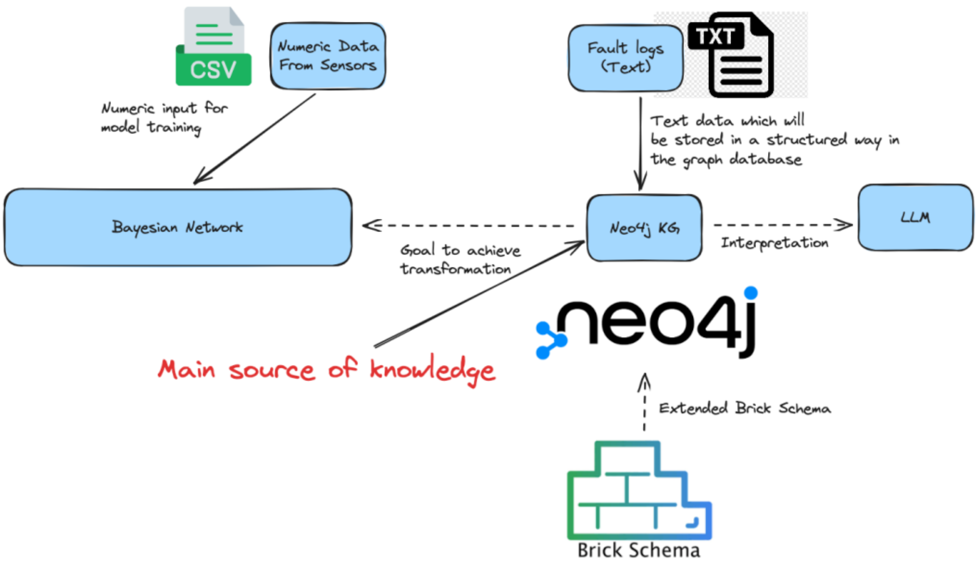

Fig. 2. Knowledge usage schema;

3.2. Knowledge graph development

The knowledge graph addresses the problem of lack of semantics in digital twins. Construction industry offers Industry Foundation Class (IFC) as an open and semantically-rich data format. However IFC doesn't have specific classes that could represent faults and has sophisticated structure itself. Whereas, Brick ontology is lightweight and addresses the HVAC domain directly, which is a scope of this paper. Brick model of a building has already defined classes of equipment and relationships between them. Therefore, the choice of Brick is preferable. However Brick ontology also isn't capable of representing faults. In the research implemented by Luo et. al [13] of Brick ontology extension in order to represent a metadata of occupants it showcased the possibility of modifying Brick according to the desired goals.

The knowledge about possible faults to occur in the HVAC system of the engineering department of the Polytechnic University of Marche was conducted from facility maintenance personnel and is stored in the plain text format. In the future work this information will be structured according to the defined extended Brick ontology and stored in the graph database same as it was done in studies [8, 9, 10].

3.3. Bayesian Network establishment

Bayesian networks (BN) are made of an acyclic graph structure and associated conditional probability tables. Among their potentials, we mention that inferences can be updated as a consequence of new evidence and diagnostic reasoning is allowed thanks to backward propagation. BN is one of the most efficient tools for FDD. In the research implemented by Cai et. al [14] it was shown the accuracy performance of 2 layers BN built based on data measured by sensors and visual inspection by FM personnel. Research implemented by Yang et. al. [15] introduces an approach of dynamic Bayesian network generation based on fault knowledge graph. It addresses the problem of uncertainty in the relationships between fault symptoms and fault causes.

Possibility of automated or semi-automated BN generation from a KG and capability of KG to store data from heterogeneous data sources, makes such a combination capable of addressing a gap of lack of semantic in a digital twin and uncertain relations between fault symptoms and fault cases. After the initial BN is generated from a KG, it is supposed to be retrained based on the numeric data from sensors.

4. Future research

The proposed framework hasn't been proved practically. However the Brick model of a 2 room facility of Polytechnic university of Marche was developed based on Dymola simulations model. The Brick model was developed manually using python by visual inspection of the Dymola model converting it into RDF triples. The model was serialised in ".owl" format. The T-Box schema of a serialised model is illustrated in Figure 3. This schema will be extended in order to be capable of representing possible faults and knowledge stored according to the defined schema and stored in Neo4j graph database will be tested for BN extraction by executing Cypher queries as it was done in the research work cited in [15].

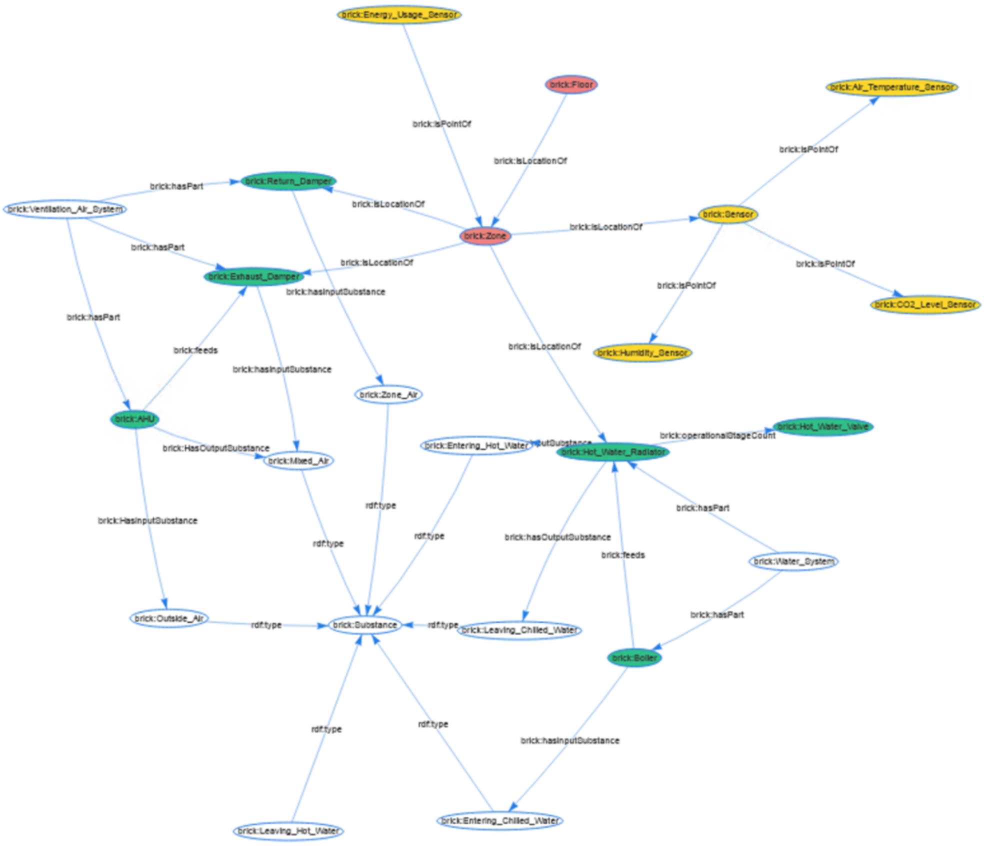

Fig. 3. T-Box of Brick classes used;

After the constraints for dynamic BN generation out of KG will be satisfied, knowledge modelling constraints for LLM interpretation should be investigated. Future role of LLM is assumed as a front end for the facility manager which can help him to manage vast amounts of information.

5. Conclusion

Existing digital twins have a significant limitation of lack of semantics of data that they represent. This makes facility management process highly dependent on a facility maintenance team because they're capable to interpret the information provided by digital twin. This gap can be addressed with semantically-rich digital twin such as a knowledge graph. However the realities of construction industry put specific constraints on knowledge graph establishment such as lack of explicit fault description and connections between fault symptoms and fault causes. Therefore, knowledge-based digital twins must be capable of establishing such connections. To address this issue the concept of dynamic bayesian network out of knowledge graph was introduced which will affect the knowledge modelling schema. As for the knowledge modelling schema, it was decided to extend existing lightweight Brick ontology which is capable for semantic representation of HVAC systems of the buildings in order to represent faults. For the future research and experiments semantic Brick ontology model of 2 rooms of Polytechnic University of Marche was developed and fault information was conducted from the facility maintenance team.

References

[1] A. R. Sahu, S. K. Palei, and A. Mishra, ‘Data-driven fault diagnosis approaches for industrial equipment: A review’, Expert Systems, vol. 41, no. 2, p. e13360, Feb. 2024. https://doi.org/10.1111/exsy.13360.
[2] M. R. Sanzana, T. Maul, J. Y. Wong, M. O. M. Abdulrazic, and C.-C. Yip, ‘Application of deep learning in facility management and maintenance for heating, ventilation, and air conditioning’, Automation in Construction, vol. 141, p. 104445, Sep. 2022. https://doi.org/10.1016/j.autcon.2022.104445.
[3] C. D. Roa, P. Raftery, R. Singla, M. Pritoni, T. Peffer, ‘Detecting Passing Valves at Scale Across Different Buildings and Systems: A Brick Enabled and Mortar Tested Application’, 2022. https://doi.org/10.20357/B7VP5H.
[4] O. Hakimi, H. Liu, and O. Abudayyeh, ‘Digital twin-enabled smart facility management: A bibliometric review’, Front. Eng. Manag., vol. 11, no. 1, pp. 32–49, Mar. 2024. https://doi.org/10.1007/s42524-023-0254-4.
[5] H. H. Hosamo, A. Imran, J. Cardenas-Cartagena, P. R. Svennevig, K. Svidt, and H. K. Nielsen, ‘A Review of the Digital Twin Technology in the AEC-FM Industry’, Advances in Civil Engineering, vol. 2022, pp. 1–17, Mar. 2022. https://doi.org/10.1155/2022/2185170.
[6] G. P. Sullivan, R. Pugh, and A. P. Melendez, ‘Operations and Maintenance Best Practices--A Guide to Achieving Operational Efficiency’, PNNL-13890, 15010224, Dec. 2002. https://doi.org/10.2172/15010224.
[7] A. Mallak, C. Weber, M. Fathi, A. Behravan, and R. Obermaisser, ‘A Graph-Based Sensor Fault Detection and Diagnosis for Demand-Controlled Ventilation Systems Extracted from a Semantic Ontology’, in 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), Las Palmas de Gran Canaria: IEEE, Jun. 2018, pp. 000377–000382. https://doi.org/10.1109/INES.2018.8523895.
[8] H. Qu, Y. Zhang, K. Liang, S. Li, and X. Huo, ‘A Knowledge-Graph-Driven Method for Intelligent Decision Making on Power Communication Equipment Faults’, Electronics, vol. 12, no. 18, p. 3939, Sep. 2023. https://doi.org/10.3390/electronics12183939.
[9] X. Tang, G. Chi, L. Cui, A. W. H. Ip, K. L. Yung, and X. Xie, ‘Exploring Research on the Construction and Application of Knowledge Graphs for Aircraft Fault Diagnosis’, Sensors, vol. 23, no. 11, p. 5295, Jun. 2023. https://doi.org/10.3390/s23115295.
[10] H. Han, J. Wang, X. Wang, and S. Chen, ‘Construction and Evolution of Fault Diagnosis Knowledge Graph in Industrial Process’, IEEE Trans. Instrum. Meas., vol. 71, pp. 1–12, 2022, https://doi.org/10.1109/TIM.2022.3200429.
[11] J. Deng, T. Wang, Z. Wang, J. Zhou, and L. Cheng, ‘Research on Event Logic Knowledge Graph Construction Method of Robot Transmission System Fault Diagnosis’, IEEE Access, vol. 10, pp. 17656–17673, 2022. https://doi.org/10.1109/ACCESS.2022.3150409.
[12] Y. Li, Y. He, R. Lian, and Q. Guo, ‘Fault Diagnosis and System Maintenance Based on Large Language Models and Knowledge Graphs’, in 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), Hangzhou, China: IEEE, Dec. 2023, pp. 589–592. https://doi.org/10.1109/RICAI60863.2023.10489566.
[13] N. Luo, G. Fierro, Y. Liu, B. Dong, and T. Hong, ‘Extending the Brick schema to represent metadata of occupants’, Automation in Construction, vol. 139, p. 104307, Jul. 2022. https://doi.org/10.1016/j.autcon.2022.104307.
[14] B. Cai et al., ‘Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network’, Applied Energy, vol. 114, pp. 1–9, Feb. 2014. https://doi.org/10.1016/j.apenergy.2013.09.043.
[15] N. Yang, G. Zhang, and J. Wang, ‘Research on Knowledge Graph and Bayesian Network in Fault Diagnosis of Steam Turbine’, in 2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai), Shanghai, China: IEEE, Oct. 2020, pp. 1–6. https://doi.org/10.1109/PHM-Shanghai49105.2020.9281007.