Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 1 (2024)

Analysis of Dynamic Knowledge Graph Construction and Clustering for Effective Knowledge Management in Machine-to-Machine Communication

Ganesh S. Pise [1*], Sachin D. Babar[2,] Parikshit N. Mahalle[3]

Research Scholar Smt. Kashibai Navale College of Engineering Pune [1], Assistant Professor in Pune Institute of Computer Technology Pune

Email- ganeshpise143@gmail.com

Department of Computer Engineering Sinhgad Institute of Technology Lonavala, Research Guide Smt. Kashibai Navale College of Engineering Pune[2]

Email- hodce.sit@sinhgad.edu

Department of Artificial Intelligent and Data Science, Vishwakarma Institute of Information Technology[3]

Email- aalborg.pnm@gmail.com

Article History:

Received: *18-09-2023*

Revised: *30-10-2023*

Accepted: *22-11-2023*

Abstract:

An era of interconnected devices that exchange data has emerged due to machine-to-machine (M2M) communication, a key component of the Internet of Things (IoT). This study explains how dynamic knowledge graph construction improves knowledge management in M2Mcommunication networks. In M2M communication, devices continuously generate and exchange data, creating a complex and dynamic information network. A dynamic knowledge graph is a promising solution for managing and addressing this level of complexity. The knowledge graph evolves in real time to capture M2M network relationships, entities, and data flows. M2M communication with dynamic knowledge graphs has many benefits. It begins with a broad overview of network components and their relationships. The structured format helps understand and make decisions by representing devices, their attributes, and their contextual relationships. The knowledge graph can also scale easily to support the rapid growth of devices and data in M2M networks. A dynamic knowledge graph lets M2M networks route data intelligently. Context-aware decisions reduce latency and improve network efficiency. The knowledge graph helps M2M networks detect and analyze anomalies and patterns. Detecting deviations from expected behavior improves security and proactive network maintenance, ensuring its integrity and reliability. Efficient knowledge management requires dynamic knowledge graphs in M2M communication networks. The data used for the proposed work is collected from the World Wide Web Consortium (W3C). It provides valuable insights into using technologies to improve learning and knowledge management. The dataset is comprehensive and useful for studying dynamic knowledge graphs and clustering in M2M. This enhances M2M networks' reliability and intelligence in the IoT era..

Keywords: M2MCommunication, Knowledge Graph Construction, NLP, Effective Knowledge Management.

Subject Classification: Primary 93A30, Secondary 49K15

1. Introduction

M2M communication is at the forefront of the IoT revolution, reshaping industries through autonomous data sharing and decision-making devices. Smart cities, healthcare, manufacturing, and environmental monitoring rely on M2M networks[1]. However, managing knowledge in this dynamic

Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 1 (2024)

landscape poses challenges, prompting researchers to explore solutions[2]. Presented study enhances M2M communication knowledge management using Natural Language Processing (NLP), creating dynamic knowledge graphs for M2M clustering[3]. The objective is to boost M2M network reliability and cognitive capabilities within the ever-evolving IoT environment. In the IoT, devices communicate to create a network of sensors with advanced cognitive abilities. These devices share, collaborate, and make decisions through M2M communication. Figure 1 illustrates a smart city powered by M2M connections[4]. Beyond smart cities, M2M technologies drive operational efficiency and decision-making in healthcare, agriculture, logistics, and manufacturing..

fig. 1 Smart city with M2M communications (source-aliga.sk)

Despite the benefits, M2M communication networks face challenges due to the sheer volume of devices and data. Effective knowledge management tools, including NLP, dynamic knowledge graph construction, and clustering, are essential[5], [6].Knowledge management in M2M involves recognizing, acquiring, retaining, organizing, and distributing knowledge to handle massive data and information from connected devices. Two key challenges include the dynamic nature of M2M networks and the diverse data formats, protocols, and semantics of devices. NLP, known for its ability to understand human language, is applied to M2M knowledge management, offering benefits such as enhanced semantics, entity recognition, and real-time processing. NLP application in M2M scenarios involves analyzing unstructured text data, like sensor logs and maintenance reports, to extract valuable insights[7].

The study delves into dynamic knowledge graphs, which visually represent entities and relationships in M2M communication. These graphs adapt to network changes, providing real-time awareness crucial for monitoring and decision-making[8]. Dynamic knowledge graphs offer scalability, representing entities and relationships to improve decision-making, regulate temperature in smart buildings, and optimize energy usage. They also enhance data routing efficiency for M2M data transmission[9], [10].To support this research, the W3C WoT use case dataset, provided by the World W3C, is used. This dataset sheds light on practical applications of WoT technologies in learning, knowledge management, and other fields. This study focuses on improving M2M knowledge management through dynamic knowledge graphs, clustering algorithms, and NLP. By leveraging insights from the W3C WoT use cases dataset, the research aims to enhance M2M network reliability and intelligence, preparing these networks for the challenges and opportunities presented by the IoT.

2. Literature review

IoT Scholars and professionals are exploring new M2M communication methods to address challenges and capitalize on opportunities. We review recent M2M communication, knowledge management, and IoT studies here. This review covers theoretical models, simulations, systematic reviews, case studies, and frameworks. Internet of Things and M2M communication differ. Knowledge management systems, QoE rate control, middleware development, knowledge modeling, trust management, and security enhancement use these methods. Research will illuminate M2M communication and IoT advances. R&D gaps should be identified. To illustrate the changing research landscape, each study's methodology, impact, results, and results are briefly examined. Table 1 lists major related work and its impact.

Table 1 Major related work

Author	Method	Methodology	Impact	Result
A. Paul et al.[11]	Theoretical	Probabilistic model	Reduced the complexity of M2M communication in IoT	Improved the performance of M2M communication
Y. Li et al.[12]	Simulation	Knowledge management system	Improved the efficiency of data-centric IoT application systems	Reduced the time and cost of developing and maintaining IoT applications
J. Yin et al.[13]	Simulation	QoE-Oriented rate control and resource allocation scheme	Improved the QoE of cognitive M2M communication	Reduced the energy consumption of cognitive M2M communication
C. Xie et al.[14]	Simulation	Multilayer IoT middleware	Improved the performance and scalability of IoT middleware	Reduced the cost of developing and maintaining IoT middleware
P. Wen et al.[15]	Systematic review	Knowledge modeling and extraction methods	Identified the state-of-the-art knowledge modeling and extraction methods for manufacturing process planning	Provided a comprehensive overview of the knowledge modeling and extraction methods for manufacturing process planning
C. Ramonell et al.[16]	Case study	Knowledge graph-based data integration system	Improved the efficiency of data integration for digital twins of built assets	Reduced the cost of developing and maintaining digital twins of built assets

Y. Xiao et al.[17]	Systematic review	Knowledge graph-based manufacturing process planning	Identified the state-of-the-art knowledge graph-based manufacturing process planning methods	Provided a comprehensive overview of the knowledge graph-based manufacturing process planning methods
X. Feng et al.[18]	Simulation	Neural generated knowledge graph	Improved the accuracy of detecting contradictions from IoT protocol specification documents	Reduced the time and cost of detecting contradictions from IoT protocol specification documents
H. Tyagi et al.[19]	Systematic review	Trust management techniques	Identified the state-of-the-art trust management techniques for security and privacy in IoT	Provided a comprehensive overview of the trust management techniques for security and privacy in IoT
Y. Wang et al.[20]	Simulation	IoT-based framework with Prognostics and Health Management and short term fire risk assessment	Improved the accuracy of fire risk assessment in smart firefighting systems	Reduced the time and cost of fire risk assessment in smart firefighting systems

This literature review illuminated the complex and dynamic IoT ecosystem machine-to-machine communication field. The analyzed studies cover many methods and fields. They enhanced field knowledge management, efficiency, security, and innovation. Although promising, the above studies have limitations. Many people have been assessed in simulated settings, but their real-life applicability needs further study. Remember that only manufacturing process planning, software engineering, and trust management are reviewed. This study shows M2M/IoT communication advancement. This development is driven by theory, practice, and improving interconnected systems. These methods can be tested, validated, and improved to build robust, flexible, and secure machine-to-machine communication networks for the IoT revolution.

3. M2M communication and Knowledge graphs

The IoT ecosystem relies on M2M communication, enabling devices to exchange data without human intervention. This communication is crucial for various applications, such as industrial sensor collaboration, healthcare data transmission, and smart home adjustments based on sensor inputs. The evolution of IoT has made intelligent devices, data connectivity, and M2M communication commonplace. In managing the vast data generated by connected devices, IoT and M2M knowledge management employ systematic methods. This includes collecting, storing, retrieving, analyzing, and using data to enhance intelligence and functionality. Knowledge graphs, utilizing a graph-based framework, visually represent entities and relationships, aiding in pattern recognition and relationship understanding.

NLP algorithms play a vital role in analyzing and manipulating human language data. They assist in organizing unstructured text, identifying entities, extracting facts, and establishing textual connections. NLP is essential for extracting and managing knowledge in the context of IoT and M2M, facilitating the organization of human-understandable text into stored, queried, and analyzed data.

4. Methodology

4.1.Dataset

This document functions as an extensive resource that delineates diverse practical applications and prerequisites for the Web of Things (WoT) technology[21]. This document encompasses comprehensive explanations regarding the utilization of WoT technology across diverse domains, including but not limited to smart homes, healthcare, agriculture, transportation, and other relevant sectors. This statement emphasizes the application of IoT devices, sensors, and web technologies in real-world situations to establish intelligent and highly interconnected environments.

	Domain	Sub Domain	Text
0	Smart Agriculture	Greenhouse Horticulture	Greenhouse Horticulture controlled by computers can create an optimal environment for growing plants. This enables to improve productivity and ensure stable vegetable production throughout the yea...
1	Smart Agriculture	Greenhouse Horticulture	Sensors (temperature, humidity, brightness, UV brightness, air pressure, and CO2) Heating, CO2 generator, open and close sunlight shielding sheet.
2	Smart Agriculture	Greenhouse Horticulture	Sensors values to clarify the gaps between conditions for maximizing photosynthesis and the current environment. Following sensors values at one or some points in the greenhouse: temperature, humi...
3	Smart Agriculture	Greenhouse Horticulture	Sensors and some facilities like heater, CO2 generator, sheet controllers are connected to the gateway via wired or wireless networks. The gateway is connected to the cloud via the Internet. All s...
4	Smart Agriculture	Open-field Agriculture	Water is vital for ensuring food security to the world□s population, and agriculture is the biggest consumer amounting for 70% of freshwater. Field irrigation application methods are one of the ma...

fig. 2 Sample dataset

4.2. Data Normalization

Step	Description	Application to the Dataset
a. Lowercase	Convert all text to lowercase to ensure uniformity.	Apply to all text in the dataset.
b. Remove Punctuation	Eliminate punctuation marks like commas and periods.	Apply to all text in the dataset.
c. Word Tokenize	Split text into individual words (tokens).	Tokenize each sentence in the dataset.
Sentence Tokenize	Split text into individual sentences.	Tokenize the entire dataset into sentences.
d. Duplicate Keywords Removal	Remove repeated words or phrases.	Identify and remove duplicate keywords within each document.
e. Unique Keywords	Extract a list of unique words or phrases.	Create a list of unique keywords for each document.

f. Remove HTML Tags	Strip out any HTML tags or markup language.	Remove HTML tags if present in the text.
g. Remove Whitespaces	Eliminate extra spaces and tabs between words.	Remove extra whitespaces within text.
h. Remove Unicode Characters	Filter out non-ASCII or special characters.	Remove any non-Unicode characters from the text.

4.3.Data Preprocessing

4.3.1. Spelling Correction

Spelling correction is a language processing technique employed to automatically correct words that have been misspelled within a given text. The primary objective of this tool is to improve the precision and legibility of written content by identifying and substituting words with their accurate forms. Spelling correction systems utilize diverse algorithms and reference dictionaries in order to propose corrections by considering the context and similarity between the misspelled word and valid words. For example, it has the capability to transform the misspelled word "teh" into the correct form "the" within the sentence "The quick brown fox jumps." The importance of spelling correction is particularly evident in contexts where accurate language usage is of utmost importance, such as in written documents, electronic correspondence, or online information retrieval.

4.3.2. Lemmatization

Lemmatization is the process of reducing words to their root, or lemma. Lemmatization converts words into their canonical or dictionary forms, unlike stemming, which removes prefixes and suffixes. This process standardizes text words for consistency and meaningful analysis. Lemmatization turns "running" into "run" and "better" into "good". This streamlines text processing and ensures that similar-meaning words share a base form. Lemmatization is important in sentiment analysis, information retrieval, and topic modeling.

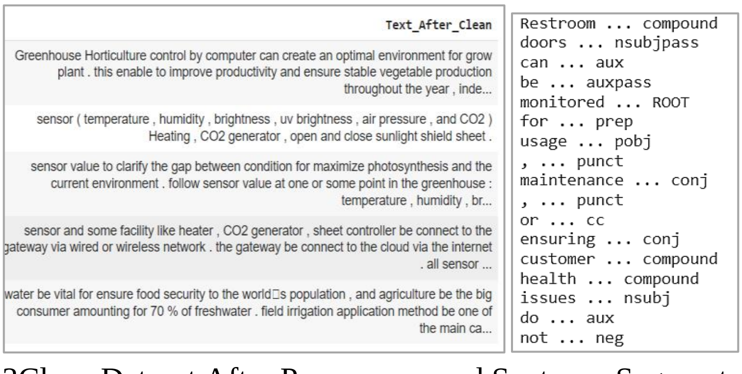

fig. 3Clean Dataset After Pre-process and Sentence Segmentation

4.4.Sentence Segmentation

The task of entity extraction in natural language processing encompasses the identification and isolation of distinct elements of information, including but not limited to individuals' names, organizational entities, geographical locations, dates, and other relevant entities, from a provided text. This procedure facilitates the organization of unstructured textual data by identifying and classifying essential components contained therein.

4.5.Entities Extraction

```
[['Fleet Dashboard system', 'visual geofencing capability'],
 ['trustful  discovery', 'public space application'],
 ['', ''],
 ['public  user', 'public  device'],
 ['', 'personalizable space service'],
 ['same slide show it', 'private smartphone slideshow'],
 ['within  they', 'such Middle Node C'],
 ['', 'smart water context'],
 ['such iot device', 'even  individual'],
 ['multiple production result', 'raw fault condition'],
 ['', ''],
 ['early quality it', 'raw material delivery'],
 ['well use form', 'such OPC UA'],
 ['typically server that', 'gateway OPC etc'],
 ['4 production line', 'control OPC purpose'],
 ['it', '2 filling module'],
 ['meaningful  it', 'operational  retailer'],
 ['operational customer understanding', 'real worker status'],
 ['such  people', ''],
 ['corporate loss prevention', 'real store support outcome']]
```

```
be                  12
include              3
service              2
sensor               2
boost                1
expect               1
need                 1
be important         1
have additional      1
be investigate       1
initiate standard    1
register with        1
result in            1
device               1
integrate            1
equip biomedical     1
build                1
```

Fig. 4 Entity Extraction and 5Relationship Extraction

4.6.Relations Extraction

The task of entity extraction in natural language processing encompasses the identification and isolation of distinct elements of information, including but not limited to individuals' names, organizational entities, geographical locations, dates, and other relevant entities, from a provided text. This procedure facilitates the organization of unstructured textual data by identifying and classifying essential components contained therein.

4.7.Dynamic Graph Construction

Node addition: Let E represent the set of nodes in the dynamic graph and set of criteria is denoted by A for adding new nodes n_{new}. The addition of new node based on knowledge acquisition is represented as eq.1. The operation updated the set of nodes E by adding the newly acquired node.

$$E \leftarrow E \cup \{n_{new}\} \text{ where } n_{new} \in A \ldots 1$$

Edge Addition: Let R represent the set of edge in the dynamic graph, and M denotes the methods for identifying and establishing relationship between nodes. The addition of a new edge e_{new} , based on established relationships as in eq.2.

$$R \leftarrow R \cup \{e_{new}\} where\ e_{new} \in M$$

Node removal: Let E represent the set of nodes in the dynamic graph and C denotes the criteria for removing outdated nodes or ireelavent nodes. The removal of node, n_{remove}based on specified criteria is represented as eq.3

$$E \leftarrow E - \{n_{remove}\} where\ n_{remove} \in C$$

Edge removal: Let R represent t the set of edges in the dynamic graph and D denote the criteria for removing outdated and irrelevant edges. The removal of an edge based on specified criteria represented as eq.4

$$R \leftarrow R - \{e_{remove}\} where\ e_{remove} \in D \ldots 4$$

4.8.Knowledge Graph Generation

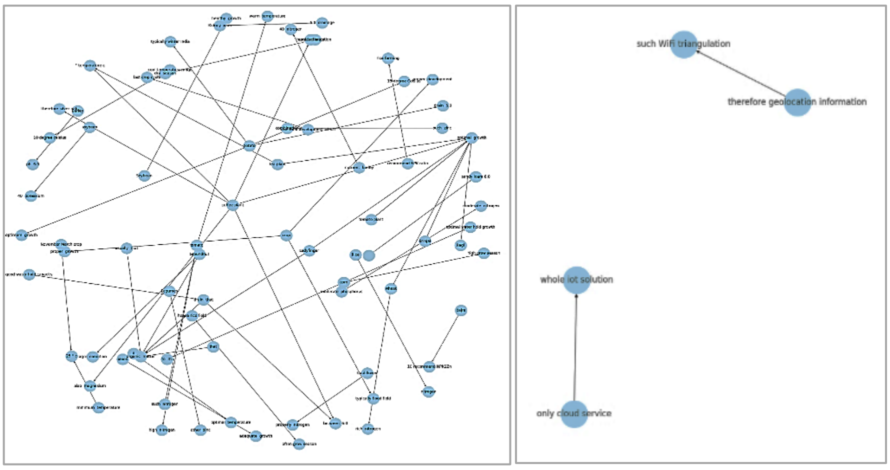

fig. 5edge " require" & Edge "grow in"

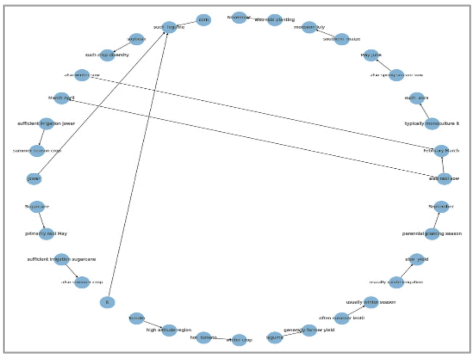

fig. 6edge "grow as"

5. Conclusion and future scope

This study improves M2M communication in the Internet of Things knowledge management using dynamic knowledge graphs and clustering. NLP solves complex M2M network development problems, as shown by our work. Results show that knowledge graphs with NLP insights improve M2M communication systems' effectiveness, scalability, and flexibility. Our findings show that dynamic knowledge graphs accurately depict M2M communication network evolution. We can keep up with instantaneous changes, improving decision-making and network knowledge. NLP has improved data routing efficiency, latency, and transmission. For machine-to-machine communication, this study improved dynamic knowledge graph construction and clustering. However, this area offers many promising research opportunities. Deep and machine learning for knowledge graph clustering are promising. These methods may improve clustering precision and dynamic system adaptability. Future research and development prioritizes system scalability. Distributed knowledge graph construction and clustering are essential to solving large-scale M2M network problems, especially as the IoT grows. To prove the system's efficacy and suitability for machine-to-machine communication, performance metrics and extensive field trials are needed. Future research on intelligent and efficient M2M communication in IoT begins with this study..

References

[1] D. M. Chang, T. C. Hsu, H. Yang, and Y. H. Tsai, “An M2M computing model for improving the performance among devices,” *Microsyst. Technol.*, vol. 27, no. 4, pp. 1467–1474, 2021, doi: 10.1007/s00542-019-04382-7.

[2] S. Iraji, P. Mogensen, and R. Ratasuk, “Recent Advances in M2M Communications and Internet of Things (IoT),” *Int. J. Wirel. Inf. Networks*, vol. 24, no. 3, pp. 240–242, 2017, doi: 10.1007/s10776-017-0362-3.

[3] Y. Mehmood, C. Görg, M. Muehleisen, and A. Timm-Giel, “Mobile M2M communication architectures, upcoming challenges, applications, and future directions,” *Eurasip J. Wirel. Commun. Netw.*, vol. 2015, no. 1, pp. 1–37, 2015, doi: 10.1186/s13638-015-0479-y.

[4] J. Lin, Y. Zhao, W. Huang, C. Liu, and H. Pu, “Domain knowledge graph-based research progress of knowledge representation,” *Neural Comput. Appl.*, vol. 33, no. 2, pp. 681–690, 2021, doi: 10.1007/s00521-020-05057-5.

[5] Z. Huang, X. Li, Y. Ye, B. Zhang, G. Xu, and W. Gan, “Multi-view knowledge graph fusion via knowledge-aware attentional graph neural network,” *Appl. Intell.*, vol. 53, no. 4, pp. 3652–3671, 2023, doi: 10.1007/s10489-022-03667-1.

[6] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, “A Survey on Knowledge Graphs: Representation, Acquisition, and Applications,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 33, no. 2, pp. 494–514, 2022, doi: 10.1109/TNNLS.2021.3070843.

[7] S. Liang, J. Shao, J. Zhang, and B. Cui, “Graph-Based Non-Sampling for Knowledge Graph Enhanced Recommendation,” *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 9, pp. 9462–9475, 2023, doi: 10.1109/TKDE.2023.3240832.

[8] C. Wu, Z. Zeng, Y. Yang, M. Chen, X. Peng, and S. Liu, “Task-driven cleaning and pruning of noisy knowledge graph,” *Inf. Sci. (Ny).*, vol. 646, no. November 2022, p. 119406, 2023, doi: 10.1016/j.ins.2023.119406.

[9] Q. Zou and C. Lu, “An Accurate Matching Query Method of Natural Language Knowledge Graph Based on Hierarchical Graph Topological Sequence,” *IEEE Access*, vol. 10, pp. 24080–24094, 2022, doi: 10.1109/ACCESS.2022.3155520.

[10] Z. Zeng, C. Xie, W. Tao, Y. Zhu, and H. Cai, “Knowledge-Graph-Based IoTs Entity Discovery Middleware for Nonsmart Sensor,” *IEEE Trans. Ind. Informatics*, vol. PP, pp. 1–13, 2023, doi: 10.1109/TII.2023.3292540.

[11] A. Paul and S. Rho, “Probabilistic Model for M2M in IoT networking and communication,” *Telecommun. Syst.*, vol. 62, no. 1, pp. 59–66, 2016, doi: 10.1007/s11235-015-9982-z.

[12] Y. Li *et al.*, “IoT-CANE: A unified knowledge management system for data-centric Internet of Things application systems,” *J. Parallel Distrib. Comput.*, vol. 131, pp. 161–172, 2019, doi: 10.1016/j.jpdc.2019.04.016.

[13] J. Yin, Y. Chen, G. Sang, B. Liao, and X. Wang, “QoE-Oriented Rate Control and Resource Allocation for Cognitive M2M Communication in Spectrum-Sharing OFDM Networks,” *IEEE Access*, vol. 7, pp. 43318–43330, 2019, doi: 10.1109/ACCESS.2019.2908681.

[14] C. Xie, B. Yu, Z. Zeng, Y. Yang, and Q. Liu, “Multilayer Internet-of-Things Middleware Based on Knowledge Graph,” *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2635–2648, 2021, doi: 10.1109/JIOT.2020.3019707.

[15] P. Wen, Y. Ma, and R. Wang, "Systematic knowledge modeling and extraction methods for manufacturing process planning based on knowledge graph," *Adv. Eng. Informatics*, vol. 58, no. April, p. 102172, 2023, doi: 10.1016/j.aei.2023.102172.

[16] C. Ramonell and R. Chac, "Automation in Construction Knowledge graph-based data integration system for digital twins of built assets," vol. 156, no. May, 2023, doi: 10.1016/j.autcon.2023.105109.

[17] Y. Xiao, S. Zheng, J. Shi, X. Du, and J. Hong, "Knowledge graph-based manufacturing process planning: A state-of-the-art review," *J. Manuf. Syst.*, vol. 70, no. August, pp. 417–435, 2023, doi: 10.1016/j.jmsy.2023.08.006.

[18] X. Feng *et al.*, "Detecting contradictions from IoT protocol specification documents based on neural generated knowledge graph," *ISA Trans.*, no. xxxx, 2023, doi: 10.1016/j.isatra.2023.04.025.

[19] H. Tyagi, R. Kumar, and S. K. Pandey, "A detailed study on trust management techniques for security and privacy in IoT: challenges, trends, and research directions," *High-Confidence Comput.*, vol. 3, no. 2, p. 100127, 2023, doi: 10.1016/j.hcc.2023.100127.

[20] Y. Wang, R. Zheng, L. Hou, M. Li, and W. Li, "A novel IoT-based framework with Prognostics and Health Management and short term fire risk assessment in smart firefighting system," *J. Build. Eng.*, vol. 78, no. April, 2023, doi: 10.1016/j.jobe.2023.107624.

[21] W3C Editors, "Web of Things (WoT): Use Cases and Requirements." 2021, [Online]. Available: https://w3c.github.io/wot-usecases/.