Knowledge Representation and Engineering for Smart Diagnosis of Cyber Physical Systems

Ameneh **Naghdi Pour¹, Benno Kruit¹, Jieying Chen¹, Godfried Webers², Peter Kruizinga³** and Stefan Schlobach*¹*

¹Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, NL ³Canon Production Printing Netherlands, Van der Grintenstraat 1, 5914 HH Venlo, NL 2 Philips Medical Systems International, Veenpluis 6, 5684 PC Best, NL

Abstract

The traditional maintenance approach used by manufactures such as Canon and Philips, which relies on service engineers' expertise to diagnose the failure causes, poses significant challenges and costs. To overcome challenges and minimize expenses, we aim to explore the construction and application of a fault knowledge graph that integrate different maintenance data and knowledge sources. This involves developing an upper-level ontology based on the requirements for the fault diagnosis of Cyber Physical Systems. This ontology draws inspiration from the Industrial Domain Ontology (IDO) and Industrial Ontology Foundry-Maintenance Reference ontology (IOF-MRO). By leveraging these two ontologies as foundation, we aim to construct a comprehensive framework that captures and represents fault-related knowledge in a structured manner. Additionally, this paper envisions the integration of different fault diagnosis methods and knowledge modeling techniques. Combining these approaches is expected to enhance the accuracy and effectiveness of fault diagnosis, leading to more efficient and reliable solutions.

Keywords

Fault diagnosis, cyber-physical systems, knowledge representation, knowledge graph

1. Introduction

Machine breakdowns result in a substantial financial burden for manufacturers and their customers, primarily due to expenses related to training service engineers for fault diagnosis, their salaries, and the provision of spare parts [\[1\]](#page--1-0). Additionally, the capacity of machines is adversely affected, as they remain out of service during downtime, further increasing costs and impacting productivity. Therefore, efficient fault diagnosis is of utmost importance for minimizing both downtime and costs. At Canon and Philips manufactures, service engineers are trained using valuable documentation and occasionally videos. However, significant challenges raised with this approach. These challenges include heavy reliance on the engineers' expertise, the complexity of navigating extensive documentation, and the high costs associated with producing and providing instructional videos in terms of time and resources.

To address these challenges, the Zorro project has been initiated, focusing on achieving zero downtime in Cyber-Physical Systems (CPS). Our goal in this project is to support service engineers in diagnosis the failure cause. Figure 1 shows the proposed framework designed to achieve this goal. The input comprises various sources that have been identified through interviews conducted with authorities in the manufacturing field. In the first phase, a Knowledge Graph (KG) is constructed based on the input. This involves manually developing an upper-level ontology, inspired by IDO and IOF-MRO. In the next phase, service engineers observe a symptom that need to be converted into a query for the knowledge

SemIIM'24: Third International Workshop on Semantic Industrial Information Modelling, co-located with ISWC'24, 12th November 2024, Baltimore, USA

 \bigcirc a.naghdipour@vu.nl (A. Naghdi Pour); b.b.kruit@vu.nl (B. Kruit); j.y.chen@vu.nl (J. Chen); godfried.webers@philips.com (G. Webers); peter.kruizinga@cpp.canon (P. Kruizinga); k.s.schlobach@vu.nl (S. Schlobach)

GLOBE <https://https://github.com/Ameneh71> (A. Naghdi Pour); <https://github.com/bennokr/> (B. Kruit); <https://github.com/JieyingChenChen> (J. Chen)

Orcid [0000-0002-7357-7906](https://orcid.org/0000-0002-7357-7906) (A. Naghdi Pour); [0000-0002-3282-1597](https://orcid.org/0000-0002-3282-1597) (B. Kruit); [0000-0002-3282-1597](https://orcid.org/0000-0002-3282-1597) (J. Chen); [0000-0002-3282-1597](https://orcid.org/0000-0002-3282-1597) (P. Kruizinga); [0000-0002-3282-1597](https://orcid.org/0000-0002-3282-1597) (S. Schlobach)

^{© 2024} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

graph reasoning. The KG then provides a root cause analysis and the corresponding repair procedure as the output.

In addition to the proposed framework, we have envisioned several methods for fault diagnosis that can offer significant advantages. Firstly, the automatic generation of Bayesian Networks (BNs) from KG. The KG excels in data integration, providing a unified view of system faults, while BNs enable probabilistic reasoning for more accurate diagnosis. This integration allows for a more precise understanding of fault scenarios and their likelihood, leading to improved diagnostic outcomes. Secondly, integrating KG with fault tree provides a means to capture complex causal mechanisms. This integration also facilitates the generation of BNs. Moreover, we envision that KG can be constructed with minimal effort using Language Models (LLMs). These LLMs can facilitate the integration of different fault diagnosis representations, such as BNs and Fault Trees, thereby enhancing the robustness of CPS by minimizing downtime. By leveraging the power of this method, we can work towards achieving zero downtime and ensuring the continuous operation of CPS in the future.

2. Related work

There are various sources of knowledge and data that can be utilized to support service engineer in diagnosing failure. These sources include machine log data from sensors, documented knowledge provided by authorities, logbook data written by service engineers, and also tacit knowledge in the brain of experts. Researchers have extensively explored different categories of these sources and proposed various methods for supporting maintenance tasks. These methods can be broadly categorized into four groups: (1) Model-based approach, which is based on constructing an accurate physical model that captures the dynamic changes within the system. It diagnoses faults by comparing the actual measurement output of the system with the predicted output of the model, relying on consistency monitoring. While this method offers high accuracy, it necessitates the establishment of precise and quantitative physical models [\[2\]](#page-6-0). (2) Signal-based approach, which diagnoses faults by analyzing and processing monitored signals. However, it can only be used to identify fault types in key components that are equipped with sensors [\[3\]](#page-6-1). (3) Quantitative-knowledge-based method utilize historical data to train a model and treat fault diagnosis as a pattern recognition problem. However, it does rely on a significant amount of historical fault data to train the model effectively[\[4,](#page-6-2) [5\]](#page-6-3). (4) Qualitative knowledgebased method which adopts a different approach by constructing a qualitative model that describes prior knowledge related to faults. It then utilizes techniques such as searching, matching, and reasoning to diagnose faults within the system. Unlike the model-based and quantitative-based approaches, this method does not require the construction of accurate physical models or rely heavily on extensive historical fault data. Instead, it needs to build a fault knowledge base [\[6\]](#page-6-4). Therefore, considering the requirements of system-level analysis and the need for searching, matching, and reasoning capabilities, the qualitative-knowledge-based fault diagnosis approach appears to be a suitable choice for this project.

A crucial aspect of qualitative methods is the requirement for a well-defined and reasonable model that accurately represents knowledge related to faults. The most popular models in the literature include Rule-Based [\[7\]](#page-6-5), Causal Models [\[8\]](#page-6-6), Fault Trees [\[9\]](#page-6-7), and Petri Nets [\[10\]](#page-6-8). Rule-based models use if-then rules to represent expert knowledge about system behavior and fault conditions, allowing for automated diagnosis and decision-making under predefined conditions. Causal models diagnose faults in CPS by mapping cause-and-effect relationships between system components and events. Fault trees represent different events that can lead to a specific system failure, helping to identify potential causes and their relationships. Nevertheless, these models have certain limitations. They typically require prior analysis of potential equipment fault modes and involve manual editing, which makes them inflexible and challenging to update dynamically. Consequently, the systematic and timely sharing of maintenance engineers' experience and expertise in fault diagnosis becomes difficult [\[6\]](#page-6-4). Therefore, this paper proposes to use knowledge graph technology to mine fault knowledge from vast and diverse fault documents and then construct a structured and interconnected fault knowledge base.

Figure 1: Construction and application framework of domain fault knowledge graph

3. Methodology

Figure [1](#page-2-0) shows our proposed framework for fault diagnosis of CPS. The following subsections provide a detailed illustration of the framework.

3.1. Input Source

Following regular interviews with authorities at Canon and Philips, we have identified various maintenance sources and categorized them into three groups including data, documented knowledge, and tacit knowledge. Each of these sources offers unique insights and information about the system, but they also have limitations. Data, which is logbook written by service engineers, contains real-life information regarding specific problems and actions taken to resolve them. However, the information provided in the logbook may sometimes be incomplete. For instance, engineers may only write "done" as the action taken without providing detailed explanations. Documented knowledge encompasses three resources including (1) Bill Of Material, which provides information about the physical structure of the system, including the location of each part within assemblies or components, but lacks details on problems that may happen for each part of the system. (2) Failure Mode Effect Analysis, offers insights into challenging failure modes and their potential effects, but may not guarantee an accurate and comprehensive solution for failures. (3) Troubleshooting Manuals, provide insights into the causes of failures and offer remedies to resolve them. Lastly, tacit knowledge refers to undocumented knowledge residing in the minds of experts. These different sources are complementary, and relying on just one of them would not result in an efficient fault diagnosis. By combining them, we can leverage their strengths and compensate for their limitations.

3.2. Knowledge Graph Construction

The first phase involves two steps to construct the knowledge graph: *Upper-level ontology construction* and *Information Extraction*. The upper-level ontology has been manually created to serve as the foundational structure for organizing the schema of the knowledge graph. This involved extensive analysis of various input sources to identify the most valuable knowledge. Additionally, the concepts and terminology have been carefully selected to ensure clarity and consistency throughout the ontology. We also drew inspiration from the IDO and IOF-MRO in developing our fault diagnosis ontology. IDO is a recent work item, approved by the ISO TC 184/SC4 Industrial Data committee in July 2023, and holds significance in the realm of industrial standardization. IOF-MRO is to support semantic interoperability through the use of modular ontologies in the maintenance domain. Due to limited space, more detailed information on the IDO and IOF-MRO models can be found in [\[11\]](#page-6-9) and [\[12\]](#page-6-10) respectively. Figure [2](#page-3-0) shows the design of the Zorro ontology, where there are classes (nodes) and relationships (edges) that represent the domain of interest. There is a class named "Component" that can represent a part,

Figure 2: Construction and application framework of domain fault knowledge graph

assembly, or subsystem, each serving a specific "Function". When a part encounter malfunction, it leads to the occurrence of a "Problem". The problem has cause(s), which can be considered as problem, result in a certain "Effect". To address these problems, there are two types of "Procedure" that can be implemented. The first is a "Solution", which directly solves the problem at hand. The second is a "Workaround", which aims to address the effects caused by the problem. Both procedures consist of multiple "Step", with each step involving a specific part of the system. For a thorough understanding, Table [1](#page-4-0) provides a depiction of the classes, their definitions and the sources containing relevant information about each class, along with an example of the classes in real field. We also compared our defined classes with the IDO and IOF-MRO. On the other hand, Table [2](#page-5-0) illustrates the object properties, including their definitions and comparison. For the information extraction step, different techniques such as Regular Expression, Named Entity Recognition, and LLMs, have been utilized to extracted required knowledge and populate it based on the ontology.

3.3. Diagnosis

The next phase, diagnosis, shows the application of the proposed method in which service engineers observe symptoms of the failure, which should be converted to a query for KG-based reasoning. As a result, a diagnosis which find the root cause of the issue along with a procedure should be suggested to solve the problem. To this end, we are planning to develop querying and reasoning systems for diagnosis, with the aim of integrating or supporting different fault diagnosis reasoning techniques, such as BNs and Fault Trees. Crucially, such systems will need to take into account the maintenance engineers' expert knowledge in order to augment their ability. However, integrating their tacit knowledge into the diagnosis process is an open research question. We urge the community to engage in this challenge, leveraging semantic technologies to enhance fault diagnosis and ensure operational resilience.

4. Discussion

The related work section has extensively analyzed the advantages and disadvantages associated with various fault diagnosis methods and knowledge modeling techniques. Additionally, the rationale for constructing a knowledge graph as the underlying model to capture and represent prior knowledge has been explained. However, we believe that the combination of fault diagnosis methods and knowledge modeling techniques holds great potential for enhancing overall capabilities by leveraging the strengths

Table 1

Zorro Class	Definition and Source	Close Match	Example
Compo- nent	A 'physical object' of industrial equipment that could be a part, a component or assembly Source: BOM, Logbook, FMEA, TM	IDO: InAnimatePhysicalObject IOF-MRO: MaintainableMateri- alltem	DC servo mo- tor
Function	A capability that meets a design requirement Source: FMEA, TM, Logbook	IDO : Function IOF-MRO: MaintainableMateri- alltemRole	Feeding paper to printer
Problem	An issue that happened for the system Source: Logbook, FMEA, TM	$IDO: -$ IOF-MRO: FailedState	Overheating
Effect	An observable result of the problem on the system Source:FMEA	$IDO: -$ IOF-MRO: FailureEffect	Wear and tear
Procedure	Sequence of steps that can be taken to solve the problem Source: Logbook, FMEA, TM	$IDO: -$ IOF-MRO: PlannedProcess	Check and clean the cooling fan
Solution	Procedure to directly solve the root cause of the failure Source: Logbook, FMEA, TM	$IDO: -$ IOF-MRO: MaintenanceActivity	Check and the clean cooling fan
	Workaround Procedure that addresses the effect of the failure, not the root cause Source: Logbook, FMEA, TM	$IDO: -$ IOF-MRO: SupportingMainte- nanceActivity	Replace damaged components closer to motor
Step	Single action that need to be in the procedure Source: Logbook, FMEA, TM	$IDO: -$ IOF-MRO: MaintenanceWorkO- rderRecord	Replace

Classes in the Zorro Ontology

of each approach. We have identified three specific combinations that hold promise in enhancing fault diagnosis:

- 1. Automatically generating a Bayesian Network from our knowledge graph: data integration from various sources is a crucial task for efficient fault diagnosis. While BNs face challenges in data integration [\[13\]](#page-6-11), KG excel in this aspect, allowing for a unified and comprehensive view of the system's faults. On the other hand, BNs offer a robust framework for probabilistic reasoning, enabling more accurate diagnosis in the presence of uncertainty. By leveraging the probabilistic capabilities of BNs, we can enhance the fault diagnosis process, considering the likelihood of different fault scenarios and their potential impact on the system. This combination harnesses the strengths of both approaches, enabling a more comprehensive and accurate fault diagnosis methodology.
- 2. Integrating our knowledge graph representation with fault tree: The ontology now contains only simple causal links, but often there is a more complex causal mechanism. Sometimes, faults only occur when multiple events occur simultaneously, or under specific conditions (such as humidity). Other times, faults have separate, orthogonal causes. These structures are often modeled in fault trees. We envision that these structures can also be modeled in a KG, at various levels of granularity, and that Fault Trees can be generated from them. Additionally, if fault trees exist that model certain mechanisms, they could be useful sources of information to be linked and integrated into the KG. Then, their content can be used to generate Bayesian Networks as described above.
- 3. Combining model-based with quantitative fault diagnosis methods: In section [2,](#page-1-0) we have

Table 2 Properties in the Zorro Ontology

Property	Definition	IDO	IOF-MRO
hasFunc-	Every part, component, and subsystem of the ma-	hasFunc-	hasRole
tion	chine has a function	tion	
define	If the system does not function properly it defines a problem	create	describe
failsVia	System fails and initiate a problem		
hasCause	A problem has a cause which the cause itself is a problem		
resultIn	Problem results in an effect		precedes
address	Workaround just addresses the effect of the problem, not the root cause		
solve	the Solution directly solves the root cause of the prob- lem		
subPart	Each part may has some subparts	contain	hasRole
involve	Each step of the procedure involves a part of the system	$\overline{}$	
consistsOf	Each procedure consist of some steps		Step "islnputOf" Procedure

described the drawbacks of these two methods. Both exist on either ends of knowledge-to-data spectrum: on one end, the model-based diagnosis requires detailed (physical) knowledge of the system, and on the other end, quantitative fault diagnosis requires comprehensive observation data to learn from. However, in most cases only a little of either is available, and it is difficult to integrate the physical models with the observed data. We envision that the extraction of fault diagnosis knowledge from diverse sources will aid the construction of models that are both physically informed and data-driven. To this end, the Knowledge Graph facilitates the integration of various views of historically collected (sensor) data within unified physical, logical and functional views of the CPS. We envision that this will significantly reduce the efforts needed to construct accurate models, by re-using modular model components, simplifying the data analysis pipeline, and allowing efficient validation on historical maintenance data.

5. Conclusion and Future work

In this study, various fault diagnosis approaches were reviewed, and a framework consisting of two primary phases, Knowledge Graph Construction and Diagnosis, was developed. By constructing a knowledge graph using diverse input sources, accurate diagnosis and efficient repair procedures can be achieved. In future work, the focus will be on refining and automating the construction of the knowledge graph using advanced techniques like LLMs. Additionally, the transformation of KG into Bayesian Network will be explored to have a probabilistic reasoning. Furthermore, the construction of model-based and quantitative-based fault diagnosis methods using the knowledge graph will also be investigated.

Acknowledgments

This publication is part of the project ZORRO with project number KICH1.ST02.21.003 of the research programme Key Enabling Technologies (KIC) which is (partly) financed by the Dutch Research Council (NWO).

References

- [1] D. S. Thomas, The costs and benefits of advanced maintenance in manufacturing, US Department of Commerce, National Institute of Standards and Technology, 2018. doi:[https://doi.org/10.](http://dx.doi.org/https://doi.org/10.6028/NIST.AMS.100-18) [6028/NIST.AMS.100- 18](http://dx.doi.org/https://doi.org/10.6028/NIST.AMS.100-18).
- [2] S.-H. You, Y. M. Cho, J.-O. Hahn, Model-based fault detection and isolation in automotive yaw moment control system, International Journal of Automotive Technology 18 (2017) 405–416. doi:[https://doi.org/10.1007/s12239- 017- 0041- 5](http://dx.doi.org/https://doi.org/10.1007/s12239-017-0041-5).
- [3] K. Zhang, Y. Xu, Z. Liao, L. Song, P. Chen, A novel fast entrogram and its applications in rolling bearing fault diagnosis, Mechanical Systems and Signal Processing 154 (2021) 107582. doi:[https://doi.org/10.1016/j.ymssp.2020.107582](http://dx.doi.org/https://doi.org/10.1016/j.ymssp.2020.107582).
- [4] J. Dong, D. Su, Y. Gao, X. Wu, H. Jiang, T. Chen, Fine-grained transfer learning based on deep feature decomposition for rotating equipment fault diagnosis, Measurement Science and Technology 34 (2023) 065902. doi:[https://doi.org/10.1088/1361- 6501/acc04a](http://dx.doi.org/https://doi.org/10.1088/1361-6501/acc04a).
- [5] J. Long, Y. Chen, Z. Yang, Y. Huang, C. Li, A novel self-training semi-supervised deep learning approach for machinery fault diagnosis, International Journal of Production Research 61 (2023) 8238–8251. doi:[https://doi.org/10.1080/00207543.2022.2032860](http://dx.doi.org/https://doi.org/10.1080/00207543.2022.2032860).
- [6] X. Tang, G. Chi, L. Cui, A. W. Ip, K. L. Yung, X. Xie, Exploring research on the construction and application of knowledge graphs for aircraft fault diagnosis, Sensors 23 (2023) 5295. doi:[https:](http://dx.doi.org/https://doi.org/10.3390/s23115295) [//doi.org/10.3390/s23115295](http://dx.doi.org/https://doi.org/10.3390/s23115295).
- [7] Z. Wang, S. Li, W. He, R. Yang, Z. Feng, G. Sun, A new topology-switching strategy for fault diagnosis of multi-agent systems based on belief rule base, Entropy 24 (2022) 1591. doi:[https:](http://dx.doi.org/https://doi.org/10.3390/e24111591) [//doi.org/10.3390/e24111591](http://dx.doi.org/https://doi.org/10.3390/e24111591).
- [8] X. Bu, H. Nie, Z. Zhang, Q. Zhang, An industrial fault diagnostic system based on a cubic dynamic uncertain causality graph, Sensors 22 (2022) 4118. doi:[https://doi.org/10.3390/s22114118](http://dx.doi.org/https://doi.org/10.3390/s22114118).
- [9] K. Pan, H. Liu, X. Gou, R. Huang, D. Ye, H. Wang, A. Glowacz, J. Kong, Towards a systematic description of fault tree analysis studies using informetric mapping, Sustainability 14 (2022) 11430. doi:[https://doi.org/10.3390/su141811430](http://dx.doi.org/https://doi.org/10.3390/su141811430).
- [10] C. Xu, J. Li, X. Cheng, Comprehensive learning particle swarm optimized fuzzy petri net for motor-bearing fault diagnosis, Machines 10 (2022) 1022. doi:[ttps://doi.org/10.3390/](http://dx.doi.org/ttps://doi.org/10.3390/machines10111022) [machines10111022](http://dx.doi.org/ttps://doi.org/10.3390/machines10111022).
- [11] ISO, ISO 15926-14:2020(E), ISO 15926 Part 14: Industrial top-level ontology, Technical Report, ISO, Geneva, CH, 2020.
- [12] https://github.com/iofoundry/ontology/tree/master/maintenance, Technical Report, ????
- [13] B. Cai, L. Huang, M. Xie, Bayesian networks in fault diagnosis, IEEE Transactions on industrial informatics 13 (2017) 2227–2240. doi:[https://doi.org/10.1109/TII.2017.2695583](http://dx.doi.org/https://doi.org/10.1109/TII.2017.2695583).