Proposed fog computing-enabled conceptual model for semantic interoperability in internet of things

Devamekalai Nagasundaram, Selvakumar Manickam, Shams Ul Arfeen Laghari, Shankar Karupayah
National Advanced IPv6 Centre, Universiti Sains Malaysia, Penang, Malaysia

ABSTRACT
Semantic interoperability has emerged as a key barrier amidst the major developments and challenges brought about by the rapid expansion of internet of things (IoT) applications. Establishing interoperability is essential for IoT systems to function optimally, especially across diverse organizations. Despite extensive research in achieving semantic interoperability, dynamic interoperability, a vital facet, remains inadequately addressed. This paper addresses this gap by presenting a fog-based conceptual model designed to facilitate dynamic semantic interoperability in IoT. The model incorporates a single-tier fog layer, providing the necessary processing capabilities to achieve this goal. The study conducts a comprehensive literature review on semantic interoperability, emphasizing latency, bandwidth, total cost, and energy consumption. Results demonstrate the proposed double skin façade (DSF) model’s remarkable 88% improvement in service delay over IoT-SIM and Open IoT, attributed to its efficient load-offloading mechanism and optimized fog layer, offering a 50% reduction in service delay, power consumption, and 86% reduction in network usage compared to existing approaches through data redundancy elimination via pre-processing at the fog layer.

Keywords:
Conceptual model
Fog environment
Internet of things
Internet of things equipment effectiveness
Semantic interoperability

1. INTRODUCTION
The internet of things (IoT) refers to a network of interconnected and autonomous objects that possess the ability to communicate data over a network, and function independently without direct human interaction [1]. The IoT encompasses a network system within the internet, designed to facilitate instantaneous and seamless interaction among objects, machines, and humans through advanced technological approaches [2]. As defined by the International Telecommunication Union (ITU), the IoT represents a global infrastructure for the knowledge society, interconnecting physical and virtual entities via interoperable information and communication technologies, thereby enabling enhanced services [3]. The IoT concept has gained widespread acceptance and application in a variety of fields [2], such as smart agriculture [4], smart supply chains [5], smart medical care [6], smart transportation [6–9], and smart cities [10], [11]. Consequently, interoperability between diverse and heterogeneous entities becomes critical for IoT systems to achieve their objectives [12].

The results of a comprehensive survey conducted by Bain Company highlight interoperability as the foremost barrier to widespread adoption of the IoT in the United States (US) [13]. Recognizing the criticality of interoperability, a recent collaborative effort involving European industrial and academic partners has emerged with the aim of constructing IoT frameworks specifically targeted at addressing this challenge [14].
Notably, semantic interoperability, which strives to establish a unified comprehension of information exchanged among heterogeneous IoT devices across diverse domains, emerges as the pinnacle of interoperability measures [16]–[18]. The management of vast volumes of data generated by IoT devices, alongside the inherent data heterogeneity, diverse capabilities of the devices, and the array of services they offer, represents a paramount concern within the IoT landscape [19]. Data collected from multiple sources exhibit a multitude of semantic representations, rendering semantic interoperability a critical challenge in facilitating seamless communication and harmonized functionalities across diverse IoT systems [20].

The scholarly community has presented a plethora of research studies and proposed solutions aimed at mitigating the challenges posed by heterogeneity in IoT environments. Nevertheless, real-time applications within the IoT domain continue to encounter obstacles in enhancing latency, bandwidth, and dynamism, primarily due to the intricate task of transferring substantial data volumes across multiple interconnected systems. The remainder of this article is structured as follows. Section 2 presents the motivation, semantic interoperability in IoT, semantic web, ontology enrichment, fog computing, and research problems. Section 3 discusses about related research work in this field. Section 4, shows the design of the proposed model. Section 5, provides the expected result of this paper. Finally, this paper will conclude in section 6.

2. BACKGROUND
2.1. Motivation

The recent global pandemic has significantly amplified the demand for IoT services, resulting in a remarkable surge in the deployment of IoT devices [21]. The adoption of smart city initiatives, intelligent classrooms, interconnected campuses, and advanced transportation systems has further propelled the utilization of IoT devices [22]. Consequently, the exponential growth of IoT installations has become a primary driver of the unprecedented proliferation of global data [23]–[25].

The staggering increase in data volume, accompanied by its heterogeneity and interconnectedness, necessitates efficient data management solutions that account for scalability, data gravity, and integration challenges [26]. Many enterprises have turned to cloud computing as a means of managing their data, drawn by its cost-effectiveness, flexibility, availability, scalability, and other inherent benefits [27]. However, establishing a cohesive data management framework remains a formidable task due to the disparate nature of data originating from a multitude of IoT devices. Hence, it becomes crucial for IoT devices within the ecosystem to possess the capability to share and leverage data, emphasizing the importance of data interoperability, commonly referred to as semantic interoperability. The processing of massive data volumes necessitates robust computing power, which can be provided by existing cloud architectures. However, the centralized nature of cloud infrastructures introduces challenges related to latency, bandwidth, and reliability.

To address these challenges, we propose a fog-based solution to facilitate dynamic semantic interoperability within IoT applications. By leveraging the fog computing paradigm, our approach aims to overcome the limitations of centralized architectures, offering enhanced processing capabilities and reducing latency issues. The fog layer acts as an intermediary between the cloud and IoT devices, enabling efficient data management, real-time analytics, and seamless communication. Our proposed solution seeks to bridge the gap in achieving dynamic semantic interoperability, thereby empowering IoT applications with improved performance, reliability, and efficiency.

2.2. Semantic interoperability in internet of things

The term “interoperability” within computer systems refers to the ability of multiple applications or components to share information and effectively utilize data [28]. Extensive research conducted by Andoec and Noura through a systematic literature review underscores the growing recognition of the interoperability challenges posed by IoT frameworks, particularly in recent years [28], [29]. Figure 1 a visual representation is provided to illustrate the fundamental aspects of interoperability in a four-dimensional framework. The four pillars of IoT interoperability encompass technical, syntactical, semantic, and organizational dimensions. Given the dynamic and intricate nature of data in cross-domain IoT applications, advanced data modeling and management techniques are imperative [28]. In this context, the concept of semantic interoperability has emerged, denoting the ability of diverse applications and entities to interpret exchanged data with a clear and consistent understanding of its meaning [30]. Semantic interoperability encompasses the ability of multiple IoT devices, systems, or applications to exchange information or knowledge in a manner that is both meaningful and accurately interpreted automatically, surpassing the scope of data transfer and clear data definition across IoT devices [31], [32]. Achieving semantic interoperability involves a streamlined process that eliminates the need for intricate programming, configuration, or human intervention in locating specific data [33].
Proposed fog computing-enabled conceptual model for semantic … (Devamakalai Nagasundaram)

The implementation of ontology in information exchange processes serves as a critical enabler of semantic interoperability [34]. The term “ontology” derives its origins from Greek and was introduced by Johannes Clauberg, a distinguished German theologian and philosopher. It denotes a structured knowledge representation comprising a collection of interrelated concepts within a specific domain. The essence of ontology lies in explicitly representing knowledge for a conceptual domain through the logical constructs conveyed by semantic web languages [35]. While a significant portion of ontologies employed to address semantic interoperability in IoT remains static, the advent of real-time services and the influx of voluminous data necessitates the incorporation of dynamicity to cater to evolving requirements, device capability changes, and application mobility [36]. Even with the ontologies, it is still challenging to accomplish semantic interoperability in diverse IoT contexts with reduced computing time, quick annotation, low complexity, and dynamic semantics.

2.3. Semantic web

The semantic web represents an extension of the world wide web aimed at enhancing the meaning and interconnectedness of online information by providing semantic context [37]. It empowers machines with the ability to comprehend and intelligently process data. At its core, the Semantic Web endeavours to enrich the existing structure of the web by incorporating additional metadata and ontologies that establish meaningful connections between diverse pieces of information [38].

The semantic web comprises several essential components, including the resource description framework (RDF), ontologies, SPARQL, and web ontology language (OWL), which collaboratively facilitate the organization, integration, and comprehension of web data [39]. RDF serves as a framework for representing knowledge on the web, establishing a standardized approach for expressing data through subject-predicate-object triples that form the core of the semantic data model [40]. Ontologies play a vital role by defining the language and connections between concepts within a specific domain. By offering a formal and structured representation of knowledge, ontologies enhance data interpretation and foster improved interoperability [41].

SPARQL serves as a powerful query language designed to retrieve and manipulate RDF data, enabling the search and extraction of specific information from the semantic web [42]. OWL, on the other hand, serves as a language specifically designed for the development of ontologies. It offers a rich set of constructs that allow for the specification of classes, properties, and relationships between entities, enhancing the robustness and expressiveness of ontologies [43]. Through the collaboration of these components, a web of data is constructed that not only caters to human readability but also facilitates machine understanding. This machine-understandable web enables intelligent search capabilities, seamless data integration, and automation of information on the web [37].

In the realm of IoT, the semantic web plays a vital role in facilitating seamless integration and interoperability among the vast array of interconnected devices [33]. By harnessing semantic technologies, the IoT can achieve enhanced levels of data interoperability and automation [44]. Devices gain the ability to comprehend and analyze the contextual information of the data they collect, leading to more informed decision-making and process automation. The rapid growth of the IoT industry underscores the necessity for standardized data models, vocabularies, and ontologies that enable devices to collaborate and communicate effectively. This standardization ultimately fosters the development of IoT applications that are more efficient and intelligent in their operations [33].
2.4. Ontology enrichment

An ontology serves as a formal representation of knowledge, capturing the concepts, relationships, and attributes within a specific domain [45]. With the increasing volume of data generated by IoT devices, it becomes crucial to ensure that ontologies evolve effectively to accommodate and articulate this growing knowledge [46]. Ontology enrichment is a subset of ontology evolution. Ontology evolution encompasses a wider range of activities that include ontology enrichment as well as others such as ontology refinement and restructuring to maintain the ontology’s relevance and usefulness over time. Ontology enrichment is the process of enhancing an existing ontology to improve its expressiveness and coverage by adding new concepts, relationships, or attributes [47]. By incorporating ontology enrichment, the dynamic and evolving nature of the IoT domain can be effectively addressed, ensuring that the ontology remains relevant and valuable over time.

Ontology enrichment plays a pivotal role in enhancing interoperability and providing a robust representation of IoT data [48]. It enables improved data integration across diverse IoT platforms and applications, leading to a deeper understanding of data and enhanced analysis capabilities [49]. By introducing new concepts and relationships, the ontology can effectively capture the complexities and interdependencies within the IoT ecosystem, enabling more meaningful and context-aware interpretation of data [48]. Moreover, in the context of the IoT, ontology enrichment facilitates information discovery and empowers advanced analytics. Researchers and practitioners gain the ability to extract valuable insights from IoT data by leveraging evolving ontologies that encompass a broader range of ideas and relationships. Richer ontologies facilitate advanced reasoning and inference processes, enabling the identification of hidden patterns, correlations, and trends amidst the vast volume of IoT data [48].

The increasing magnitude of data generated by the IoT underscores the importance of ontology enrichment [49]. By enriching ontologies, we can achieve a more comprehensive representation of the IoT domain, enhance data integration and interoperability, and enable sophisticated knowledge discovery and analytics. Ontology enrichment empowers us to unlock the full potential of IoT data and explore new avenues for innovation, decision-making, and optimization across various industries and applications. It paves the way for leveraging the richness of IoT data to drive advancements and unlock valuable insights in an interconnected world [48].

2.5. Fog computing

Fog computing is a computer architecture advancement with a resource pool where one or more universal and distributed nodes are enabled to interact and communicate at the extreme edge level instead of being supported by cloud computing [50]. Fog computing was coined to define the computational paradigm that tries to bring the benefits of the cloud closer to end devices, similar to how fog is seen whenever a cloud is near the ground [51]. Figure 2 illustrates the fog computing architecture.

Fog nodes operate autonomously and collaborate within a hierarchical environment to provide computational agility, enhanced communication, increased storage capacity, and a wide range of novel market services. They eliminate the need for third-party intervention and cater to the growing number of devices, clients, and end-users, offering a seamless and efficient infrastructure for distributed computing and service provisioning. The National Institute of Standards and Technology (NIST) summarised six basic properties of fog computing, as illustrated in Figure 3 [52].
Fog computing offers low-latency services by placing the fog nodes in close proximity to end users. This proximity enables faster analysis and response to user-generated data compared to traditional cloud services. The distributed deployment of fog computing creates a decentralized support system, allowing for geographically distributed services, and applications to operate on the fog computing infrastructure. This decentralized approach enhances performance, scalability, and reliability, enabling efficient data processing and real-time interactions in diverse locations [52].

The heterogeneous nature of fog computing enables the collection and analysis of data from diverse sources obtained through multiple network communication methods. Given the time-sensitive and heterogeneous characteristics of IoT, fog computing offers crucial features such as interoperability, federation, scalability, and real-time interaction support. These capabilities facilitate seamless integration and collaboration among different devices, systems, and applications within the IoT ecosystem. Fog computing acts as a bridge between edge devices and centralized cloud services, enabling efficient data processing, resource management, and dynamic adaptation in a distributed and diverse environment [51].

The prominent characteristics of fog computing offer solutions that have the potential to significantly enhance revenues, reduce costs, and expedite product rollouts in the IoT industry. By minimizing request-response time from IoT devices and providing local computational power, fog computing addresses key challenges of cloud computing such as latency, storage limitations, resource constraints, and high bandwidth requirements [51]. This enables more efficient and responsive processing of data at the edge, closer to the source of generation. The capabilities of fog computing contribute to improved system performance, enhanced user experiences, and increased overall productivity in IoT deployments [50].

2.6. Research problem

To ensure an efficient and interconnected IoT framework, devices must be able to effectively communicate and share data within heterogeneous systems. This facilitates the creation of a cost-effective and easily implementable application chain [53]. When one device can utilize data collected by another device, it reduces the need for redundant information gathering and optimizes resource utilization [54]. Seamless data availability is crucial for new devices integrated into the framework. However, traditional centralized semantic systems are less practical for real-time IoT applications [55]. High response time is not tolerable in time-sensitive IoT applications [56]. Cloud-based semantic interoperability systems also incur high bandwidth usage for real-time data transfer, leading to increased network utilization [20]. Additionally, existing solutions often fail to provide dynamic semantic interoperability processing due to resource limitations in IoT gateways, which lack the necessary energy and processing capability [55], [57]. To address these challenges, we propose a fog-based conceptual model that leverages the capabilities of fog computing to tackle dynamic semantic interoperability issues in IoT. This approach offers a more efficient and responsive solution for achieving real-time interoperability in IoT applications.
3. RELATED WORK

This section discusses the semantic interoperability solution contributed for IoT systems. In order to address the crucial need for semantic interoperability in IoT systems, various solutions have been proposed [58]. The IoT domains are advancing with ongoing research and concepts for re-utilizing, fusing, and abstracting current technology to attain interoperability semantically [59]. One prominent approach is the utilization of ontologies, which are widely employed to facilitate semantic interoperability. Additionally, fog computing has gained attention as a potential enabler of semantic interoperability in IoT.

A study conducted by Chen and Smys [55] focused on using fog computing to achieve IoT semantic interoperability. They developed a two-layer hierarchical fog network, where data was mapped and then transmitted to the cloud for comprehensive processing. The architecture employed a single-owl file ontology, and a series of operations including screening, consolidation, composing, modeling, and data mapping was performed within the fog layers. The proposed model aimed to reduce processing costs, improve energy efficiency, minimize latency, and optimize network utilization [55].

Rahman and Hussain [60] has presented a comprehensive solution for achieving semantic interoperability in IoT through the utilization of the fog computing framework. In their model, certain semantic tasks are offloaded from the cloud to the fog nodes, leading to reduced job execution time and energy consumption. An efficient offloading technique is employed to facilitate communication between fog-to-fog and fog-to-cloud devices [60]. The fog nodes in the system are organized hierarchically based on their roles. The L1-fog nodes are responsible for operations such as composing, modeling, and connecting, utilizing the high-quality aggregated data obtained. The generated data is semantically annotated using a single-owl ontology. However, it should be noted that the static nature of the deployed ontology restricts the introduction of new components into the system. In contrast to the traditional semantic paradigm where computations are solely performed in the cloud, the fog computing architecture allows for pre-processing of data packets by the fog devices. This approach eliminates delays, energy waste, network utilization, and computation costs. However, it has been observed that service delays may increase as the number of fog nodes grows [60]. One significant advantage of this method is the reduction in energy consumption since semantic operations take place in nearby fog nodes rather than in the cloud. Additionally, the system benefits from improved network bandwidth utilization, leading to decreased service delays, lower energy consumption, and overall cost reduction [60]. Overall, Rahman’s approach demonstrates the potential of the Fog computing framework in enhancing semantic interoperability in IoT systems by offloading semantic tasks and optimizing data processing and communication within the fog infrastructure, but using multiple layers of fog nodes introduces prominent delay and additional cost.

In another research paper, Rahman and Hussain [57] introduces the concept of a “lightweight ontology” as a means to achieve semantic interoperability in IoT systems. They emphasizes the use of radio frequency identification (RFID), sensors, and actuators as elements of the IoT framework. The main objective is to develop a lightweight ontology, referred to as light-weight ontology (LiO-IoT), that facilitates semantic interoperability in a simplified manner. The proposed solution focuses on utilizing widely used and readily available ontologies to achieve semantic interoperability. The ontology employed in the research is designed to be lightweight, prioritizing simplicity and efficiency [57]. However, Rahman and Hussain [20] suggests that further enhancements can be made by incorporating a lightweight dynamic ontology using a machine learning approach, which would further improve the capabilities of the proposed solution for IoT semantic interoperability.

This lightweight dynamic ontology using a machine learning approach introduces the capability to automatically discover and incorporate new individuals and concepts into the ontology, thereby enhancing its dynamic nature [20]. The primary objective of dynamic ontology is to facilitate real-time information exchange and dynamic interaction among multiple IoT devices. Machine learning techniques are employed to leverage the wealth of data within the ontology, enabling the identification of hidden patterns and unexpected knowledge [20]. The clustering approach significantly accelerates the search and concept discovery process within the IoT infrastructure. However, the introduction of dynamic ontologies also introduces a delay in clustering, leading to increased response time and processing time. Additionally, delays have been observed in detecting and integrating new concepts into the ontology [20]. On the other hand, the lightweight dynamic ontology demonstrates superior performance and resource efficiency in large-scale networks. Despite a slightly higher average response time compared to solutions without a dynamic ontology, the benefits of incorporating a dynamic ontology outweigh the minimal increase in response time. The clustering mechanism and the process of identifying and including new nodes contribute to this response time difference.

Overall, the use of a lightweight dynamic ontology offers a promising approach to achieving real-time and dynamic semantic interoperability in IoT systems, enabling efficient information exchange and knowledge discovery while considering the resource constraints of large-scale networks. By utilizing existing
ontologies and exploring machine learning techniques for dynamic ontology development, Rahman’s work presents a promising approach for achieving efficient and effective semantic interoperability in lightweight IoT environments. However, high response time and clustering delay are present in the method proposed in [20]. Table 1 indicates the differences between each reviewed related work.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Fog based</th>
<th>Dynamic/static</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>[55]</td>
<td>Yes</td>
<td>Static</td>
<td>i) medium latency, energy usage; ii) lightweight design; iii) two level of fog enables strong processing; and iv) processing occurred at the fog device; as a result, network utilization and computational costs were reduced.</td>
<td>i) the inclusion of a new feature was not possible; ii) a multi-layer approach requires an extra device, which would raise the cost; and iii) dynamic data mapping not supported by the single owl model.</td>
</tr>
<tr>
<td>[57]</td>
<td>No</td>
<td>Static</td>
<td>i) lightweight; ii) concentrate on RFID, actuators, and sensors; iii) low computational cost; and iv) high delay.</td>
<td>i) clustering delay is present; ii) high response time; and iii) compared to non-dynamic networks, there is a longer latency for new node recognition and inclusion.</td>
</tr>
<tr>
<td>[20]</td>
<td>No</td>
<td>Dynamic</td>
<td>i) lightweight; ii) identify new ideas automatically and include them in the ontology; iii) implement a real-time framework; iv) use a machine learning approach to accelerate the process of searching for and learning new ideas; and v) suitable for large-scale.</td>
<td>Static ontology and could not add a new component to the system.</td>
</tr>
</tbody>
</table>
| [60] | Yes       | Static         | i) using lightweight ontology; ii) reduce execution time and energy consumption; and iii) edge processing at the fog nodes helps reduce network usage and response time. | | }

4. DESIGN OF PROPOSED MODEL

IoT devices by their very nature are compact and constrained, they have limited processing capability. As a result, IoT devices have limited resources in the processing capacity to process real-time data and make it interoperable across various devices. We proposed a methodology for IoT apps to update ontology automatically in real-time to accomplish dynamic semantic interoperability. As time-sensitive processing power is required, fog computing architecture is applied to provision processing capabilities.

The proposed fog-based semantic interoperability model effectively manages ontology with real-time IoT data processing. Processing data from IoT devices to be interoperable requires multiple tasks such as data mapping, data parsing, RDF transformation, and ontology enrichment. Therefore, the fog nodes provide processing power to carry out all these tasks. The proposed model of real-time automatic ontology update to achieve dynamic semantic interoperability contains three layers:

- Edge layer-comprises end devices from the IoT framework. The data and attributes from the end devices in the edge are passed to the fog layer to be processed and analyzed.
- Fog layer-is responsible for providing processing capabilities to enable real-time automatic ontology enrichment to achieve dynamic semantic interoperability. This layer receives data from the edge layer and regularly uploads data to the cloud. Data is constantly transferred between the edge and fog layers, while data is only transferred from fog to the cloud when needed.
- Cloud layer-comprises a centralized database in which an ontology is stored. Only when necessary, do uploads and downloads from a centralized cloud ontology take place. For example, the cloud ontology was uploaded when new data was contributed to the fog ontology.

The fog layer handles data procurement, RDF transformation, and ontology enrichment processes. Figure 4 shows the functional architecture of the model proposed. The architecture consists of 3 primary phases: Data procurement, comma-separated values (CSV) to RDF transformation, and ontology enrichment. The data procurement phase provides a method for obtaining structured and unstructured data from IoT devices. The information gathered will be parsed and filtered. The data parser is responsible for collecting data and converting it into CSV format. Then the data filtration will check the availability of the data in the ontology. This phase helps to improve ontology to become dynamic.

CSV to RDF transformation phase uses the CSV to RDF engine to create the RDF from the data acquired in the procurement phase. This phase is crucial as the RDF triples generated from this phase are the ontology enrichment phase benchmark. Ontology enrichment is the final phase responsible for enhancing and improvising the ontology by including additional data. The proposed model’s ontology is implemented using the learn-and-grow technique. Ontology requires constant evolution due to the massive amount of data and the heterogeneity of IoT devices, therefore the process of enhancing and enriching the ontology is performed for real-time IoT data. Each phase is discussed in detail in the following section.

Proposed fog computing-enabled conceptual model for semantic ... (Devamekalai Nagasundaram)
Figure 4. Process flow diagram of proposed model

4.1. Data procurement

Processing and analyzing large volumes of structured and unstructured IoT data in real-time has posed significant challenges. Extracting critical contextual information from real-time data streams is particularly difficult due to factors such as heterogeneity, complexity, dynamism, and inadequate data representation. In our research, we propose a novel approach where data from IoT devices is directly received into the fog device, which is located in close proximity to the information source within the network. The fog node’s data procurement agent is responsible for gathering the raw data and performing analyses to identify any new outliers.

To differentiate between outlier and typical data, a data parser is employed. Outlier data is identified as a potential new data set within the system, while data deemed typical undergoes real-time annotation using sensor metadata. This annotation process enhances the understanding of the data and its context. To better illustrate the data flow within the system, we provide a figure depicting the data procurement module. This module showcases the seamless flow of data from IoT devices to the fog node, where preliminary analysis and outlier detection take place. It highlights the importance of efficient data handling and the role of fog computing in facilitating real-time data processing and analysis. By implementing this approach, we aim to address the challenges associated with processing and extracting valuable insights from real-time sensor data, ultimately enabling more effective decision-making and utilization of IoT-generated data. Figure 5 illustrates the data flow of the data procurement module.

Figure 5. Data procurement modal process flow

4.1.1. Data procurement agent

The data procurement agent plays a crucial role in the data flow process within the fog computing environment. It receives the raw data streams transmitted from IoT devices, which can include various types of sensory information such as temperature, humidity, pressure, and more. The agent’s primary function is to process this incoming data and identify any potential outliers. To identify outliers, the data procurement agent...
utilizes algorithms and statistical techniques that compare the received data with expected values or predefined thresholds. If the obtained data significantly deviate from the anticipated range or show abnormal patterns, they are classified as outliers. Outlier data can arise from various factors such as sensor malfunction, environmental anomalies, or unexpected events. When an outlier is detected, the data procurement agent flags it as new data that requires further analysis. This identification of outliers helps in capturing unusual or unexpected events, anomalies, or trends in the data, which may hold significant insights or pose potential risks.

4.1.2. Data parser

The data parser module receives the raw data streams from the data procurement agent and performs essential preprocessing tasks to transform the data into a structured format. It converts the raw data streams into a semi-structured CSV format, which is widely used for data representation and interoperability. During the parsing process, the data parser extracts relevant information from the raw data, such as sensor readings, timestamps, and metadata. It organizes this information into appropriate fields and columns within the CSV file, ensuring that the data is properly labelled and structured for further processing. The CSV file generated by the data parser contains valuable properties and tags associated with the data. These properties provide essential context and characteristics of the data, while the tags, such as outlier or usual, serve as indicators for subsequent stages in the data flow pipeline.

4.1.3. Data clarification

The data clarification module focuses on establishing the novelty of the data that has been identified as outliers in the previous stage. It aims to validate the outliers by comparing them with existing data in the local database maintained by the fog nodes. To achieve this, the data clarification module employs a search algorithm that queries the local database for similar or related data instances. This comparison enables the module to determine if the identified outliers are indeed unique or represent known patterns within the dataset. By confirming the novelty of the outlier data, redundancy, and duplication in the data collection process can be minimized, ensuring efficient utilization of storage resources and maintaining data integrity. Once confirmation is obtained and the outlier data is validated as unique, the data clarification module proceeds to forward this data to the subsequent component in the data flow pipeline. This confirmation step ensures that only relevant and distinct outlier data is further processed, preventing unnecessary redundancy or duplication in subsequent analysis and interpretation stages. The combination of the data procurement agent, data parser, and data clarification modules establishes a systematic and efficient approach to handling raw data in fog computing environments. These modules collectively contribute to data refinement, organization, and validation, enhancing the accuracy and reliability of subsequent data analysis and decision-making processes.

4.2. CSV to RDF transformation

This phase uses the CSV to RDF engine to create the RDF from the data acquired in the procurement phase. This phase is crucial as the RDF triples generated from this phase are the ontology enrichment phase benchmark. Data from the data clarification is received using RDF transformation. As was already indicated, this area only accepts data marked as an outlier, indicating that it is a new entry into the system and requires updating in the ontology. The three primary components of the CSV to RDF transformation step are the CSV parser, metadata annotation, and generate RDF. The RDF generating part is vital to the framework’s ability to be dynamic since the output from this module serves as a guide for adding new data to the ontology. The output from this part contributes to achieving the second research goal, which is to enable the addition of new data to the ontology.

- CSV parser: the CSV parser module in the transformation phase processes the CSV files and extracts the embedded metadata. It scans each row and column in the CSV table to gather relevant metadata associated with the data. This metadata includes information such as column names, data types, and constraints. The CSV parser ensures that the metadata is captured accurately and serves as a foundation for subsequent processing steps.

- Metadata annotation: the metadata annotation module takes the extracted metadata from the CSV Parser and creates an annotated tabular data model. It annotates the tables with additional information, such as names, titles, descriptions, and semantic annotations. This step adds semantic meaning to the data by enriching it with metadata that describes it is structure, semantics, and relationships. The annotated tabular data model enhances the understanding and interpretation of the data during further processing.

- RDF generator: the RDF generator module receives the annotated tabular data from the metadata annotation module and transforms it into RDF triples. It leverages the annotated information to create RDF representations of the data, which consist of subject-predicate-object triple statements. The RDF triples capture the semantic relationships and attributes of the data, enabling interoperability and
integration with other semantic systems. The RDF generator ensures that the generated RDF triples adhere to the ontology’s structure and semantics.

The CSV to RDF transformation phase serves as a pivotal step in the ontology enrichment process. By converting CSV data into RDF triples, it enables the integration of heterogeneous data sources, enhances data interoperability, and facilitates advanced reasoning and knowledge discovery. The resulting RDF triples become a key resource for dynamic and context-aware applications in the IoT domain, supporting efficient data integration, analysis, and decision-making.

4.3. Ontology enrichment

The proposed model incorporates an ontology enrichment phase using the learn-and-grow technique to ensure the ontology remains up-to-date and relevant in the context of the IoT ecosystem. This phase is essential to accommodate the ever-increasing volume of data and the diverse nature of IoT devices. It focuses on dynamically incorporating new terms, concepts, and relationships into the ontology through a combination of search and update techniques. Figure 6 illustrates the stages of the ontology enrichment module.

![Figure 6. Illustrates the process flow of the update ontology module](image)

- Search triple query: the search triple query module plays a crucial role in identifying the appropriate class within the ontology to classify the new data. When RDF triples are received, a keyword identification process is initiated. The subject of the RDF triple is used as the initial keyword, and a search operation is performed within the ontology to find a matching class. If a matching class is found, it is details are passed to the next stage. In cases where the initial search is unsuccessful, subsequent keywords derived from the subject are sequentially searched until a suitable class is identified. The search process is considered complete when both the subject and object keywords fail to find a match.

- Query result analysis: the query result analysis module analyzes the results obtained from the search operation to determine the accuracy of the identified class. A comparison and lookup process is performed, leveraging existing database resources, to validate and refine the selection of the correct class to be updated in the ontology. In situations where multiple classes are identified, an additional step of referring to the ontology is carried out to ensure the most appropriate class is chosen.

- Class verification: the query result analysis module analyzes the results obtained from the search operation to determine the accuracy of the identified class. A comparison and lookup process is performed, leveraging existing database resources, to validate and refine the selection of the correct class to be updated in the ontology. In situations where multiple classes are identified, an additional step of referring to the ontology is carried out to ensure the most appropriate class is chosen.

- Ontology enrichment: the ontology enrichment module receives the identified class details from the previous stages and utilizes this information to update the ontology. The new data is added to the ontology as either a class, subclass, or relation, based on the keyword used during the search operation in the search triple query module. For instance, if the keyword used for detecting the class is the subject, the object and predicate of the RDF triple will be added as a subclass and relation, respectively, within that class in the ontology.

The ontology enrichment phase is a critical component of the proposed model, ensuring the ontology remains adaptable and reflects the evolving nature of the IoT ecosystem. By incorporating new terms, concepts, and relationships dynamically, the ontology becomes more comprehensive and capable of capturing the semantics and relationships of the IoT data, enabling effective knowledge discovery, reasoning, and decision-making within IoT applications.

5. EXPECTED RESULT

The projected outcomes for the analysis and functionality of the proposed fog-based conceptual model are described in this section. First, we perform theoretical analysis by comparing semantic-fog,
IoT-Sim, and Open-IoT performance in terms of network usage, service latency, power consumption, and total cost. Then we will evaluate the expected performance of the proposed mechanism. Then we will evaluate the expected performance of the proposed mechanism. These parts will be described as follows.

5.1. Performance metrics

5.1.1. Service delay

The period of time between when an IoT device submits a service request and when it receives the response to that request is referred to as service delays in the IoT. The sum of the processing, propagation, and transmission times for the request and answer is the service delay. The IoT devices send their data packets to fog nodes for immediate processing and to the cloud for additional processing as needed.

5.1.2. Power consumption

The computation and data transmission from IoT nodes to fog nodes and from fog nodes to the cloud account for the majority of the power consumption in the proposed method. The sum of the energies needed to process each unit data packet at the fog nodes and cloud, along with the energy needed to transmit this data, is used to compute the overall amount of power consumed by this mechanism.

5.1.3. Network usage

The bandwidth utilization of an IoT network is the quantity of bandwidth needed to send packet data from a source point to a destination. The transfer of redundant and meaningless data packets between the cloud and the IoT layer via fog device, and vice versa, results in bandwidth being wasted in IoT systems. The sum of bandwidth used to send data from the IoT device to the fog node and the fog node to the cloud is used to compute the total bandwidth used one way.

5.1.4. Total cost

Calculating the combined cost of storage, computing, and connectivity allows an IoT system to evaluate it is overall cost. The entire cost of each cloud and fog component is included in the IoT system’s overall computing cost. Instead of using dollars or rupees, the entire cost is defined in the form of million instructions per second (MIPS).

5.2. Performance analysis

This section discusses about the overall comparative performance analysis of all the solutions discussed based on network usage, service latency, power consumption, and total cost. Based on our expected result, overall the proposed double skin façade (DSF) model performed around 88% better in terms of service delay compared to IoT-SIM and Open IoT, as shown in Table 2. The total processing time of DSF model improved significantly compared to IoT-SIM and Open IoT due to the load-offloading mechanism introduced in the DSF, which helps reduce the processing time by managing the load. DSF model has a reduced fog layer compared to semantic fog which results in a 50% reduction in service delay of the DSF model. The power consumption and network usage of DSF mechanism are expected to reduce by 50% and 86% compared to IoT-SIM and Open IoT due to pre-processing of information in the fog layer which removes the redundancy of data transmitted. While there is a slight improvement which is 2% for power consumption and 30% in network usage compared to the semantic fog model as there is a lesser fog node processing level in the DSF model. While the total cost of implementing DSF was around 80% reduced compared to the IoT-SIM and Open IoT model while 30% reduction compared to the semantic fog model. Making use of single-layer fog nodes with the load offloading mechanism helps in improving the performance of semantic interoperability in IoT.

<table>
<thead>
<tr>
<th></th>
<th>[55]</th>
<th>[57]</th>
<th>[20]</th>
<th>[60]</th>
<th>DSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service delay</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Power consumption</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Network usage</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Total cost</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
</tbody>
</table>

6. CONCLUSION

In this paper, we address the important issue of semantic interoperability in IoT applications. Our study underlines the critical requirement for establishing semantic interoperability among various businesses in order to fulfil the goals of IoT systems. Despite significant research efforts in this area, the issue of dynamic interoperability has only received scant attention in the literature so far. We have presented a...
fog-based conceptual approach to close this gap and help the IoT interoperability industry advance. Our methodology intends to promote dynamic semantic interoperability in IoT by using the characteristics of fog computing. The processing power required to achieve dynamic semantic interoperability is made possible by our model’s single-tier fog layer. The constraints of static interoperability approaches are overcome by this innovative method, enabling IoT devices to adapt and effortlessly exchange information in real-time.

While our study makes an important contribution to the subject of IoT semantic interoperability, we must accept it is limits. First, the suggested fog-based conceptual model has been verified through simulation and small-scale experimentation; nevertheless, additional real-world implementations are required to evaluate its scalability and performance in different IoT scenarios. The efficacy of the approach is also dependent on the availability of dependable fog computing infrastructure, which may not always be available depending on the deployment scenario. These restrictions offer priceless information for upcoming studies and advancements. There are many chances to improve and build upon our work as the IoT industry continues to develop. Future research projects can concentrate on carrying out thorough empirical analyses of the fog-based model in various IoT scenarios, examining optimizations to improve its scalability, and researching methods to deal with the difficulties of fog infrastructure availability.

Finally, by introducing a fog-based conceptual paradigm, our paper contributes to the growth of dynamic semantic interoperability in IoT applications. We stress the need for dynamic interoperability in maximizing the potential of IoT systems as well as the necessity for complete semantic interoperability solutions. In order to encourage additional developments in the area of IoT interoperability and ultimately promote a more interconnected and effective IoT ecosystem, we strive to address the restrictions and open up new research directions.

REFERENCES


Proposed fog computing-enabled conceptual model for semantic ... (Devanmekai Nagasundaram)


BIOGRAPHIES OF AUTHORS

Devamekalai Nagasundaram is a Software Enabling and Optimization Engineer at Intel Microelectronics (M) Sdn. Bhd and currently pursuing a Ph.D. degree at the National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia. Her research interests include internet of things, IPv6, semantic interoperability, fog computing, and Industry 4.0. She can be contacted via email at: mekalai.dev@gmail.com.

Selvakumar Manickam is currently working as an Associate Professor at National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia. His research interests include cybersecurity, internet of things, Industry 4.0, and machine learning. He has authored and co-authored more than 160 articles in journals, conference proceedings, and book reviews and graduated 13 PhDs. He has 10 years of industrial experience prior to joining academia. He is a member of technical forums at national and international levels. He also has experience building IoT, embedded, server, mobile, and web-based applications. He can be contacted at email: selva@usm.my.

Shams Ul Arfeen Laghari is a computer scientist and dedicated researcher, who currently serves as a Postdoctoral Researcher at the National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia (USM). He holds a Ph.D. in Internet Infrastructures Security from NAv6 at USM, consistently demonstrating an unwavering commitment to academic excellence. Dr. Laghari’s academic journey includes institutions such as the University of Sindb, Jamshoro, Pakistan, where he earned a B.Sc. (Hons.) and M.Sc. in computer science, and PAF-KIET in Karachi, Pakistan, where he obtained an M.S. in computer science. Over his 24-year career in teaching and research, he has contributed at both undergraduate and graduate levels. He has notably held positions as the Head of the Department of Computer Science and as an Additional Director of the quality enhancement cell (QEC) at Hamdard University in Karachi, Pakistan. Dr. Laghari’s research interests span various domains, including cybersecurity, IoT, Industry 4.0, distributed systems, cloud computing, and mobile cloud computing. He can be contacted at email: shamsularfeen@gmail.com.

Shankar Karuppayah received the B.Sc. degree (Hons.) in computer science from Universiti Sains Malaysia, in 2009, the M.Sc. degree in software systems engineering from the King Mongkut’s University of Technology North Bangkok (KMUTNB), in 2011, and the Ph.D. degree from TU Darmstadt, in 2016. His Ph.D. dissertation is titled “Advanced Monitoring in P2P Botnets.” He has been a Senior Lecturer at the National Advanced IPv6 Centre (NAv6), Universiti Sains Malaysia, since 2016. He has also been a Senior Researcher/Postdoctoral Researcher at the Telecooperation Group, TU Darmstadt, since July 2019. He is currently working on several cybersecurity projects and groups, such as the National Research Center for Applied Cybersecurity (ATHENE), formerly known as the Center for Research in Security and Privacy (CRISP). He can be contacted at email: kshankar@usm.my.