Journal of Reliable Intelligent Environments
https://doi.org/10.1007/s40860-024-00224-0

ORIGINAL ARTICLE

# A multilevel graph approach for IoT-based complex scenario management through situation awareness and semantic approaches

Mario Casillo[1] · Francesco Colace[1] · Angelo Lorusso[1] · Domenico Santaniello[1] · Carmine Valentino[1]

Received: 28 January 2024 / Accepted: 16 May 2024


**Abstract**
Modern reality and the environments in which we live are characterized by technology components geared toward automatic management through pervasive services. Thanks to the advent of the Internet of Things, such environments can provide information such as pollution levels, public transport conditions, efficiency of energy distribution networks, and identification of suspicious activities by generating complex scenarios. The profitable management of such scenarios can be performed through context modeling and methodologies that can extract and understand environmental information by preventing certain events through artificial intelligence techniques by increasing Situation Awareness. This paper focuses on developing a methodology with predictive capabilities and context adaptability for managing complex scenarios. The use of semantic and graph-based approaches, unlike many approaches used, leads to better integration of knowledge, resulting in improved system performance. In addition, such approaches allow understanding of what is happening in the system at a given time, enabling manipulation and integration of semantic information. Graph-based approaches chosen for this purpose are Ontologies, Context Dimension Trees, and Bayesian Networks, which are able to support the end-user or expert user in handling complex scenarios. The proposed methodology has been validated and applied to real complex scenarios based on the IoT paradigm. The proposed approach validation was conducted using open data from the city of London; a practical scenario case study was conducted in the field of automated management of a Smart Home. In both cases, the system achieved promising results.

**Keywords** Artificial intelligence · Bayesian network · Complex scenarios · Internet of Things · Ontology · Situation awareness

## 1 Introduction

Modern reality is characterized by a vital technological component oriented to providing intelligent and pervasive services. Through modern smartphones, people are always connected to the Internet, accessing and exchanging considerable information. Even the environments and the services are managed, sometimes autonomously, through technological and pervasive systems. This concept has led to modern complex environments, one of which is represented by Smart Cities. Smart Cities are designed to optimize resources for economic growth, environmental efficiency, and stability by utilizing integrated technological solutions. Contemporary settings use various data sources, including traffic congestion metrics, power usage figures, and records of public safety incidents, to enhance urban services [1]. A crucial concept that drives the creation of these intricate environments is the Internet of Things (IoT) [2–4]. The term IoT denotes the concept that objects, or "devices," are connected and able to exchange information among themselves and humans. Through data acquisition and exchange, these devices improve the awareness of the entities and environments in which they are located. Specifically, the IoT allows for the characterization of monitored environments by describing their main features through acquired data [5]. In this sense, IoT permits to reach a digital copy of the environment in which devices are located. However, this paradigm, oriented toward exchanging information, allows information generation according to the Big Data phenomena. In fact, ever-increasing digitization and the attempt of IoT to represent those complex scenarios involve producing specific data that conventional methods cannot process. It is possible

✉ Domenico Santaniello
dsantaniello@unisa.it

1 University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, SA, Italy

to define a scenario as a set of conditions in which systems, humans, and the environment interact [6]. Usually, systems work to support humans in performing a task; therefore, it is used for decision-making. A scenario became complex where the interaction among these three entities requires a significant effort to decide because of the number of variables involved in the decision process. Therefore, complex scenarios can involve more different entities, requiring systems to expend additional effort to support humans [7].

This development poses several questions regarding using this data and the capacity to process it to enhance the quality of living spaces. Could this information be leveraged to turn scenarios from complex to useful for humans?

Properly filtered data can be used to achieve the Context-Awareness [8–10]. An in-depth comprehension of the context is critical to choosing or designing a suitable framework, which allows extracting relevant information to process data, model reality, and provide answers. However, describing the context is not enough since performing actions to improve the environment's livability is necessary. For this reason, it is possible to exploit the Situation Awareness paradigm [11, 12]. This concept refers to systems able to extract and understand environmental information and predict the occurrence of certain events. This paradigm can be improved by artificial intelligence techniques oriented to machine learning models, leading systems to perform autonomous decisions. Many approaches in the scientific literature address the management of complex scenarios in which several machine learning and deep learning techniques are employed. However, many of these approaches are complex. In many cases, even if the algorithms provide understandable results, it is difficult for users to visualize and understand the processes that lead to the resolution of the problem. Therefore, it would be interesting to introduce a methodology that integrates contextual data and manages analysis and prediction processes available to users.

This paper focuses on methodology development with predictive capabilities and context adaptability for managing complex scenarios. The added value of the proposed approach lies in integrating contextual analysis into semantic values using graph-based techniques, as detailed subsequently. Specifically, leveraging context to manage domain knowledge significantly facilitates a probabilistic analysis to enhance system performance [13, 14]. Additionally, adopting selected graph methodologies enhances the explainability of the decision-making process. Graph-based approaches allow an understanding of what is happening in the system at a given time, allowing the manipulation and integration of semantic information. The graph approaches chosen for this purpose are Ontologies [15], Context Dimension Trees (CDT) [16], and Bayesian Networks [17] Ontologies, which facilitate the formal representation of specific domains, enabling collaboration between humans and machines. Specifically, they offer an explicit specification of conceptualizations, abstractions, and simplifications of the environments in which systems are designed to assist humans [15]. Ontologies should be explicit, coherent, and extendible to reach this goal. This study leverages task ontologies designed around application domains, which could provide enhanced support and improve system performance [18]. A Context Dimension Tree is a specific model for context representation, representing all possible contexts. Bayesian networks are probabilistic graph models that can predict the occurrence of specific events. In summary, the proposed approach, defined as the Multilevel Graphic or MuG approach, exploits three graphical approaches: Ontology, Context Dimension Tree, and Bayesian Networks. The system aims to provide answers and applications that can impact the real environment, helping the end-user or expert user manage complex scenarios. The main contribution of this work is the introduction of a methodology based on the three graphs (Context Dimension Tree, Ontology, and Bayesian Network) to manage contextual data to achieve situation awareness. In particular, by integrating the context identified through the CDT with the semantic relations of the Ontology, we can model the structure of the Bayesian Network to fit complex scenarios. In addition, this work includes the proposed methodology in an architecture that allows humans to be supported through services developed through data elaboration. The proposed method has been validated and applied to several case studies based on the IoT paradigm, obtaining promising results. Therefore, this work's main contribution combines three graph approaches to manage complex scenarios and support users in the decision-making phase.

The document is organized as follows. The next section reports related works based on the Internet of Things approach to achieve Situation Awareness; Sect. 3 focuses on the proposed approach; Sect. 4 reports an experimental campaign designed for the proposed approach testing; Sect. 5 presents the conclusions.

## 2 Related works

This section outlines the progression of the Internet of Things (IoT), exploring attaining Context and Situation Awareness via IoT and applying graph-based methodologies like Ontologies, Context Dimension Trees (CDT), and Bayesian Networks in research.

The IoT concept encapsulates the interaction and data transfer between devices and humans over networks. This constant data flow catalyzes a prominent phenomenon in managing intricate scenarios: Big Data. Crafting algorithms and solutions to decipher and engage with this extensive data corpus is a crucial contemporary challenge [19]. Data management has expanded across three axes: volume, velocity, and variety, collectively known as the 3Vs, which are pivotal in defining Big Data infrastructure characteristics [20]. Merging IoT with Big Data paradigms profoundly impacts various facets of daily life and user behavior. This integration aims to harness and interpret data to meet critical human-centric goals like enhanced quality of life, economic progress, and improved environmental sustainability and resilience. These objectives are furthered by designing urban spaces that leverage integrated technologies and optimize resources, advancing crucial areas such as mobility, communication, economy, employment, environment, and governance. Moreover, in the contemporary landscape of Industry 4.0, the adoption of IoT has prompted a reevaluation of all manufacturing processes. Mainly, sensors are employed to monitor and control processes in real-time, deriving valuable insights for supply chain, maintenance, and production enhancement [2, 21].

The concept of context has been extended to all aspects that characterize the user concerning the capabilities that information systems can deliver. For this reason, complex and general context models have been proposed to support context-aware applications, which leverage them, for example, to adapt interfaces, tailor a set of application-relevant data [8], increase the accuracy of information retrieval, discover services, and compose services [9], make user interaction implicit, and create intelligent environments [22]. Several other examples of context-aware systems leverage different technologies for context management. In particular, Sarker et al. [10] examine the context in mobile services by proposing a comparison of supervised machine learning with the goal of classification with graphical techniques for context representation. This work emphasizes the importance of context-aware systems in providing personalized services to users exploiting mobile devices capable of managing contextual domains through acquired information. Ranganathan and Campbell [23] propose an infrastructure that obtains contextual information through sensors. The system processes information through ontology, which allows the management of the information acquired. The proposed model allows, therefore, not only the acquisition of contextual information but also the automatic processing of the context. Gu et al. [24] consider context management using ontologies. In particular, the goal is to provide personalized services to users based on processed contextual information. The planning of SOCAM (Service-Oriented Context-Aware Middleware) architecture exploits an additional property introduced in the ontology and uses of contextual information, distinguishing them into sensed, defined, and deduced. Sensed refers to the context acquired through physical smart devices such as sensors. Defined refers to tailoring a user and designing to meet his specific needs. Finally, processing the direct context allows for obtaining the deduced context. Instead, Yoon et al. [25] propose a Bayesian context-based approach to handle multi-object tracking. Stenneth et al. [26] also exploit a Bayesian methodology for transportation mode detection. The location context allows them to obtain better accuracy for detecting external services, e.g., public transportation. Kooij et al. [27] exploit context-awareness computing to model a Dynamic Bayesian Network that aims to predict the path per pedestrian.

The Situation Awareness concept covers many application areas, and diverse scenarios exploit it. The use cases are usually complex; one of the most popular consists of aviation. In particular, in the aviation field, SA assumes a relevant role in the training of pilots [28]. They must adequately acquire information from the surrounding environment, process information, and predict what will happen in the short and long term. Indeed, the aviation field conceives the most widely used formal models developed for SA. Gluck et al. [29] propose a concurrent verbal protocol supported by eye movement data to determine the degree of identified Situation Awareness. Other application areas, such as road driving, also take advantage of this approach [30]. Liu et al. [31] propose a model for quantitative analysis of Situation Awareness utilizing probability theory, focusing on conditional probabilities in the context of Bayes' Theorem. This model, determined as the attention allocation model, also proposes a quantitative analysis of SA. The advent of the Internet of Things (IoT) [3] contributes to environmental monitoring, allowing for a more efficient determination of SA. Situation awareness is a significant advantage when working with the Internet of Things paradigm. The increasing use of smart devices and the development of wireless technology can represent an essential advantage for achieving the first two levels of SA naturally and efficiently. Many approaches in the literature leverage the potential of IoT in SA; for example, [4, 32] proposed systems that can combine graph approaches for context representation, such as Ontologies, for knowledge construction. In this domain, other approaches can exploit Ontologies for user-defined role construction to increase SA. In detail, after acquiring heterogeneous information provided by sensors, the data is processed and integrated into the ontology that allows the system to provide personalized alerts based on user-defined rules [33]. In an attempt to support users in risk scenarios, Krytska et al. [34] propose an IoT-based system that can handle dangerous scenarios by suggesting safe sites to people during an emergency. The system takes advantage of IoT technologies integrated with the SA process. In this way, Glowacka et al. [11] propose a model in which IoT-based Smart Devices can acquire Situation Awareness and be ready to react to possible threats. This model exploits several modules: The information acquisition and analysis module, the inference module that evaluates objects based on observations and recommendations, the classification module aimed at the final evaluation, and the Reaction module that exploits previous modules to enable possible reaction threats.

Other recent works related to employing graph approaches are present in the literature. Moreover, the objective and the application fields are various. For instance, the authors [35] in the paper exploit Ontologies and Bayesian Networks for introducing a methodology that exploits a probabilistic approach to model knowledge. Instead, Ellouze and Belguith [36] takes advantage of an ontology to improve the dependence among variables of a Bayesian network in an application to social media. This application field usually integrates social context to improve semantic analysis through ontology, as reported in [37]. Riali et al. [38] incorporate fuzzy logic for ontological and Bayesian analysis in the medical field. Finally, in [39], authors exploit a probabilistic ontology for supporting people by developing an intervention system. In particular, they consider several factors that enrich the ontology that exploits Bayesian networks. As reported in the scientific literature, there are various proposals capable of leveraging the IoT paradigm to attain Context and Situation Awareness for event management. Some of the proposed articles also utilize graph-based approaches such as Ontologies and Context Dimension Tree and probabilistic graph approaches like Bayesian Networks or other Artificial Intelligence techniques. Indeed, developing a methodology that combines these tools to gather contextual information and make predictions to navigate modern complex scenarios remains challenging. Therefore, this study aims to introduce a method for integrating three graph-based approaches, incorporating representation methodologies for context awareness (via a Context Dimension Tree), Ontologies, and Bayesian Networks (Table 1). The following section will detail the proposed method, which, harnessing data from sensors, can model context by exploiting three graph-based approaches. The acquired information, in conjunction with probabilistic graph-based approaches, enables the support of various user profiles, thereby facilitating the management of complex scenarios.

**Table 1** Related work using Context-Awareness, Ontology, and Bayesian Network approaches comparison

| Paper | Context-awareness | Ontology | Bayesian network |
|---|---|---|---|
| Sarker et al. [10] | X | | |
| Ranganathan and Campbell [23] | X | X | |
| Gu et al. [24] | X | X | |
| Yoon et al. [25] | X | | X |
| Stenneth et al. [26] | X | | X |
| Kooij et al. [27] | X | | X |
| Mozzaquatro et al. [4] | X | X | |
| Kolbe et al. [32] | X | X | |
| Freedman et al. [35] | | X | X |
| Ellouze and Belguith [36] | | X | X |
| Ben Fraj and Nouri [37] | X | X | |
| Riali et al. [38] | | X | X |
| Gueddes and Mahjoub [39] | | X | X |

## 3 The proposed approach

This paper aims to introduce a new methodology that can support users in managing complex scenarios based on IoT using graph approaches. This section focuses on presenting a methodology and its architecture that is adaptable to the context and situation and able to perform predictions to assist humans. The architecture provides the structural basis for applying the novel methodology in real complex scenarios to achieve situation awareness and provide predictions that enable decision-making.

Many approaches in the scientific literature rely on machine learning techniques to manage complex IoT scenarios. Among the most used methodologies include Support Vector Machine, K-Nearest Neighbors, Artificial Neural Networks, and all the types related to Deep Learning [40–42].

However, these approaches often do not adequately exploit contextual information from the environment and use methodologies that can usually be represented as black boxes where the user cannot intervene in decisions and choices [43–45]. The paradigms presented in the previous chapters related to the Internet of Things, Context Awareness, and Situation Awareness are crucial in this field. In fact, smart devices belonging to the Internet of Things paradigm can acquire a wide range of information related to a user, domain, and environment. Context Awareness could exploit the specific conditions acquired through the 5W + 1H model to personalize further needs and actions related to the user. Finally, the models associated with the Internet of Things and Context Awareness are integrated into perception and understanding phases typical of Situation Awareness. The latter provides the prediction assumptions to the system.

In summary, the proposed methodology aims to exploit the IoT paradigm to monitor the reference environment and approaches based on Context-Awareness and Situation Awareness, which are able to formally map the reference domain and provide answers in terms of predictions to support users in managing complex scenarios.

Environments for managing information, or in general pervasive data models, can benefit from context depiction strategies and be further augmented through probabilistic methods like Bayesian Networks (BN) [46–48]. BNs present an analytical structure that proficiently aids users by modeling data. The merits of probabilistic methods include the capability to depict intricate systems, forecast outcomes, diagnose issues, assess event probabilities, update predictions with new data, handle diverse variables, and provide a comprehensible, concise graphical format [49]. Additionally, enhancing these benefits through context representation methods, such as the Context Dimension Tree (CDT), is noteworthy. The CDT is instrumental for applications that select specific areas of interest [50]. Context-aware strategies generally lead to more effective organization and personalized data distribution [51]. Ontologies are another prevalent method for reflecting reality and significantly support pervasive Context-Aware systems [52]. There's a notable synergy between Ontology and Bayesian networks [53], with possibilities to construct BNs using Ontologies [54] and, conversely, to automatically generate Ontologies through BNs [55].

Considering the context, as mentioned earlier, depiction techniques, and the capabilities of BNs to identify potential events through experimental data and probabilistic methods, it's essential to introduce strategies that can dynamically manage context to enhance complex scenario handling. This study aims to detail and elucidate a Multilevel Graph (MuG) Approach, merging Ontologies, Context Dimension Trees, and Bayesian Networks. This fusion seeks to assist expert and everyday users in addressing their needs and providing recommendations to augment the quality of life in complex scenarios. In particular, the contextual data acquired through IoT devices enables the Context Dimension Tree to determine the context in which the system works. Therefore, Ontology allows the representation of the knowledge domain [56] and takes advantage of the context to identify the appropriate semantic relations in the specific context. Identifying concepts, attributes, and relations exploits the application domain and allows the definition of the Ontology. Experts in the application field support these steps in this phase. Thus, the Ontology is constructed, configured, and validated. The role of ontology is crucial, and it represents the application field by consulting domain experts. Ontology relations enable the construction of constraints on the possible or impossible dependency among variables, namely, the Bayesian Network (BN) nodes. Therefore, the system improves performance by integrating structural learning algorithms with the identified constraints.

The following subsection will present the general architecture for including the given system in real complex scenarios. Then, the formalization of the proposed MuG Approach will be reported. In particular, the theoretical description will focus on introducing context acquired through IoT devices and managed through the Context Dimension Tree (CDT) in the semantic analysis of the Ontology. The system will use the first two graphs to enhance the Bayesian Network structure, taking advantage of semantic-contextual constraints.

### 3.1 A proposed architecture

Developing an architecture based on IoT systems that integrates context and situation awareness approaches for managing complex scenarios is challenging. Based on IoT systems, the architecture must be supported by typical IoT paradigm multilayer architectures. In this case, the proposed architecture will be based on the most generic and high-level architecture consisting of three layers: Perception, Network, and Application [57]. The proposed architecture must be scalable, interoperable, distributive, and able to operate with low-resource devices [58]. In addition, it must be able to manage Big Data coming from different heterogeneous sources. Moreover, it must extract and manipulate semantic information using context and situation management tools. To this end, it is possible to refer to Endsley's model shown in Fig. 1 [59], which represents a fundamental starting point concerning the characteristics of the presented system. The purpose of the system, in fact, is to acquire data from the surrounding environment, store, process, and use data by making it available to different categories of users. In particular, users can obtain applications from the system for data visualization and interaction with the system and the external environment. Enclosing the features of the proposed system in a general architecture can be represented in Fig. 1.

The proposed architecture provides several blocks with different functionalities. The top layer of the system is the Environmental Data Acquisition Layer. This layer collects methodologies and techniques for retrieving data from various sources. At this stage, the purpose is to store and filter as much data as possible, homogenize them, and make them available to the system and users. IoT sensors and actuators and API or Open services form data acquisition sources. IoT sensor nodes are capable of acquiring information from the reference domain; these data are acquired by specific sensors that can report information about general environmental conditions or the asset under consideration, depending on the application cases. In the IoT sensor network, actuators can also interact with the environment, i.e., interacting with

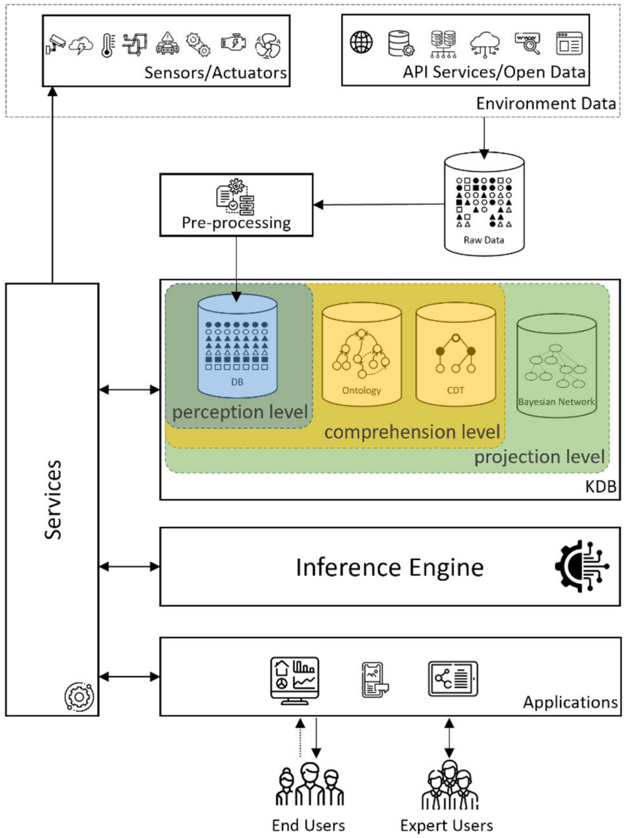


**Fig. 1** General system architecture

environmental conditioning systems, triggering alarms, and activating security actions remotely. The system is able to interact with API and Open data services, integrating data acquired from IoT sensors. These data enrich the system in different aspects, i.e., services that can provide information on events, weather, booking ticketing services, and descriptive insights into the assets under study. All data stored in the top layer are acquired, pre-processed, and transferred to the next layer: the Knowledge Database. In order to define the proper use of the acquired data, the information is structured in the Knowledge Database, making it valuable and ready for processing.

The knowledge database represents the information core of the proposed architecture. In fact, the three phases of Situation Awareness proposed by Endsley [59] are performed inside this layer. The KDB collaborates with the Inference Engine to construct and extract knowledge from the graph structures included in the database. In particular, the perception phase is developed through the mentioned graph structures, representing the structured knowledge base. In addition, the comprehension phase is performed by building and exploiting graph structures that are able to interpret and manage context, such as Ontologies and Context Dimension Tree (CDT).

Ontologies facilitate the detailed illustration of a particular field by establishing a formal, shared, and distinct representation. Precisely, they form an axiomatic framework that is articulated using descriptive logic. In their graphical format, ontologies can effectively communicate with instruments like the Context Dimension Tree (CDT) and Bayesian Networks [60]. The CDT is adept at organizing all conceivable contexts within a hierarchical tree structure. This tree categorizes nodes into two types: dimensional and conceptual. Dimensional nodes outline the potential dimensions within the application field, while conceptual nodes accumulate all the possible values a dimension might take. This arrangement allows for efficient interfacing and querying of databases to select appropriate services based on the chosen context.

The third stage of Situation Awareness, the projection phase, is executed using probabilistic graphical structures like Bayesian Networks. These networks endow the system with a vital feature: the capacity for prediction. These models utilize Bayes' Theorem to forecast the likelihood of an event's occurrence by representing a collection of stochastic variables and their conditional interdependencies via a Directed Acyclic Graph (DAG). Additionally, due to their structured nature, Bayesian networks can aptly connect with both the CDT and Ontologies, creating a robust framework for data analysis and decision-making [61].

To understand in detail the interaction between the graph structures for the MuG Approach operation, it is possible to refer to Fig. 2, where the workflow of the proposed methodology is shown. The workflow starts from the data from the domain of interest provided by the sensors spread in the environment. A Task Ontology can be developed from the information collected and through the contribution and comparison with experts [18, 62]. Subsequently, a CDT must be designed to describe all the possible contexts within the reference domain. At this point, the Ontology and the CDT are combined to obtain semantic relations, constituting semantic constraints valuable for building the Bayesian Network. The semantic constraints, in combination with structural learning algorithms [17] and the available data, develop the Bayesian Network to predict events. To summarize, in this way, the need to perform a decision in a given context can be satisfied by using the correct information provided by the architecture. This information is characterized by innovative elements based on knowledge management and organization, formal context representation, and inferential approaches.

After the knowledge construction and organization phases, in which the KDB and the Inference Engine are involved, the Application block exploits the acquired information. The application layer contains all the possible applications that users can use. Different applications suitable for various user categories can coexist in this layer. In detail, expert users can benefit from the technical and specific information provided and are able to interact with the system.

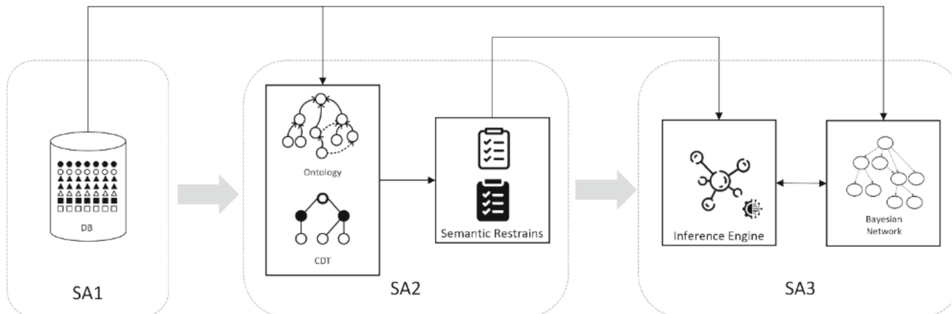


**Fig. 2** Multilevel graph approach workflow

In contrast, standard users can only access selected contents and cannot interact with the system to perform actions; however, on particular occasions, to enrich the KDB, the system can ask for feedback from standard users.

All operations between the KDB, Inference Engine, and Applications layer are performed through the Services Module. This module collects all the services that are able to communicate through the different blocks, allowing the various operations of the system to be processed between the different layers within the KDB, providing predictions, representing information through applications, etc. In addition, this module allows user interaction through actuators.

The following section will detail the formal definition of the proposed Multilevel Graph Approach.

### 3.2 A multilevel graph approach: formal definition

This paragraph allows the formalization of the Multilevel Graph Approach. In particular, it will show how MuG develops the lists of semantic relations to correct the structural learning algorithm to obtain the Bayesian network. In the following, we will deepen the theoretical treatments for Ontology, Context Dimension Tree, and Bayesian Network.

Defined the generic Ontology $O$

$$O = \{C, A, H, R^{+}, R^{-}\} \tag{1}$$

with $C$ set of concepts, $A$ set of attributes, $H$ hierarchical relations, and the sets $R^{+}$ and $R^{-}$ sets denote the relations of axiomatic type ($A_x$), semantic ($R_T$) or non-semantic ($R$) relations, which constitute a dependency or independence link, respectively, among the ontology concepts.

This allows defining the $G_O =< N_O, E_O, r >$ associated with the ontology, where $r$ is the root of the graph, $N_O = C$ is a set of graph nodes and coincides with the set of Ontology concepts, $E_O = \{(n_1, \alpha, n\_2) : n_1, n_2 \in N_O, \alpha \in H \cup R^{+} \cup R^{-}\}$ set of arcs. It contains all links among nodes. These links can be hierarchical or specify dependency or independence between concept nodes. From the definition of the set $E_O$ the labels of the arcs are defined by the relations of the given Ontology.

The previous paragraphs formally presented the Context Dimension Tree (CDT) graphical approach. In particular, the CDT consists of a graph:

$$G_{CDT} = \langle N_{CDT}, E_{CDT}, r \rangle. \tag{2}$$

In the case of the proposed approach, the Context Dimension Tree provides nodes that are in common with the constructed Ontology. Therefore, both the domain d and the sets of concepts $C$ and attributes $A$ of the Ontology, the following sets are defined:

$$C_d \subseteq C, \ A_d \subseteq A \tag{3}$$

such that

$$N_D \cup N_C = N_{def} \cup C_d, \ N_A = A_d \tag{4}$$

with $N_{def} = \{who, why, when, how, \ldots\}$ of predefined dimension nodes necessary for creating the CDT. We note, therefore, that the set of nodes of the CDT is generated through the ontology and the analysis of the domain d.

$$N_{CDT} = N_{def} \cup C_d \cup A_d \tag{5}$$

Moreover, the non-default dimension nodes and the concept nodes of the CDT are chosen among the set $C$ elements of the Ontology. In contrast, the parameters are chosen among the attributes $A$ of the Ontology.

The set of CDT arcs is partitioned into three sets:

$$E_{CDT} = E_{def} \cup E_C \cup E_A \tag{6}$$

where $E_{def}$ the set of arcs connecting the root to the predefined dimension nodes $N_{def}$ and the arcs connecting the latter to the related concept nodes, $E_A$ is the set of arcs connecting the concept nodes with the relative attributes evaluated in the specific contextual domain, while $E_C$ it is the set of the arcs for which there is a connection of hierarchical type between two nodes in the Ontology $O$.

$$E_C = \{(n_1, \alpha, n_2): n_1, n_2 \in \phi_1^{Pr}(C), \alpha \in H\} \tag{7}$$

Thus, the Context Dimension Tree defined via the Ontology takes the following form:

$$G_{CDT} =< N_{CDT},\ E_{CDT},\ r \\ \geq \langle N_{def} \cup C_d \cup A_d,\ E_{def} \cup E_C \cup E_A,\ r\rangle \tag{8}$$

In Ontology $O = \{C,\ A,\ H,\ R^{+},\ R^{-}\}$ specific roles have been assigned to concepts, attributes, and hierarchical relations. At this point, instead, a role will be assigned to the relationships of dependence $R^{+}$ and independence $R^{-}$ between ontological concepts. In particular, those relationships will play a key role in designing the lists of semantic relations. To this end, some preliminary definitions are provided, preparatory to formalizing their use.

**Definition** Let $m_1 \rightarrow_{\alpha_1} m_2 \rightarrow_{\alpha_2} \cdots \rightarrow_{\alpha_t} m_{t+1}$ a path. We define in the graph $G_O$ associated with the Ontology a dependency path, a path that enjoys the property

$$\exists!\ j \in \{1,\ldots,t\} : \alpha_j \in R^{+},\quad \alpha_k \in H''\ ''\forall k \in \{1,\ldots,t\} - \{j\} \tag{9}$$

Instead, a path of independence is defined as a path that enjoys the property

$$\exists!\ j \in \{1,\ldots,t\} : \alpha_j \in R^{-},\quad \alpha_k \in H''''\forall k \in \{1,\ldots,t\} - \{j\} \tag{10}$$

Finally, a classpath is defined as a path that has the property

$$\alpha_k \in H \quad \forall k \in \{1,\ldots,t\} \tag{11}$$

For convenience, it is possible to denote the set of dependency paths $P^{(D)}$ and the set of dependency paths $P^{(I)}$. At this point, it is necessary to define the concepts of dependence and independence between the nodes of the ontology $O$.

**Definition** Let $m_1,\ m_2 \in C$ distinct concept nodes of the Ontology, such nodes of define dependent nodes if and only if $\exists v \in C$ such that $m_1$ e $v$ are connected by a dependency path and $m_2$ e $v$ are connected through a classpath or vice versa.

Instead, nodes are defined to be independent if $\exists v \in C$ such that $m_1$ e $v$ are connected by an independent path, e $m_2$ e $v$ are connected by a classpath or vice versa.

At this point, through the relations just introduced, it is possible to define the sets

$$R^{(D)} = \{(m_1,\ m_2) : m_1,\ m_2\, dependent\ nodes\} \tag{12}$$

of dependent nodes and

$$R^{(I)} = \{(m_1,\ m_2) : m_1,\ m_2\, independent\ nodes\} \tag{13}$$

of independent nodes. The first one constitutes the list of semantic dependence relations, while the second one constitutes the independent ones. Building the Bayesian network involves analyzing the $O$ ontology along with the Context Dimension Tree. The graph that represents the network is denoted with:

$$G_B = \langle N_B,\ E_B\rangle \tag{14}$$

where $N_B$ is the set of nodes, $E_B$ is the set of the edges. The nodes of the network are identified through the analysis of the specific context using the CDT, i.e.

$$\exists \psi_d N_B = \psi_d(C) \tag{15}$$

The generation of the nodes happens through the action of both the Ontology and the Context Dimension Tree. In fact, the latter shares some of the Ontology nodes, as shown above.

The set $N_B$ contains a finite number of nodes, and each can assume a known value within the dataset. This statement is supported by the very construction of the set of nodes of the Bayesian network, carried out through the synergy between ontology and context dimension tree.

The following assumptions are met according to Cooper et al. [63]:

- **Hypothesis 1**: The variables in the database $Z$ are discrete since the construction is done through the dataset and the work done between ontologies and CDT.
- **Hypothesis 2**: Given a belief-network model, the cases are independent.
- **Hypothesis 3**: There are no cases where the variables have unknown values since the nodes are selected through the information provided by the dataset and processed by the preliminary graph approaches.

An in-depth analysis has to be performed on the possible belief networks through the relations introduced among the nodes. In fact, the goal of the proposed approach is to reduce the number of belief networks to be evaluated and to select, among those left available, the most consistent with the dataset exploited through the analysis performed by ontology and context dimension tree. In particular, the Bayesian network B of which the nodes are already known $N_B$ must be determined through the identification of the set of arcs $E_B$ which will include the arcs obligatorily joining the dependent nodes of the list of dependency relations $R^{(D)}$ and must not present arcs connecting the independent nodes $R^{(I)}$ of the list of independent relations. In particular, by definition, the dependency list $R^{(D)}$ represents the set of oriented arcs joining the dependent nodes, while the independence list

$R^{(D)}$ represents the set of arcs connecting the independent nodes.

The target network B will be determined among the belief networks belonging to the set:

$$Q = \{B_i = \langle N_B, E_i \rangle : ND \subseteq E_i, IN \cap E_i = \varnothing\} \tag{16}$$

On this set, the assumptions are assumed to be valid:

**Hypothesis 5** It is possible to determine an order for the variables.

**Hypothesis 6** The probability $P(B_i)$ are equal $\forall B_i \in Q$.

Once determined the conditional probabilities, set a priori through the list of semantic dependency relations, and eliminated the impossible conditional probabilities through the tabu list, the remaining vectors of the possible conditional probabilities $B_P$ can be chosen to respect the conditions set through Ontologies and Context Dimension Trees. For all the possible choices for $B_P$ we assume that the following hypothesis is valid:

**Hypothesis 4** The density function $f(B_P, B_i)$ which links the vector of conditional probabilities $B_P$ to the belief-network structure $B_i$ is uniform.

Denoted with $(\pi_h)$ the instance of the relatives of the random variable $X_h$, this instance will respect both the dependence and independence semantic relations list. In fact, the aleatory variables $X_1,\ldots,X_t$ connected through an arc to the aleatory variable $X_h$ will respect the following conditions:

$$\{(X_1, X_h), \ldots, (X_t, X_h)\} \cap R^{(I)} = \varnothing \tag{17}$$

$$\nexists j \notin \{1, \ldots, t\} : (X_j, X_h) \in R^{(D)} \tag{18}$$

In particular, the second relation emphasizes that the relatives $(\pi_i)$ of the aleatory variable $X_h$ include all dependency relations set via the list of semantic dependency relations.

Recalling that $P((\pi_h^s) \rightarrow B_i)$ represents the probability that in the belief-network $B_i$ variable $X_h$ has as relatives $(\pi_i^s)$ and that we have previously defined the function

$$g(h, \pi_h) = \prod_{j=1}^{q_i} \frac{(r_h - 1)!}{\left(N_{hj} + r_h - 1\right)!} \prod_{k=1}^{r_i} N_{hjk}! \tag{19}$$

The Bayesian network will be determined such that

$$\max_{B_i \in Q} P(B_i, D) = \prod_{h=1}^{n} \max_{\pi_h} P\left(\pi_i^s \rightarrow B_i\right) g(h, \pi_h) \tag{20}$$

In the proposed approach, this computation is performed by modifying the heuristic procedure used by Cooper et al. [63]. In particular, the following pseudocode is executed in which a maximum number of relatives are fixed, and an order on the nodes of the set $N_B$ is set. The set of possible relatives' nodes $(\pi_i)$ is constructed to exploit the list of the relations of independence. Therefore, in the inputs, it is requested the insertion of the set $R^{(I)}$ even if in the pseudocode, it is not used explicitly.

| |
|---|
| **Algorithm** |
| *Input:* $N_B$, ND, ID, node order, u |
| *Output*: a printout of the node's relatives for each node. |
| **for** $\boldsymbol{i = 1,\ldots,n}$ **do** |
| $\pi_i = \{(X,X_i) \in ND\} \subseteq ND$ |
| $P_{old} := g(h,\bar{\bar{\pi}}_h)$ |
| $OK\text{-}to\text{-}Proceed := TRUE$ |
| **while** $\boldsymbol{(Ok\text{-}to\text{-}Proceed = TRUE) AND (\lvert\pi_h\rvert \le u)}$ **do** |
| Let be $z \in Prec(X_h)$ and be $\bar{\bar{\pi}}_h$ which maximizes $g(h,\bar{\bar{\pi}}_h \cup \{z\})$ |
| $P_{new} := g(h,\pi_h \cup \{z\})$ |
| **if** $\boldsymbol{P_{new} > P_{old}}$ **do** |
| $P_{old} := P_{new}$ |
| $\bar{\bar{\pi}}_h = \bar{\bar{\pi}}_h \cup \{z\}$ |
| **else** |
| $OK\text{-}to\text{-}Proceed := FALSE$ |
| write ('Node: ', $X_h$, 'Parents of this node: ', $\bar{\pi}_h$) |

Once the belief network has been determined based on the lists of dependence and independence relations, it will be possible to perform the training phase of the network to determine its weights. In particular, through the dataset, it is possible to calculate the conditional probabilities associated with the dependence relations that identify the Bayesian network built through the synergy work of the Ontology $G_O$ and Context Dimension Tree $G_{CDT}$.

## 4 Experimental results

This section reports the case studies developed using the proposed Multilevel Graph Approach. The first subsection shows a case study used for model validation; the case study is related to the Smart City of London, chosen for reasons of data availability. The second subsection provides another case study related to real life, which concerns supporting users in managing services and resources within a Smart Home.

In particular, the experiments in both study cases exploit IoT devices for acquiring data that enables individuating context through the Context Dimension Tree and improving the construction of constraints through the relations of the ontology. Therefore, the system can identify the appropriate Bayesian Network (BN) structure by integrating constraints with the structural learning algorithm, as suggested by Algorithm 1. This BN allows decision-making. Therefore, this section aims to estimate the system's dependability through accuracy measures and user evaluations. Accordingly, the experimental phase evaluates the introduced methodology overall through the system's reliability in supporting users.

### 4.1 Model validation: London Smart City

This subsection reports the case study that validates the proposed approach. For data availability, this pilot case study is focused on London Smart City, in which, due to open data availability policies, it was possible to obtain a large amount of information. Based on the collected data, the MuG Approach is exploited in this scenario to predict different events, such as traffic accidents, rainfall, and bike-sharing service availability. Several steps are necessary to validate the proposed model. In particular, this section aims to compare the performance of the proposed approach, which involves predicting phenomena through Bayesian Networks built with the support of semantic information derived from the context. The results will be compared with traditional approaches, such as constructing Bayesian Networks with the support of experts in the field or through structural learning algorithms. In detail, the proposed methodology will be validated in an attempt to predict three phenomena in the Smart City scenario: traffic accidents, rainfall, and bike-sharing service availability.

The chosen study area is the city of London; in particular, the complete dataset of information collected refers to the neighborhoods of Islington, Westminster, Hammersmith and Fulham, Lambeth, and Tower Hamlets. Only for these boroughs is it possible to collect enough data to provide an example that allows us to show the ability of the system to predict different events. Depending on the phenomena to be predicted and according to data availability, all areas or only some correspond to London's central areas. The data come from different sensors, providing information in various formats and time intervals. The data is aggregated at different time intervals for the analysis, depending on the event to be predicted. The full dataset contains 12 variables and several records aggregated from 2017 to 2019. The one-year observation period exceeds 35,000 instances. In detail, the information acquired is related to whether information such as Evapotranspiration, Pressure, Radiation, Rainfall, RainRate, Temperature, UV Index, Wind Direction, Wind Speed, and others related to public transport information, such as the use of bike sharing service and vehicle road accidents.

As mentioned above, the Multilevel Graph Approach exploits the semantic relationships that arise from Context-Awareness. Considering the Situation Awareness (SA) paradigm defined by Endsley, through this phase, SA 2, namely, the Comprehension phase, is reached through context description, performed by graph structures such as Ontologies and CDT. These graph structures are designed with expert domain supported by scientific literature-defined approaches. In order to pursue the objective of evaluating the proposed approach, it is necessary to deepen the study of the two graph approaches by focusing on scientific literature desk research to design Ontologies and CDTs that meet the specific needs of context representation. This representation will allow the system to integrate semantic relations functional to build constraints between the data. Due to this, it will be possible to design more reliable Bayesian Networks. The experimental results regarding the first case study are performed to predict three phenomena within the city of London. The phenomena to be expected were road accidents, rainfall, and the availability of bicycle-sharing services.

In all three cases, the proposed method is compared using about 75% of the dataset available as a training set and about 25% as a test set and through the three phases exposed in the paragraph on the experimental approach. In particular, the predictions related to a Bayesian network structure built by a group of experts were evaluated. Using the K2 algorithm [64], the Bayesian network built by machine learning is evaluated, and then analysis is conducted using the proposed methodology.

However, each application preserved its peculiarities. In the case of traffic accidents, an attempt is made to predict the

**Table 2** London Smart City Case of Study: System performance in Accident prediction

| | Low (%) | Medium (%) | High (%) |
|---|---|---|---|
| Step 1 | | | |
| Precision | 79.02 | 87.77 | 77.44 |
| Recall | 91.52 | 74.37 | 71.43 |
| $F_1$-Score | 84.81 | 80.52 | 74.31 |
| Step 2 | | | |
| Precision | 77.25 | 54.32 | 16.23 |
| Recall | 61.38 | 53.74 | 63.64 |
| $F_1$-Score | 68.41 | 54.03 | 25.86 |
| Step 3 | | | |
| Precision | 82.78 | 89.40 | 83.51 |
| Recall | 93.87 | 78.92 | 75.94 |
| $F_1$-Score | 87.98 | 83.83 | 79.55 |

**Table 3** London Smart City Case of Study: System performance in Rainfall prediction

| | Yes (%) | No (%) |
|---|---|---|
| Step 1 | | |
| Precision | 59.96 | 88.77 |
| Recall | 73.47 | 81.05 |
| $F_1$-Score | 66.03 | 84.73 |
| Step 2 | | |
| Precision | 45.90 | 63.17 |
| Recall | 42.06 | 66.71 |
| $F_1$-Score | 43.90 | 64.89 |
| Step 3 | | |
| Precision | 85.17 | 91.06 |
| Recall | 83.84 | 91.85 |
| $F_1$-Score | 84.50 | 91.45 |

probability of accident occurrence by classifying it as Low, Medium, and High risk. In this case, the analysis is conducted by observing the data collected for all areas available by collecting aggregate instances every three hours. In particular, among the 7920 events analyzed, 5960 (about 75%) were used for the training set and 1960 (about 25%) for the test set. The results obtained using the test set, in terms of Precision, Recall, and F1-Score, are reported in Table 2

In the case of rainfall prediction, the variable under consideration, which measures the accumulation of rain during a rainfall event, defines the event itself. For this reason, the parameter is summarized in a Boolean variable (Yes/No). Moreover, in this case, the prediction is improved using data correlation techniques, which allow us to find relationships between data with different observation times. This process increased the system's reliability concerning a difficult-to-predict event, such as rainfall forecasting, by exploiting cross-correlated data as precursors or indicators.

The results were obtained through the dataset from all the central zones aggregated at hourly intervals. In particular, among the 23,760 events analyzed, 17,850 (about 75%) were used for the training set and 5910 (about 25%) for the test set. The results obtained in terms of prediction, recall, and $F_1$-Score are reported in Table 3.

In the third case, the system is tested based on the forecast for using the bike-sharing service in the city of London. In this case, two variables representing the beginning (Bicycle Rental) and the end (Bicycle Return) of the service's use are considered to provide users with the probability of availability or unavailability of the service aggregated into three ranges: Low, Medium, and High. In this case, the results were obtained using data from the Westminster area only at 30-min intervals. In particular, among the 47,520 events analyzed, 35,747 (about 75%) were used for the training set and 11,773 (about 25%) for the test set. In terms of precision, recall, and F1-Score, the results are reported in Table 4

As can be seen from the experimental results shown in Tables 2, 3, and 4, the system can predict the variables with varying degrees of accuracy. In particular, the first and the third steps bring higher accuracy to the second one, representing the limits of automatic learning algorithms. The step that achieves better results is based on the proposed approach in all three cases. In particular, in all cases, due to graph structures used for context modeling, the proposed methodology is able to correct any associations learned by the automatic structural learning algorithms that work exclusively on data and possibly highlight new hidden semantic relationships.

From the results obtained, it is possible to observe that, thanks to the innovative approach used, the number of instances classified correctly has increased, and the cases classified incorrectly have decreased. Moreover, it is possible to infer that the MuG Approach application increases the Bayesian Network's reliability. The tables show that the Precision, Recall, and $F_1$-Score results related to the proposed approach (step 3) in all cases are better than the others. These results were obtained with a few seconds of training time, which is perfectly comparable to the other steps and could represent an advantage in the dynamic training of the system. This encouraging result confirms the MuG approach's benefit in constructing more reliable Bayesian networks.

### 4.2 Smart Home case study

This case study concerns the management of a Smart Home. In recent years, IoT applications and technologies have become increasingly within people's reach, interfacing with their daily activities and occupations, even in the home environment.

**Table 4** London Smart City Case of Study: System performance in Bicycle usage prediction

| | Rental | | | Return | | |
|---|---|---|---|---|---|---|
| | Low (%) | Medium (%) | High (%) | Low (%) | Medium (%) | High (%) |
| Step 1 | | | | | | |
| Precision | 87.07 | 90.08 | 83.64 | 87.82 | 90.52 | 80.66 |
| Recall | 93.84 | 85.49 | 75.87 | 93.41 | 86.16 | 76.02 |
| $F_1$-Score | 90.33 | 87.73 | 79.57 | 90.53 | 88.29 | 78.27 |
| Step 2 | | | | | | |
| Precision | 82.19 | 71.43 | 52.94 | 82.38 | 74.05 | 53.41 |
| Recall | 77.64 | 72.09 | 61.76 | 81.55 | 69.54 | 63.42 |
| $F_1$-Score | 79.85 | 71.76 | 57.01 | 81.96 | 71.72 | 57.99 |
| Step 3 | | | | | | |
| Precision | 92.33 | 92.78 | 86.53 | 93.66 | 92.70 | 86.20 |
| Recall | 95.69 | 91.34 | 81.18 | 95.58 | 92.34 | 81.78 |
| $F_1$-Score | 93.98 | 92.05 | 83.77 | 94.61 | 92.52 | 83.93 |

To be defined as such, smart homes must own specific technological characteristics. The architecture of such systems is based on a communication network between different components and smart devices that allow control, action, and acquiring information for data management. Smart Devices, such as home appliances and various sensors (i.e., temperature and humidity sensor, motion sensor, etc.) able to monitor the environment, and actuators able to perform actions are interconnected in a virtual component that generates a Digital Twin of the System [65]. This environment can integrate data processing and management algorithms, allowing the various elements to interact with each other, and also allows management and control via interface [66] The smart component can address one or more features of the home. The most common are kitchen equipment, lighting, temperature control, and infotainment systems [67]. Another emerging field of application concerns energy and comfort management [68]; this area is growing due to interaction through voice assistants. In this scenario, designing a system that improves the quality of life by optimizing available resources is interesting. Some services allow scheduling tasks and activities and remotely defining routines for smart systems, helping save time in daily home management. Such systems can also help optimize the use of energy resources and help to improve security through more effective control of the environment. This case study focuses on applying the MuG approach in the Smart Home field. Due to the ability to understand contextual information, the proposed method could support the sensor network by defining rules and triggering actions that allow the devices to intervene proactively in the occupations and lives of users and suggest activities tailored to their preferences.

This case study facilitates the evaluation of the MuG approach via an experimental phase, focusing on assessing two critical facets of the suggested method: the system's capacity to aid users in managing home environments by offering recommendations and making independent choices and its ability to display information effectively. To achieve this objective, a prototype featuring both server and client components accessible via the web and applications is utilized. For the server aspect, particularly concerning the inference engines, technologies based on Python and the Django REST framework were implemented. Additionally, the development of the mobile application is realized using the Flutter framework, founded on the Dart language.

This application allows the system to support users in managing a smart home. Its development aims to show users an overview of all the main parameters monitored in the application to choose System actions. In addition, the application design enables assisting users by providing helpful information about household activities, home care, and food preparation.

For experimental purposes, the proposed system takes advantage of Raspberry Pi 4 and Pi 0, connected with various indoor sensors and installed into two rooms of a corporate building. Raspberries enable the obtainment of a smart environment. A similar typology of board (Raspberry Pi 4) allows monitoring weather parameters through a Weather Station as an integration of the collected data. A small solar system for experimental purposes also enables monitoring of electrical absorption. In particular, sensors can collect data related to Air Temperature, Air Humidity, VOC, Luminosity, Weather Description, Voltage, Wind Speed, Wind Direction, Rainfall.

The monitored parameters allow the building of a dataset through which the system can acquire indoor-outdoor environmental parameters such as weather conditions, humidity, indoor temperature, air quality, and the presence of people.

**Table 5** Results related to energy saving and ventilation

| | Yes | No | | Yes | No |
|---|---|---|---|---|---|
| Energy saving | | | Ventilation | | |
| Yes | 136 | 18 | Yes | 189 | 11 |
| No | 22 | 774 | No | 19 | 731 |
| Overall Accuracy: | | **95.79%** | Overall Accuracy: | | **96.84%** |

**Table 6** Smart home experimental results

| | Yes (%) | No (%) | Overall accuracy (%) |
|---|---|---|---|
| Energy saving mode | | | |
| Precision | 89.42 | 98.36 | 95.79 |
| Recall | 88.17 | 98.12 | |
| Ventilation | | | |
| Precision | 95.15 | 98.05 | 96.84 |
| Recall | 91.48 | 97.99 | |

This experimental campaign aims to foresee the need to control the lighting system and ventilate the building when necessary. This experiment was preliminary and aimed at evaluating the system's effectiveness in reduced actions to allow its eventual use in a real environment.

The initial stage of the experimental phase is crucial for determining the system's ability to support users autonomously. This experiment utilizes data gathered by the system over six months, including weather parameters (such as temperature, humidity, and rainfall), indoor environmental factors (like temperature, humidity, and occupancy), conditions of the solar panel system's energy production, and actions implemented to enhance these conditions (such as lighting and room ventilation), all recorded from May to November 2021. The dataset comprises over 4000 entries and is split into a training set (75%) and a test set (25%) to facilitate the MuG approach in learning the structure of the Bayesian network. The evaluation focuses on the system's 'on' and 'off' phases to discern whether the deduced BN structure could autonomously identify necessary actions. Specifically, it predicts parameters related to energy conservation and ventilation. Energy conservation involves dimming lights when rooms are unoccupied or when natural light renders additional illumination redundant. Ventilation, executed via exhaust fans, aims to aerate and purify the space. Despite the limited data volume, experimental findings, as depicted in Tables 5 and 6, indicate the system consistently achieved Overall Accuracy rates above 95%, along with promising Precision and Recall metrics.

In the second stage of the experimental phase, the focus is on the system's capability to effectively present information acquired via an application, as illustrated in Fig. 3, which displays images of the developed App. To assess this, 15 participants were engaged and introduced to a prototype of the application. This App consolidates all the details the system can furnish regarding indoor and outdoor environmental conditions, along with recommended actions, including managing lighting, heating, and ventilation systems.

To evaluate the system, participants are provided with a questionnaire following their interaction. This questionnaire is segmented into five distinct parts. All individuals involved possessed a mobile device and ranged in age from 25 to 58 years. Each statement within a particular section was accompanied by five potential responses: "I totally disagree"—TD, "I disagree"—D, "Undecided"—U, "I agree"—A, "I totally agree"—TA.

Section A: Recommendation.

1. Did the recommendations align with your specific needs and preferences in the given context?
2. How effectively did the system modify its recommendations based on environmental changes?

Section B: User Engagement.

1. How intuitive were your interactions with the application?
2. Was the application's method of communication straightforward and uncomplicated?

Section C: Display.

1. How suitable did you find the presentation of content and services?
2. Was the information provided detailed and easy to understand?

Section D: User Experience.

1. How would you rate the navigability and friendliness of the system's interface?

Section A: Recommendation.

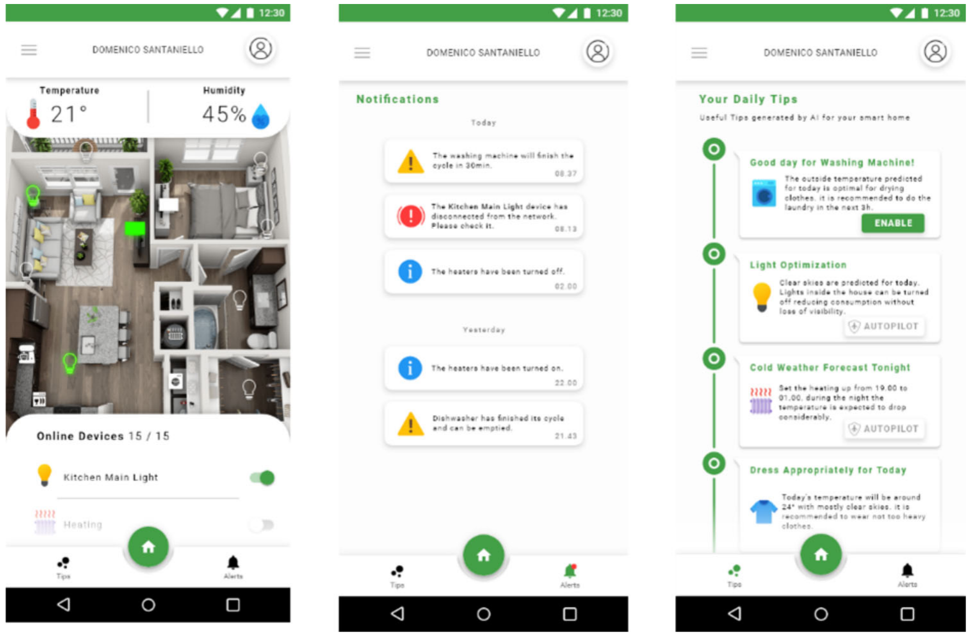


**Fig. 3** Images related to the developed App

1. Did the recommendations align with your specific needs and preferences in the given context?
2. How effectively did the system modify its recommendations based on environmental changes?

Section B: User Engagement.

1. How intuitive were your interactions with the application?
2. Was the application's method of communication straightforward and uncomplicated?

Section C: Display.

1. How suitable did you find the presentation of content and services?
2. Was the information provided detailed and easy to understand?

Section D: User Experience.

1. How would you rate the navigability and friendliness of the system's interface?
2. How satisfactory were the system's response times during usage?

Figure 4 presents the consolidated results by category. Reviewing all feedback indicates a high level of user contentment. The system was evaluated for its reliability and effectiveness, its capability to showcase services, and the App's user-friendliness, yielding valuable insights into the system's interactive prowess.

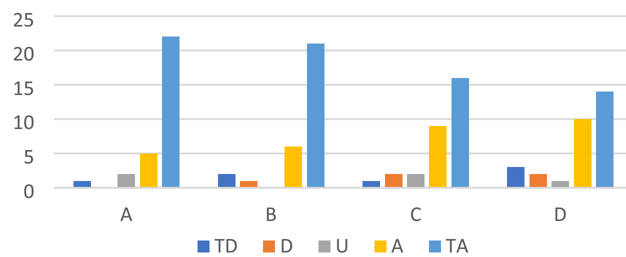


**Fig. 4** Results related to the questionnaire

Even in its initial stages, the system demonstrates a notable index of user acceptance, effectively suggesting the right actions at the appropriate times. This was particularly evident in sections A and B, which underscored its proficiency in delivering accurate information and facilitating user interaction.

This case study aims to assess the effectiveness of the MuG Approach in helping users navigate a Smart Home environment. This approach entailed a two-step experimental phase, which involved developing a prototype and enabling the system to adapt and learn from the gathered data. Despite being in the early stages, the experiments yielded optimistic results, demonstrating the system's capacity to learn and assist users. The application has shown promising potential, indicating its suitability for managing individual and cooperative settings. Considering recent global health concerns, the system could be adapted to ensure cleanliness and safety in communal workspaces with a suitable increase in sensors and proper data handling.

## 5 Conclusions

The increase in technology leads to the creation of increasingly complex and sophisticated scenarios that bring advantages to modern society if managed efficiently. In particular, in literature, several applications exploit context awareness, ontologies, and Bayesian networks (BN) in different application fields without combining all of them. This work aimed to deepen the technological aspects, degree, and probabilistic approaches to propose a methodology for managing complex scenarios based on IoT and able to exploit context-awareness, ontologies, and Bayesian networks. In particular, the main contribution of our approach consists of exploiting contextual data acquired by IoT devices and modeled through the Context Dimension Tree to integrate the semantic analysis of the Ontology through the context. Thus, the next step involves the identification of constraints that can improve the Bayesian network structure learned through data. In particular, Algorithm 1 describes the process of discovering the appropriate structure of the BN, combining available data and constraints. Moreover, the proposed approach is contextualized in an architecture that gives consistency to the developed methodology. Finally, the experimental phase exploits two study cases to validate the theoretical formalization introduced. In particular, the proposed method is tested through accuracy measures and evaluation of users, proving the ability to make decisions in complex scenarios.

The added value of the proposed methodology consists of exploiting graph approaches that are able to include and manage semantic information. Such approaches permit users to perform better while allowing users to understand the relations between entities, providing a greater understanding of data processing and improving performance as the available data increases. In addition, a further advantage of the system lies in interfacing with other systems based on the same formalisms to improve its performance further. However, future developments involve integrating the ability to adapt the Context Dimension Tree and the Ontology construction with data. In particular, the proposed approach depends on the structure of the graphs built a priori and specific for the application domain. In addition, the proposed approach can benefit from integrating other inference engines based on different machine-learning techniques to improve its reliability.

**Author contributions** F.C. and D.S. primarily worked on the conception and design of this study, defining the research objectives and methodological approach. M.C. and A.L. actively participated in data collection, while D.S. and C.V. further contributed to this phase by processing the available data and ensuring the accuracy of the results. A.L., D.S., and C.V. collaborated in drafting the initial version of the article, with M.C. providing a significant contribution at this stage. F.C. played a fundamental role in the manuscript's revision, editing, and refinement, ensuring the content's consistency and quality. F.C. and D.S. provided supervision and guidance throughout the entire research and writing process, offering their academic and professional support to ensure the fulfillment of the research objectives.

**Funding** Open access funding provided by Università degli Studi di Salerno within the CRUI-CARE Agreement.

**Data availability** The data supporting the findings of this study are available upon request from the corresponding author. Some data cannot be shared publicly due to privacy or intellectual property restrictions.

### Declarations

**Conflict of interest** The authors declare no competing interests.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

**Ethical approval** Every method employed in research with human participants was conducted in accordance with the ethical norms of the respective institutional and/or national research committee, as well as the 1964 Helsinki Declaration and its subsequent updates or equivalent ethical guidelines.



## References

1. Zanella A, Bui N, Castellani A, Vangelista L, Zorzi M (2014) Internet of things for smart cities. IEEE Internet Things J 1(1):22–32. https://doi.org/10.1109/JIOT.2014.2306328
2. Khan WZ, Rehman MH, Zangoti HM, Afzal MK, Armi N, Salah K (2020) Industrial internet of things: recent advances, enabling technologies and open challenges. Comput Electr Eng. https://doi.org/10.1016/j.compeleceng.2019.106522
3. Zhang WE, et al (2020) The 10 research topics in the internet of things. In: Proceedings—2020 IEEE 6th International Conference on Collaboration and Internet Computing, CIC 2020. https://doi.org/10.1109/CIC50333.2020.00015
4. Mozzaquatro BA, Jardim-Goncalves R, Agostinho C (2018) Situation awareness in the Internet of Things. In: 2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017—Proceedings. https://doi.org/10.1109/ICE.2017.8279988.
5. Pradeep P, Krishnamoorthy S, Pathinarupothi RK, Vasilakos AV (2021) Leveraging context-awareness for Internet of Things ecosystem: representation, organization, and management of context. Comput Commun 177:33–50. https://doi.org/10.1016/j.comcom.2021.06.004
6. Riedmaier S, Ponn T, Ludwig D, Schick B, Diermeyer F (2020) Survey on scenario-based safety assessment of automated vehicles.

IEEE Access 8:87456–87477. https://doi.org/10.1109/ACCESS.2020.2993730

7. Bai X, Tsai WT, Paul R, Feng K, Yu L (2022) Scenario-based modeling and its applications. In: Proceedings of the Seventh IEEE International Workshop on Object-Oriented Real-Time Dependable Systems. (WORDS 2002), IEEE Computer Society, pp 253–260. https://doi.org/10.1109/WORDS.2002.1000060.
8. Orsi G, Tanca L, Zimeo E (2011) Keyword-based, context-aware selection of natural language query patterns. In: ACM International Conference Proceeding Series. https://doi.org/10.1145/1951365.1951390.
9. Furno A, Zimeo E (2014) Context-aware composition of semantic web services. Mobile Netw Appl. https://doi.org/10.1007/s11036-014-0494-y
10. Sarker IH, Kayes ASM, Watters P (2019) Effectiveness analysis of machine learning classification models for predicting personalized context-aware smartphone usage. J Big Data. https://doi.org/10.1186/s40537-019-0219-y
11. Glowacka J, Krygier J, Amanowicz M (2015) A trust-based situation awareness system for military applications of the internet of things. In: IEEE World Forum on Internet of Things, WF-IoT 2015—Proceedings. https://doi.org/10.1109/WF-IoT.2015.7389103.
12. Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors. https://doi.org/10.1518/001872095779049543
13. Meditskos G, Kontopoulos E, Vrochidis S, Kompatsiaris I (2020) Converness: Ontology-driven conversational awareness and context understanding in multimodal dialogue systems. Expert Syst. https://doi.org/10.1111/exsy.12378
14. Rodriguez Lera FJ, Martín Rico F, Guerrero Higueras AM, Olivera VM (2020) A context-awareness model for activity recognition in robot-assisted scenarios. Expert Syst. https://doi.org/10.1111/exsy.12481
15. Gruber TR (1995) Toward principles for the design of ontologies used for knowledge sharing? Int J Hum Comput Stud 43(5–6):907–928. https://doi.org/10.1006/ijhc.1995.1081
16. Bolchini C, Curino CA, Quintarelli E, Schreiber FA, Tanca L (2009) Context information for knowledge reshaping. Int J Web Eng Technol. https://doi.org/10.1504/IJWET.2009.025015
17. Scutari M, Graafland CE, Gutiérrez JM (2019) Who learns better Bayesian network structures: accuracy and speed of structure learning algorithms. Int J Approx Reason. https://doi.org/10.1016/j.ijar.2019.10.003
18. Abrahao E, Hirakawa AR (2017) Task ontology modeling for technical knowledge representation in agriculture field operations domain. In: Proceedings - 2017 2nd International Conference on Information Systems Engineering, ICISE 2017. https://doi.org/10.1109/ICISE.2017.18
19. Sivarajah U, Kamal MM, Irani Z, Weerakkody V (2017) Critical analysis of Big Data challenges and analytical methods. J Bus Res. https://doi.org/10.1016/j.jbusres.2016.08.001
20. Tole AA (2013) Big data challenges. Database Syst J 4(3):31–40
21. Yin Y, Stecke KE, Li D (2018) The evolution of production systems from Industry 2.0 through Industry 4.0. Int J Prod Res. https://doi.org/10.1080/00207543.2017.1403664
22. Carbone M, Colace F, Lombardi M, Marongiu F, Santaniello D, Valentino C (2021) An adaptive learning path builder based on a context aware recommender system. In: Proceedings - Frontiers in Education Conference, FIE. https://doi.org/10.1109/FIE49875.2021.9637465.
23. Ranganathan A, Campbell RH (2003) An infrastructure for context-awareness based on first order logic. Pers Ubiquitous Comput. https://doi.org/10.1007/s00779-003-0251-x
24. Gu T, Pung HK, Zhang DQ (2005) A service-oriented middleware for building context-aware services. J Netw Comput Appl. https://doi.org/10.1016/j.jnca.2004.06.002
25. Yoon JH, Yang MH, Lim J, Yoon KJ (2015) Bayesian multi-object tracking using motion context from multiple objects. In: Proceedings—2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015. https://doi.org/10.1109/WACV.2015.12.
26. Stenneth L, Wolfson O, Yu PS, Xu B (2011) Transportation mode detection using mobile phones and GIS information. In: GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems. https://doi.org/10.1145/2093973.2093982
27. Kooij JFP, Schneider N, Flohr F, Gavrila DM (2014) Context-based pedestrian path prediction. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-10599-4_40
28. Nguyen T, Lim CP, Nguyen ND, Gordon-Brown L, Nahavandi S (2019) A review of situation awareness assessment approaches in aviation environments. IEEE Syst J. https://doi.org/10.1109/JSYST.2019.2918283
29. Gluck KA, Ball JT, Krusmark MA (2012) Cognitive control in a computational model of the predator pilot. In: Integrated models of cognitive systems. https://doi.org/10.1093/acprof:oso/9780195189193.003.0002
30. Thomas M, Goode N, Grant E, Taylor NZ, Salmon PM (2015) Can we talk about Speed? The effect of verbal protocols on driver speed and perceived workload. Proc Manuf. https://doi.org/10.1016/j.promfg.2015.07.608
31. Liu S, Wanyan X, Zhuang D (2014) Modeling the situation awareness by the analysis of cognitive process. Bio-Med Mater Eng. https://doi.org/10.3233/BME-141044
32. Kolbe N, Zaslavsky A, Kubler S, Robert J, le Traon Y (2017) Enriching a situation awareness framework for IoT with knowledge base and reasoning components. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-57837-8_4
33. Xu G, Cao Y, Ren Y, Li X, Feng Z (2017) Network security situation awareness based on semantic ontology and user-defined rules for internet of things. IEEE Access. https://doi.org/10.1109/ACCESS.2017.2734681
34. Krytska Y, Skarga-Bandurova I, Velykzhanin A (2017) IoT-based situation awareness support system for real-Time emergency management. In: Proceedings of the 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017. https://doi.org/10.1109/IDAACS.2017.8095228
35. Freedman H, Metzger J, Abolhassani N, Tudor A, Tomlinson B, Paul S (2024) A Bayesian approach to constructing probabilistic models from knowledge graphs. Int J Semant Comput 18(01):25–49. https://doi.org/10.1142/S1793351X24410022
36. Ellouze M, Belguith LH (2024) Approach based on Bayesian network and ontology for identifying factors impacting the states of people with psychological problems from data on social media, pp 128–141. https://doi.org/10.1007/978-3-031-49333-1_10.
37. Ben Fraj F, Nouri N (2024) OntoFD: a generic social media fake news ontology, pp 173–185. https://doi.org/10.1007/978-3-031-49333-1_13.
38. Riali I, Fareh M, Ibnaissa MC, Bellil M (2023) A semantic-based approach for hepatitis C virus prediction and diagnosis using a fuzzy ontology and a fuzzy Bayesian network. J Intell Fuzzy Syst 44(2):2381–2395. https://doi.org/10.3233/JIFS-213563

39. Gueddes A, Mahjoub MA (2024) Remote intervention assistance system for a person in difficulty based on probabilistic ontologies. Expert Syst Appl 238:121840. https://doi.org/10.1016/j.eswa.2023.121840
40. Kousis A, Tjortjis C (2021) Data mining algorithms for smart cities: a bibliometric analysis. Algorithms. https://doi.org/10.3390/a14080242
41. Akhter R, Sofi SA (2021) Precision agriculture using IoT data analytics and machine learning. J King Saud Univ Comput Inf Sci. https://doi.org/10.1016/j.jksuci.2021.05.013
42. Ghazal TM et al (2021) IoT for smart cities: Machine learning approaches in smart healthcare—a review. Future Internet. https://doi.org/10.3390/fi13080218
43. Janiesch C, Zschech P, Heinrich K (2021) Machine learning and deep learning. Electron Mark 31(3):685–695. https://doi.org/10.1007/s12525-021-00475-2
44. McGovern A et al (2019) Making the black box more transparent: Understanding the physical implications of machine learning. Bull Am Meteorol Soc. https://doi.org/10.1175/BAMS-D-18-0195.1
45. Rudin C (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. https://doi.org/10.1038/s42256-019-0048-x
46. Colace F, De Santo M, Moscato V, Picariello A, Schreiber FA, Tanca L (eds) Data management in pervasive systems. In: Data-centric systems and applications. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-20062-0.
47. Goel D, Chaudhury S, Ghosh H (2017) An IoT approach for context-Aware smart traffic management using ontology. In: Proceedings—2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017. https://doi.org/10.1145/3106426.3106499
48. Zhong C, Huang X, Müller Arisona S, Schmitt G, Batty M (2014) Inferring building functions from a probabilistic model using public transportation data. Comput Environ Urban Syst. https://doi.org/10.1016/j.compenvurbsys.2014.07.004
49. Weber P, Medina-Oliva G, Simon C, Iung B (2012) Overview on Bayesian networks applications for dependability, risk analysis and maintenance areas. Eng Appl Artif Intell. https://doi.org/10.1016/j.engappai.2010.06.002
50. Colace F, et al (2014) A context-aware framework for cultural heritage applications. In: 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems. IEEE, pp 469–476. https://doi.org/10.1109/SITIS.2014.14.
51. Panigati E, Rauseo A, Schreiber FA, Tanca L (2012) Aspects of pervasive information management: An account of the green move system. In: Proceedings - 15th IEEE International Conference on Computational Science and Engineering, CSE 2012 and 10th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, EUC 2012. https://doi.org/10.1109/ICCSE.2012.93.
52. Chen H, Finin T, Joshi A (2003) An ontology for context-aware pervasive computing environments. Knowl Eng Rev 18(3):197–207
53. Trifonova N, Maxwell D, Pinnegar J, Kenny A, Tucker A (2017) Predicting ecosystem responses to changes in fisheries catch, temperature, and primary productivity with a dynamic Bayesian network model. ICES J Mar Sci. https://doi.org/10.1093/icesjms/fsw231
54. Helsper EM, Gaag LC (2022) Building Bayesian networks through ontologies. In: Proceedings of the 15th European Conference on Artificial Intelligence. IOS Press, pp. 680–684
55. Colace F, De Santo M (2010) Ontology for E-learning: a Bayesian approach. IEEE Trans Educ 53(2):223–233
56. Husáková M, Bureš V (2020) Formal ontologies in information systems development: a systematic review. Information 11(2):66. https://doi.org/10.3390/info11020066
57. Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M (2015) Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun Surv Tutorials. https://doi.org/10.1109/COMST.2015.2444095
58. Abdmeziem MR, Tandjaoui D, Romdhani I (2016) Architecting the internet of things: state of the art. In: Studies in systems, decision and control, vol 36. https://doi.org/10.1007/978-3-319-22168-7_3.
59. Endsley MR, Garland DJ (2000) Situation awareness analysis and measurement. https://doi.org/10.1201/b12461
60. Pan R, Ding Z, Yu Y, Peng Y (2005) A Bayesian network approach to ontology mapping. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/11574620_41
61. Helsper EM, van der Gaag LC (2002) Building Bayesian networks through ontologies. In: ECAI2002, Proceedings of the 15th European Conference on Artificial Intelligence, pp. 680–684.
62. Ikeda M, Seta K, Kakusho O, Mizoguchi R (1998) Task ontology: ontology for building conceptual problem solving models. In: Proceedings of ECAI98.
63. Cooper GF, Herskovits E (1992) A Bayesian method for the induction of probabilistic networks from data. Mach Learn. https://doi.org/10.1023/A:1022649401552
64. Cooper GF, Herskovits E (1992) A Bayesian method for the induction of probabilistic networks from data. Mach Learn 9(4):309–347. https://doi.org/10.1023/A:1022649401552
65. Nazarenko AA, Camarinha-Matos LM (2020) The role of digital twins in collaborative cyber-physical systems. IFIP Adv Inf Commun Technol. https://doi.org/10.1007/978-3-030-45124-0_18
66. Bhat O, Bhat S, Gokhale P (2007) Implementation of IoT in smart homes. Int J Adv Res Comput Commun Eng ISO 3297(December)
67. Marikyan D, Papagiannidis S, Alamanos E (2019) A systematic review of the smart home literature: a user perspective. Technol Forecast Soc Change 138:139–154. https://doi.org/10.1016/j.techfore.2018.08.015
68. Minoli D, Sohraby K, Occhiogrosso B (2017) IoT considerations, requirements, and architectures for smart buildings-energy optimization and next-generation building management systems. IEEE Internet Things J 4(1):269–283. https://doi.org/10.1109/JIOT.2017.2647881