

Article

Building Topic-Driven Virtual IoTs in a Multiple IoTs Scenario

Paolo Lo Giudice ^{1,2}, Antonino Nocera ³ , Domenico Ursino ^{4,*}  and Luca Virgili ⁴

¹ Deloitte, Via Tortona, 25, 20144 Milano MI, Italy

² DIIES, University “Mediterranea” of Reggio Calabria, 80122 Reggio Calabria, Italy

³ DIII, University of Pavia, 27100 Pavia, Italy

⁴ DII, Polytechnic University of Marche, 60131 Ancona, Italy

* Correspondence: d.ursino@univpm.it

Received: 30 May 2019; Accepted: 2 July 2019; Published: 4 July 2019



Abstract: In the last years, several attempts to combine the Internet of Things (IoT) and social networking have been made. In the meantime, things involved in IoT are becoming increasingly sophisticated and intelligent, showing a behavior that tends to look like the one of users in social networks. Therefore, it is not out of place to talk about profiles of things and about information and topics exchanged among them. In such a context, constructing topic-driven virtual communities starting from the real ones operating in a Multi-IoT scenario is an extremely challenging issue. This paper aims at providing some contributions in this setting. First of all, it presents the concept of profile of a thing. Then, it introduces the concept of topic-guided virtual IoT. Finally, it illustrates two approaches (one supervised and one unsupervised) to constructing topic-guided virtual IoTs in a Multi-IoT scenario.

Keywords: Internet of Things; Multiple IoTs; Profile of a Thing; topic-guided virtual IoTs; unsupervised and supervised approaches to virtual IoT construction

1. Introduction

The Internet of Things (hereafter, IoT) is currently considered the new frontier of the Internet. As a matter of fact, a lot of research results, along with the continuous emergence of increasingly challenging issues to address, can be found in the literature [1–7].

One of the most effective ways to represent and handle the IoT scenario leverages social networking paradigm [8]. In this direction, several social network-based approaches to modeling and managing IoTs have been presented in the literature. Three of the most advanced ones are the SIoT (Social Internet of Things) [9–12], the MIE (Multiple IoT Environment) [13] and the MIoT (Multiple IoTs) [14] paradigms. The MIoT paradigm is the last of these proposals; it aims at extending both SIoT and MIE in such a way as to preserve their strengths and avoid their weaknesses [14]. Roughly speaking, a MIoT can be seen as a set of related IoTs, i.e., as a set of related networks of things. Actually, a more precise definition of MIoT requires the introduction of the concept of instance of a thing in an IoT. Specifically, the instance of a thing in an IoT represents a virtual view of that thing in the IoT. The nodes associated with a thing in a MIoT represent the instances of the same thing in the different IoTs of the MIoT. Indeed, a thing can have several instances, one for each IoT which it participates to. The existence of more instances for one thing plays a key role in the MIoT paradigm because it allows the definition of cross relationships among the different IoTs.

We adopted the MIoT paradigm as the reference one in this paper. There are several reasons which justify this choice. Indeed:

- The MIIoT paradigm, like the SIIoT and the MIE ones, introduces the idea that objects can show a social behavior in the environment where they operate. This feature allows several advantages, like the possibility of resource sharing (see [10–12] for a comprehensive idea of these advantages).
- Differently from SIIoT, which introduces a social behavior of objects but still models IIoT as one huge network of objects extended worldwide, MIE, and much more MIIoT, allow the “breakdown” of the whole huge IIoT into multiple networks of smart objects interconnected with each other. This way to proceed is analogous to the evolution of social networking into social internetworking [15]. In particular, MIIoT allows the management of situations in which the same object shows different behaviors in different networks it joined. Furthermore, MIIoT makes an object to act as a bridge between two objects allowing them to communicate even if they belong to different networks and, therefore, are not directly connected with each other.

Another important trend characterizing the current IIoT scenario regards the existence of increasingly sophisticated and intelligent things. These are becoming increasingly smart and social, as well as more and more capable of performing computations and storage on their own. Furthermore, they are increasingly connected to each other through more and more complex and sophisticated frameworks, often based on cloud and edge computing [10–12]. The new smart and social capabilities of things and of the environments handling their interoperability paves the way to a sort of “humanization” of things, i.e., to apply to things concepts and ideas typically considered prerogative of humans. One of them is certainly the presence of a profile of a thing. Indeed, if a thing interacts with other things and exchange data with them, it is possible to determine what are the most common concepts handled by it and, based on them, to construct a corresponding profile. Analogously to the profile of a human, the one of a thing depends on its past behavior and on the profile of the other things with which it interacts. As a consequence, it could be possible to think about both a content-based and a collaborative-filtering approach to handling thing profiles.

Furthermore, starting from the real IIoTs of a MIIoT, it is possible to construct virtual communities of things, based on common interests. Once again, this is an attempt to transfer behaviors typical of humans to things. As a matter of fact, in Social Network Analysis, it is well recognized that, accordingly to the homophily concept [16,17], humans tend to group together in communities sharing the same interests.

In the literature, a lot of efforts have been made to investigate human profiles and virtual communities of people, especially (but not only) in Social Network Analysis [18,19]. Instead, these topics have been little investigated in the Internet of Things.

In this paper, we aim at providing a contribution in this direction. First of all, we introduce the concept of profile of a thing. As the profile of a human, the one of a thing has two components. The former denotes its past behavior and can be used, for instance, to support content-based recommendations. The latter reflects its neighbors, i.e., the other things with which it most frequently comes into contact; it can be exploited, for instance, to support collaborative filtering recommendations.

After this, we introduce the concept of topic-guided virtual IIoTs in a MIIoT and we propose two approaches (one supervised and one unsupervised) to the construction of them in a MIIoT. Differently from the real IIoTs of a MIIoT, which may encompass things with very heterogeneous profiles, topic-guided virtual IIoTs should include all and only those things whose profile refers to specific topics. The supervised approach requires a user to provide a set of keywords of her interest. It aims at constructing a thematic IIoT comprising all the keywords specified by the user. If such an IIoT does not exist, it returns more thematic IIoTs that, in the whole, comprise all the keywords specified by the user. She can choose whether to accept this set of virtual IIoTs or to modify her query. The unsupervised approach tries to partition a MIIoT into a set of virtual IIoTs characterized by the maximum internal cohesion (in terms of topics present in the profiles of the corresponding things) and the minimum external coupling. Virtual IIoTs in a MIIoT provide a logic representation of the objects of a MIIoT, which is not based on real links but on the content exchanged by them. As will be clear in the following, this can favor the effectiveness of information exchange, the construction of communities of

objects (and, possibly, of the corresponding users) sharing the same interests and the suggestions of the objects most adequate to a given exigency.

This paper is organized as follows: in Section 2, we examine related literature. In Section 3, we provide an overview of the MIoT paradigm, because its comprehension is necessary to understand the rest of this paper. In Section 4, we introduce our definition of a thing's profile. In Section 5, we propose our approaches to construct topic-guided virtual IoTs in a MIoT. In Section 6, we present our tests devoted to verify the performance of our approach. Finally, in Section 7, we draw our conclusions and have a look at future developments of our research efforts.

2. Related Literature

Since its introduction some years ago, the term “Internet of Things—IoT” has been associated with a huge variety of concepts, technologies and solutions [5,20–22]. In the latest years, with the advent of new technologies, such as big data and social networking, the very definition of this term is continuously changing. What IoT will become in the future depends on the evolution of these technologies [23] and their interaction with several other ones, such as Information Centric Networks [24–30] and Cloud [3,31,32]. As a matter of fact, the strengths of these last ones are exactly the features necessary to overcome the weaknesses of the current IoT concept [33]. Some examples of this combination can be already found in the literature [6,10,34,35].

The first attempts to apply social networking to the IoT domain can be found in [36–39]. In these papers, the authors propose to use human social network relationships to share services provided by a set of things. An important step forward is performed in [9], where the SIoT paradigm is introduced. Here, the authors propose an approach to creating relationships among things, without requiring the owner intervention. Thanks to this idea, things can autonomously crawl the network to find services and resources of their interest provided by other things. In [40], the same authors clearly highlight what are the main strengths of SIoT. Specifically: (i) the SIoT structure can be dynamically modified to ensure network navigability and to find new things; (ii) scalability is guaranteed, like in human social networks; (iii) a level of trustworthiness among things can be established; (iv) the past social network approaches can be redefined to solve problems typical of the IoT context [41].

One of the major drawbacks of the current IoT scenario is the presence of different technologies and solutions proposed by independent vendors to enable networking among objects. This poses the basis to a subsequent set of issues ranging from concept matching to technical compatibility, if heterogeneous smart-object-network solutions should be involved in the creation of a unique interoperable IoT [42,43]. In this research context, different works partially addressing and solving these problems have been proposed. Specifically, [44] presents a study on how ontologies and semantic data processing can be used to improve interoperability across heterogeneous IoT platforms. The authors consider two use cases, namely *Health Care* and *Transportation and Logistics*, and, for each of them, provide a survey on the main ontologies available to describe and generalize concepts and relations.

In [45], instead, the authors focus their attention on the definition of a new framework for a fully functional mobile ad-hoc social network. In this paper, the term “mobile ad-hoc social network” refers to an IoT made of mobile devices. Of course, communication between this type of objects may happen in such a wide range of modes so that the referring scenario can be considered as a constellation of mobile networks interacting with each other. Concepts from real social networks are borrowed to define user profiles, which are built starting from the objects they own and the social network they belong to. One of the main contributions of this proposal is the definition of a profile-matching strategy based on semantics.

Another contribution in the context of interoperability is the one proposed in [46]. Here, the authors illustrate a novel architecture in which objects interact with each other by leveraging an open source cloud platform. The interaction among smart devices is information-and-service-driven and can be performed in both a centralized and a peer-to-peer mode. In [47], the authors propose *Acrost*, a system capable of retrieving data spread among heterogeneous IoT platforms by leveraging topics

and semantics awareness. To build the metadata, Acrost uses two methodologies: the former exploits regular expression-based approaches, whereas the latter makes use of random fields-based strategies.

In order to address the issues arising when the interoperability among heterogeneous IoTs must be guaranteed, another research line proposes the extension of the results concerning Social Internetworking [15,48] (instead of social networking) to the Internet of Things. By following this strategy, the MIE (Multiple IoT Environment) [13] and the MIoT (Multiple IoTs) [14] paradigms have been proposed. As specified in the Introduction, this last paradigm is the reference one for this paper.

In [49], the authors present an approach to constructing a virtual data mart on which several knowledge discovery tasks can be performed. Clearly the kinds of virtual source constructed in the approach of [49] and in our own are very different. However, the general ideas underlying the two approaches are similar.

In the past, a lot of efforts have been made to investigate human profiles and virtual communities of people, especially (but not only) in Social Network Analysis ([18,19] provide two surveys about these topics). Instead, these issues have been little investigated in the Internet of Things. Specifically, to the best of our knowledge, a comprehensive, high-level abstraction approach to building and managing a profile of a thing, which also takes into account the content it exchanges during its interactions with other things, has not yet been proposed. Instead, some approaches focusing on community detection in IoT have been presented in the very recent literature. Even if they are very different (both in their purposes and in their ways proceed) from the ones of our approach, in the following we present an overview of some of them.

The approach of [50] uses structural information derived from the complex graph of an IoT to extract communities. It exploits a neighbor-based strategy to detect also overlapping communities. The approach of [51] uses data produced by sensors to define a multi-dimensional clustering. The obtained clusters are then mapped to communities of nodes in the original IoT network. To cope with the size of the data graph, the authors leverage state-of-the-art community detection approaches. Finally, they present a new community detection approach that enhances the Girvan-Newman algorithm by using hyperbolic network embedding.

Other works, instead, use knowledge from social networks to refine their results. A similar method is proposed in [52], even though here the strategy works in the opposite way. In fact, first communities are derived from structural information of owners' social networks and, then, objects are seen as resources available inside each community.

Finally, the authors of [53] propose a new community detection algorithm working in a Social Internet of Things (SIoT) scenario. To achieve their objective, they make use of three metrics, namely social similarity, preference similarity and movement similarity. Social similarity is defined according to the concept of cooperativeness and community interest proposed in [54]. Preference similarity takes into account resource and service preferences of the involved things in the network. Finally, movement similarity specifies how much and how long two or more nodes are spatially close.

In [55], the authors propose a community detection approach working on an architecture capable of integrating the Internet of Things and social networking. This approach assumes that two nodes belong to the same community only if they are at most one hop apart and have at least two mutual friends. In order to construct communities, it exploits graph mining techniques.

As a consequence, it does not consider semantics and contents, but leverages only on network structure.

3. The MIoT Paradigm

In this section, we provide an overview of the MIoT paradigm, described in detail in [14], because it is the reference one for our definitions of virtual IoTs in a MIoT.

A MIoT \mathcal{M} consists of a set of m Internets of Things. Formally speaking:

$$\mathcal{M} = \{\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_m\} \quad (1)$$

where \mathcal{I}_k is an IoT.

Let o_j be an object of \mathcal{M} . We assume that, if o_j belongs to \mathcal{I}_k , it has an instance ι_{j_k} , representing it in \mathcal{I}_k . The instance ι_{j_k} consists of a virtual view (or, better, a virtual agent) representing o_j in \mathcal{I}_k . For example, it provides all the other instances of \mathcal{I}_k , and the users who interact with \mathcal{I}_k , with all the necessary information about o_j . Information stored in ι_{j_k} is represented according to the format and the conventions adopted in \mathcal{I}_k .

A MIoT \mathcal{M} can be represented by means of a graph-based notation. In particular, each IoT $\mathcal{I}_k \in \mathcal{M}$ can be modeled by means of a graph $G_k = \langle N_k, A_k \rangle$. In this case:

- N_k is the set of the nodes of G_k ; there is a node n_{j_k} for each instance $\iota_{j_k} \in \mathcal{I}_k$, and vice versa.
- A_k is the set of the arcs of G_k ; there is an arc $a_{jq_k} = (n_{j_k}, n_{q_k})$ if there exists a physical link from n_{j_k} to n_{q_k} .

Finally:

$$\mathcal{M} = \langle N, A \rangle \quad (2)$$

Here:

$$N = \bigcup_{k=1}^m N_k; \quad (3)$$

$$A = A_I \cup A_C, \quad (4)$$

where

$$A_I = \bigcup_{k=1}^m A_k \quad (5)$$

and

$$A_C = \{(n_{j_k}, n_{j_q}) | n_{j_k} \in N_k, n_{j_q} \in N_q, k \neq q\}. \quad (6)$$

A_I is the set of the inner arcs (hereafter, *i-arcs*) of \mathcal{M} ; they relate instances (of different objects) belonging to the same IoT. A_C is the set of the cross arcs (hereafter, *c-arcs*) of \mathcal{M} ; they relate instances of the same object belonging to different IoTs.

The description of the MIoT paradigm presented above highlights that it is possible to model a MIoT at two abstraction levels. The former represents a MIoT as a network and exploits concepts typical of this environment (such as nodes, arcs and so on). The latter models a MIoT as a set of IoTs and makes use of concepts closer to this scenario (such as instances, objects and so forth). Clearly, these two representations are simply two viewpoints of the same environment, and the concepts adopted by them can be used interchangeably. For example, there is a biunivocal correspondence between a node and an instance. However, in the reality, there are some cases in which it is better to use the concept of a node (for example, when we discuss about paths in a network—see below), whereas there are other situations in which it is better the use of the concept of instance (for example, when we discuss about the transactions carried out by two smart objects).

Furthermore, in a MIoT context, a set MD_j of metadata can be associated with an object o_j . Our metadata model refers to the one of the IPSO (Internet Protocol for Smart Objects) Alliance [56]. Specifically MD_j consists of three subsets, namely: (i) MD_j^D , i.e., the set of *descriptive metadata*; (ii) MD_j^T , i.e., the set of *technical metadata*; (iii) MD_j^B , i.e., the set of *behavioral metadata*. All details about these metadata can be found in [14].

4. Definition of a Thing's Profile

In this section, we present our definition of a thing's profile, which represents a first important contribution of this paper. As pointed out in the Introduction, analogously to what happens for human profiles, the profile of a thing can have two components. The former registers its past behavior and is extremely useful for content-based recommendations; for this reason, we call it "content-based component" in the following. The latter registers the main features of those things with which it mostly interacted in the past and can be used for collaborative filtering recommendations; for this reason, we call it "collaborative filtering component" in the following.

Before illustrating in detail the profile of a thing, we must introduce some preliminary concepts. First of all, given two instances ι_{j_k} of o_j and ι_{q_k} of o_q in \mathcal{I}_k , we can define the set $tranSet_{jq_k}$ of the transactions from ι_{j_k} to ι_{q_k} as follows:

$$tranSet_{jq_k} = \{T_{jq_{k_1}}, T_{jq_{k_2}}, \dots, T_{jq_{k_v}}\} \quad (7)$$

A transaction $T_{jq_{k_t}} \in tranSet_{jq_k}$ is represented as:

$$T_{jq_{k_t}} = \langle reason_{jq_{k_t}}, source_{jq_{k_t}}, dest_{jq_{k_t}}, start_{jq_{k_t}}, finish_{jq_{k_t}}, success_{jq_{k_t}}, content_{jq_{k_t}} \rangle \quad (8)$$

Here:

- $reason_{jq_{k_t}}$ denotes the reason why $T_{jq_{k_t}}$ occurred, chosen among a set of predefined values.
- $source_{jq_{k_t}}$ indicates the starting node of the path followed by $T_{jq_{k_t}}$.
- $dest_{jq_{k_t}}$ represents the final node of the path followed by $T_{jq_{k_t}}$.
- $start_{jq_{k_t}}$ denotes the starting timestamp of $T_{jq_{k_t}}$.
- $finish_{jq_{k_t}}$ indicates the ending timestamp of $T_{jq_{k_t}}$.
- $success_{jq_{k_t}}$ denotes whether $T_{jq_{k_t}}$ was successful or not; it is set to true in the affirmative case, to false in the negative one, and to NULL if $T_{jq_{k_t}}$ is still in progress.
- $content_{jq_{k_t}}$ indicates the content "exchanged" from ι_{j_k} to ι_{q_k} during $T_{jq_{k_t}}$. In its turn, $content_{jq_{k_t}}$ presents the following structure:

$$content_{jq_{k_t}} = \langle format_{jq_{k_t}}, fileName_{jq_{k_t}}, size_{jq_{k_t}}, topics_{jq_{k_t}} \rangle \quad (9)$$

Here:

- $format_{jq_{k_t}}$ indicates the format of the content exchanged during $T_{jq_{k_t}}$; the possible values are: "audio", "video", "image" and "text".
- $fileName_{jq_{k_t}}$ denotes the name of the transmitted file.
- $size_{jq_{k_t}}$ indicates the size in bytes of the content.
- $topics_{jq_{k_t}}$ indicates the set of the content topics; it consists of a set of keywords representing the subjects exchanged during $T_{jq_{k_t}}$. It can be formalized as: $topics_{jq_{k_t}} = \{(kw_{jq_{k_t}}^1, nkw_{jq_{k_t}}^1), (kw_{jq_{k_t}}^2, nkw_{jq_{k_t}}^2), \dots, (kw_{jq_{k_t}}^w, nkw_{jq_{k_t}}^w)\}$. In other words, the set of the topics of the t^{th} transaction from ι_{j_k} to ι_{q_k} consists of w pairs; each pair consists of a keyword and the corresponding number of occurrences.

Now, we can define the set $tranSet_{j_k}$ of the transactions performed by ι_{j_k} in \mathcal{I}_k . Specifically, let $Inst_k$ be the set of the instances of \mathcal{I}_k . Then:

$$tranSet_{j_k} = \bigcup_{\iota_{q_k} \in Inst_k, \iota_{q_k} \neq \iota_{j_k}} tranSet_{jq_k} \quad (10)$$

In other words, the set $tranSet_{j_k}$ of the transactions performed by an instance ι_{j_k} is given by the union of the sets of the transactions from ι_{j_k} to all the other instances of \mathcal{I}_k .

After having defined $tranSet_{j_k}$, we must introduce the following operators:

- \uplus : it receives a set $\{entitySet_1, entitySet_2, \dots, entitySet_t\}$ of entity sets and performs their union not eliminating the duplicates but reporting the number of their occurrences. Therefore, this operator returns a set of pairs $\{(entity_1, ne_1), (entity_2, ne_2), \dots, (entity_w, ne_w)\}$ in which the pair $(entity_r, ne_r)$ indicates the r^{th} entity and the number of its occurrences. In counting it, \uplus takes the presence of synonymies and homonymies into account. These properties can be computed (for terms, images, etc.) by applying the classical approaches proposed in the past literature [57,58].
- $avgFileSize$: it receives a set of files and computes their average size.

We are now able to define the profile $\mathcal{P}_{j_{q_k}}$ of the relationship existing between two instances l_{j_k} and l_{q_k} , which performed a set $tranSet_{j_{q_k}} = \{T_{j_{q_{k_1}}}, T_{j_{q_{k_2}}}, \dots, T_{j_{q_{k_v}}}\}$ of transactions. As we will see in the following, this profile plays a crucial role in the definition of the content-based component of a thing's profile and is indirectly used also in the definition of the collaborative filtering component of it. Specifically:

$$\mathcal{P}_{j_{q_k}} = \langle reasonSet_{j_{q_k}}, sourceSet_{j_{q_k}}, destSet_{j_{q_k}}, avgSzAudio_{j_{q_k}}, avgSzVideo_{j_{q_k}}, avgSzImage_{j_{q_k}}, avgSzText_{j_{q_k}}, successFraction_{j_{q_k}}, topicSet_{j_{q_k}} \rangle \quad (11)$$

where:

- $reasonSet_{j_{q_k}} = \uplus_{t=1..v}(reason_{j_{q_{k_t}}})$;
- $sourceSet_{j_{q_k}} = \uplus_{t=1..v}(source_{j_{q_{k_t}}})$;
- $destSet_{j_{q_k}} = \uplus_{t=1..v}(dest_{j_{q_{k_t}}})$;
- $avgSzAudio_{j_{q_k}} = AvgFileSize_{t=1..v}\{fileName_{j_{q_{k_t}}} | format_{j_{q_{k_t}}} = "audio"\}$;
- $avgSzVideo_{j_{q_k}} = AvgFileSize_{t=1..v}\{fileName_{j_{q_{k_t}}} | format_{j_{q_{k_t}}} = "video"\}$;
- $avgSzImage_{j_{q_k}} = AvgFileSize_{t=1..v}\{fileName_{j_{q_{k_t}}} | format_{j_{q_{k_t}}} = "image"\}$;
- $avgSzText_{j_{q_k}} = AvgFileSize_{t=1..v}\{fileName_{j_{q_{k_t}}} | format_{j_{q_{k_t}}} = "text"\}$;
- $successFraction_{j_{q_k}} = \frac{|\{T_{j_{q_{k_t}}} | T_{j_{q_{k_t}}} \in tranSet_{j_{q_k}}, success_{j_{q_{k_t}}} = true\}|}{v}$;
- $topicSet_{j_{q_k}} = \uplus_{t=1..v}(topics_{j_{q_{k_t}}})$.

If we introduce the operator \sqcup , which compactly represents the set of operations for obtaining a profile of a pair of instances $\mathcal{P}_{j_{q_k}}$ starting from the corresponding transactions, we can formalize the previous tasks by means of only one operation as follows:

$$\mathcal{P}_{j_{q_k}} = \sqcup_{t=1..v} T_{j_{q_{k_t}}} \quad (12)$$

Now, let l_{j_k} be the instance of the object o_j in the IoT \mathcal{I}_k . Let $Inst_{j_k}$ be the set of the instances of \mathcal{I}_k with which l_{j_k} performed at least one transaction in the past. In this case, we can define the content-based component of the profile \mathcal{P}_{j_k} of l_{j_k} as:

$$\mathcal{P}_{j_k} = \sqcup_{l_{q_k} \in Inst_{j_k}} \mathcal{P}_{j_{q_k}} \quad (13)$$

Finally, let o_j be an object and let $\{\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_l\}$ be the set of the IoTs which it participates to. Let $ObjInst_j$ be the instances of o_j in the IoTs of the MIoT. We can define the content-based component of the profile \mathcal{P}_j of o_j as:

$$\mathcal{P}_j = \sqcup_{l_{j_k} \in ObjInst_j} \mathcal{P}_{j_k} \quad (14)$$

After having defined the content-based component of an instance and an object, in order to present the corresponding collaborative filtering components, we must introduce the concept of neighborhoods of an instance ι_{j_k} in an IoT \mathcal{I}_k . Specifically, the structural neighborhood $sNbh(\iota_{j_k})$ of ι_{j_k} is defined as:

$$sNbh(\iota_{j_k}) = sNbh^{out}(\iota_{j_k}) \cup sNbh^{in}(\iota_{j_k}) \quad (15)$$

where:

$$sNbh^{out}(\iota_{j_k}) = \{\iota_{q_k} | (n_{j_k}, n_{q_k}) \in A_I\} \quad (16)$$

$$sNbh^{in}(\iota_{j_k}) = \{\iota_{q_k} | (n_{q_k}, n_{j_k}) \in A_I\} \quad (17)$$

Furthermore, we can also define the behavioral neighborhood $bNbh(\iota_{j_k})$ of ι_{j_k} as:

$$bNbh(\iota_{j_k}) = bNbh^{out}(\iota_{j_k}) \cup bNbh^{in}(\iota_{j_k}) \quad (18)$$

where:

$$bNbh^{out}(\iota_{j_k}) = \{\iota_{q_k} | \iota_{q_k} \in sNbh^{out}(\iota_{j_k}), |tranSet_{\iota_{j_k} \iota_{q_k}}| > 0\} \quad (19)$$

$$bNbh^{in}(\iota_{j_k}) = \{\iota_{q_k} | \iota_{q_k} \in sNbh^{in}(\iota_{j_k}), |tranSet_{\iota_{q_k} \iota_{j_k}}| > 0\} \quad (20)$$

In other words, $bNbh(\iota_{j_k})$ consists of those instances directly connected to ι_{j_k} from the structural viewpoint that shared at least one transaction with ι_{j_k} .

We are now able to present the collaborative filtering component \mathcal{P}'_{j_k} of the profile of an instance ι_{j_k} in \mathcal{I}_k . It can be defined as follows:

$$\mathcal{P}'_{j_k} = \bigsqcup_{\iota_{q_k} \in bNbh(\iota_{j_k})} (\mathcal{P}_{q_k} \sqcup \mathcal{P}'_{q_k}) \quad (21)$$

Clearly, this definition is recursive and an accurate computation would require the resolution of a system with a number of equations and variables equal to the number of instances. In real situations, as there could be thousands or millions of instances in a MIoT, the time necessary to solve this system may easily become unacceptable. As a consequence, it appears reasonable to consider an approximate definition of \mathcal{P}_{q_k} that is much simpler to handle. It is formalized as:

$$\mathcal{P}'_{j_k} = \bigsqcup_{\iota_{q_k} \in bNbh(\iota_{j_k})} \mathcal{P}_{q_k} \quad (22)$$

After having introduced the two components of the profile of an instance ι_{j_k} of \mathcal{I}_k , we can combine them for defining the overall profile $\overline{\mathcal{P}}_{j_k}$ of ι_{j_k} . It is defined as the union of the profiles \mathcal{P}_{j_k} and \mathcal{P}'_{j_k} performed by means of the operator \sqcup :

$$\overline{\mathcal{P}}_{j_k} = \mathcal{P}_{j_k} \sqcup \mathcal{P}'_{j_k} \quad (23)$$

Finally, we can define the overall profile of an object o_j as follows:

$$\overline{\mathcal{P}}_j = \bigsqcup_{k=1..l} \overline{\mathcal{P}}_{j_k} \quad (24)$$

5. Topic-Guided Virtual IoTs in a MIoT and Approaches to Constructing Them

In this section, we present a supervised and an unsupervised approach to constructing topic-guided virtual IoTs in a MIoT.

5.1. Supervised Approach

The supervised approach for the construction of topic-guided virtual IoTs in a MIIoT requires the user to specify a query Q consisting of some keywords of her interest. It tries to construct a thematic virtual IoT in such a way that each of its instances contains at least one keyword of Q in the content-based component of its profile. If such a virtual IoT does not exist, our approach returns a minimal set of thematic IoTs that, on the whole, contain, in the content-based component of the profile of their instances, all the keywords specified by the user. In this last case, she can choose whether to accept this set of IoTs or modify her query.

Before describing in detail this approach, we must introduce a new operator J^* that represents a modified Jaccard coefficient, as we will see below.

J^* receives two sets of topics (We recall that, in our context, a topic is a pair (kw, nkw) , where kw is a keyword and nkw is the corresponding number of occurrences.) $topicSet = \{(kw_1, nkw_1), (kw_2, nkw_2), \dots, (kw_p, nkw_p)\}$ and $topicSet' = \{(kw'_1, nkw'_1), (kw'_2, nkw'_2), \dots, (kw'_p, nkw'_p)\}$ and computes the Jaccard coefficient between them. In carrying out this task, it considers the number of occurrences of each keyword and its possible synonyms.

More formally, first it computes the set:

$$commonTS = \{(kw, nkw + nkw') | (kw, nkw) \in topicSet, (kw', nkw') \in topicSet', kw \text{ is identical to or synonymous of } kw'\} \quad (25)$$

Then, it computes the final result as:

$$J^*(topicSet, topicSet') = \frac{\sum_{(kw, nkw) \in commonTS} nkw}{\sum_{(kw, nkw) \in topicSet} nkw + \sum_{(kw', nkw') \in topicSet'} nkw'} \quad (26)$$

After having introduced J^* , we can describe our approach. Specifically:

- It starts when a user specifies a query Q consisting of r keywords:

$$Q = \{kw_1, kw_2, \dots, kw_r\} \quad (27)$$

It searches for all the instances of the MIIoT having at least one topic whose keyword is identical to, or synonymous of, at least one keyword specified in Q . These instances, as a whole, represent the set of candidate instances to be included in the new thematic view. We call this set \mathcal{CI} (Candidate Instances).

- However, the fact that an instance $\iota \in \mathcal{CI}$ has a keyword in common with Q is necessary but not sufficient for it to be chosen. In fact, it is advisable that ι has more keywords in common with Q and, possibly, that the common keywords are among the ones of ι with the highest number of occurrences. This condition can be guaranteed by the usage of the operator J^* .

In particular, our approach first constructs $Q' = \{(kw, 1) | kw \in Q\}$ in such a way as to make the application of J^* on the keywords specified by the user possible. Then, it constructs the set \mathcal{RI} (Real Instances) of those instances of \mathcal{CI} whose topics have a significant similarity with the keywords of Q :

$$\mathcal{RI} = \{\iota \in \mathcal{CI} | J^*(topicSet_\iota, Q') > th_j\} \quad (28)$$

Here, th_j is a suitable tuning threshold.

- Now, our approach can start to construct the thematic view \mathcal{V}_Q corresponding to Q .

- It first creates a node n_i in \mathcal{V}_Q for each instance i of \mathcal{RI} . Let n_{i_1} and n_{i_2} be the nodes corresponding to two instances i_1 and i_2 belonging to \mathcal{RI} .
 - * If an i-arc exists between the nodes corresponding to i_1 and i_2 in the MIoT \mathcal{M} , then an i-arc is also created between the nodes n_{i_1} and n_{i_2} in \mathcal{V}_Q .
 - * Instead, if a c-arc exists between the nodes corresponding to i_1 and i_2 in \mathcal{M} , then n_{i_1} and n_{i_2} are merged in a unique node $n_{i_{12}}$ in \mathcal{V}_Q . This task is motivated by the fact that n_{i_1} and n_{i_2} represent different instances of the same object in different real IoTs, but they represent the same instance in the same virtual IoT; as a consequence, they must be merged and no cross arc can exist between them. The profile $\overline{\mathcal{P}}_{i_{12}}$ of $n_{i_{12}}$ is obtained by applying the operator \sqcup on the profiles $\overline{\mathcal{P}}_1$ of i_1 and $\overline{\mathcal{P}}_2$ of i_2 .
- Finally, our approach adds a disconnected node in \mathcal{V}_Q for each keyword in Q such that there is no MIoT instance having at least one topic whose keyword is identical to, or synonymous of, it (The rationale underlying this step will be clearer in the following.).
- At this point, two cases may occur. In particular:
 - It could happen that \mathcal{V}_Q is connected. In this case, it is returned as the answer to the query Q submitted by the user.
 - If \mathcal{V}_Q is not connected and if the number of its connected components is less than a certain threshold, our approach adds the minimum number of “fictitious” i-arcs necessary to make \mathcal{V}_Q connected.
 - Otherwise, if the number of connected components of \mathcal{V}_Q is higher than a certain threshold, our approach concludes that a unique thematic virtual IoT corresponding to the keywords specified by the user does not exist and returns the thematic views related to the connected components of \mathcal{V}_Q . At this point, the user can decide whether to accept these thematic views or to modify the query in such a way as to construct a unique thematic view by re-applying all the above mentioned steps starting from the new query.

5.2. Unsupervised Approach

The unsupervised approach begins with the construction of a support network \mathcal{N} starting from the MIoT \mathcal{M} . In particular:

- For each node n_{i_k} of \mathcal{M} , a node \overline{n}_{i_k} is added in \mathcal{N} .
- For each i-arc $(n_{i_{j_k}}, n_{i_{q_k}})$ in \mathcal{M} , an (unoriented) arc $(\overline{n}_{i_{j_k}}, \overline{n}_{i_{q_k}})$ is added in \mathcal{N} . The arcs of \mathcal{N} are weighted. The weight of the arc $(\overline{n}_{i_{j_k}}, \overline{n}_{i_{q_k}})$ is obtained by applying the operator J^* on the topic sets $topicSet_{j_k}$ and $topicSet_{q_k}$ of i_{j_k} and i_{q_k} , respectively. Therefore, the weight of an arc in \mathcal{N} belongs to the real interval $[0, 1]$; the higher this weight the higher the semantic similarity between the topics of the profiles $\overline{\mathcal{P}}_{j_k}$ and $\overline{\mathcal{P}}_{q_k}$ of i_{j_k} and i_{q_k} , respectively.
- For each c-arc in \mathcal{M} , which relates two instances $n_{i_{j_k}}$ and $n_{i_{j_q}}$ of the same object o_j in two different IoTs \mathcal{I}_k and \mathcal{I}_q , the two nodes $\overline{n}_{i_{j_k}}$ and $\overline{n}_{i_{j_q}}$ in \mathcal{N} , corresponding to the nodes $n_{i_{j_k}}$ and $n_{i_{j_q}}$ in \mathcal{M} , are merged into a unique node \overline{n}_{i_j} . This node inherits all the arcs of $\overline{n}_{i_{j_k}}$ and $\overline{n}_{i_{j_q}}$.

At the end of these steps, it could happen that two or more arcs relate the same nodes \overline{n} and \overline{n}' in \mathcal{N} . In this case, all these arcs must be merged into a single arc. Clearly, it is necessary to determine the weight of this arc. Here, it appears reasonable that it must be higher than or equal to the maximum weight of the merged arcs. To reach this objective, our approach operates as follows. Let $\{(\overline{n}, \overline{n}', \overline{w}^1), (\overline{n}, \overline{n}', \overline{w}^2), \dots, (\overline{n}, \overline{n}', \overline{w}^s)\}$ be the arcs to merge, ordered by decreasing weight. The new arc $(\overline{n}, \overline{n}', \overline{w})$ will have a weight equal to:

$$\overline{w} = \min \left(1, \overline{w}^1 + \alpha \sum_{k=2..s} \overline{w}^k \right) \quad (29)$$

In other words, in the computation of \bar{w} , the arcs with the maximum weight will contribute with all their weight. All the other arcs will contribute to a lesser extent, with a fraction of their weight. This last is determined by means of the coefficient α .

Once the construction of \mathcal{N} has been completed, the thematic views are derived by applying on \mathcal{N} a graph clustering algorithm among the ones already existing in the literature (see [59] for a survey on them).

5.3. Discussion

An important issue about the supervised and the unsupervised approaches to address regards their scalability or, better, the possibility to use them in MIoTs comprising thousands or even millions of nodes.

With regard to this issue, first of all we observe that both approaches aim at deriving virtual IoTs which are, then, exploited by users to perform their desired tasks (such as querying). As a consequence, we can distinguish two moments in the life of a MIoT, namely: (i) the construction of virtual IoTs, which can be performed *offline*, and (ii) their usage, which is generally carried out *online*.

The first moment is computationally expensive because it involves several network operations in the supervised approach and a clustering activity in the unsupervised one. Clustering's computational cost is intrinsically exponential even if all the corresponding methods adopted in the reality are heuristic and most of them have a linear or a quadratic computational complexity. In any case, as pointed above, this task is performed offline and rarely because it is necessary only when many changes have been made in the MIoT.

The second moment is certainly less expensive; its cost depends on the size of the involved clusters; in fact, each user activity generally involves one or a few clusters. Concerning this aspect, it is important to verify: (i) if clustering is possible in presence of huge MIoTs, and (ii) how the size of clusters increases against the growth of the MIoT. As for the first point, we observe that, in the past, several algorithms have been specifically conceived to cluster a huge amount of elements [60]. Concerning the second point, instead, first we observe that the size of clusters can be determined by suitably tuning the parameters of the selected clustering algorithm. However, it could be interesting to verify how much the size of clusters increases if we maintain constant all the clustering algorithm parameters and the MIoT size increases. We decided to perform this experiment. It is described in detail in Section 6.6. Here, we evidence the obtained results, i.e., that when the MIoT size highly increases, the cluster size slightly grows, whereas the number of clusters increases very much. This is a positive result for our purposes because the parameter to monitor for investigating the performance obtained during the second moment is just cluster size.

Another important issue to investigate regards the possible existence of a unique framework handling all the objects of the MIoT and, therefore, in principle, thousands or millions of objects. With regard to this aspect, we evidence that, in the past, several attempts have been successfully performed in this direction (think, for instance, of the SIoT framework proposed in [9,40]). Clearly, we understand that, in the future, the number of objects possibly belonging to a MIoT is enormously higher than the number of objects available in the past IoT frameworks. However, we point out that: (i) our approach needs to store only the metadata of the involved objects, and these are small; (ii) the real objects can operate in a distributed environment thanks to the new available technologies, such as cloud, edge and fog computing, which can ease the organization and the management of distributed contexts.

6. Experiments

In this section, we present the experimental campaign that we carried out to evaluate the performance of our approach from several viewpoints. Specifically, we describe our dataset in a subsection, whereas, in the next ones, we illustrate our tests, along with the underlying motivations and the obtained results.

6.1. Adopted Dataset

To perform our experiments, we had the necessity to create several MIoT with different sizes, ranging from hundreds to thousands of nodes. Since, currently, real MIoT with the size and the variety handled by our model do not exist yet, we had to realize a MIoT simulator, i.e., a tool that, starting from real data, is capable of simulating MIoT with certain characteristics specified by the user.

The MIoT created by our simulator follow the model described in Section 3. In order to perform its task, our simulator carries out the following steps: (i) creation of objects; (ii) creation of object instances; (iii) creation of instance connections; (iv) creation of instance profiles.

Our MIoT simulator is also provided with a suitable interface allowing a user to “personalize” the MIoT to construct by specifying the desired values for several parameters, such as the number of nodes, the maximum number of instances of an object, and so forth.

To make “concrete” and “plausible” the created MIoT, our simulator leverages a real dataset. It regards the taxi routes in the city of Porto from 1 July 2013 to 30 June 2014. It can be found at the address <http://www.geolink.pt/ecmlpkdd2015-challenge/dataset.html>. Each route contains several Points of Interests corresponding to the GPS coordinates of the vehicle.

We partitioned the city of Porto in six areas and associated a real IoT with each of them. Our simulator associates an object with a given route recorded in the dataset and an object instance for each partition of a route belonging to an area. It creates a MIoT node for each instance and a c-arc for each pair of instances belonging to the same route. Furthermore, it creates an i-arc between two nodes of the same IoT if the length of the time interval between the corresponding routes is less than a certain threshold th_t . The weight of the i-arc indicates the length of this time interval. The value of th_t can be specified through the constructor interface. Clearly, the higher th_t the more connected the constructed MIoT.

As far as instance profiles are concerned, since there are no thing profiles available (indeed, the concept of thing profile is one of the main novelties introduced in this paper), we had to simulate them. However, we aimed to make them as real as possible. In order to increase the likelihood of constructed MIoT, we performed a sentiment analysis task for each of the six areas in which we partitioned the city of Porto and for each day which the dataset refers to. For this purpose, we leveraged IBM Watson on the social media and blogs it uses as default. Having this data at disposal, our simulator assigns to each instance the most common topics (along with the corresponding occurrences) discussed in that area in the day on which the corresponding route took place. The constructed MIoT are returned in a format that can be directly processed by the cypher-shell of Neo4J (see below).

Some features of the constructed MIoT are reported in Table 1. The interested reader can find the MIoT adopted in the experiments described in this section at the address <http://daisy.dii.univpm.it/miot/datasets/virtualIoT>.

Table 1. Main features of the constructed MIoT.

MIoT (Size)	Number of Arcs	Mean In-Degree	Mean Out-Degree	Number of i-arcs	Number of c-arcs
\mathcal{M}_1 (176)	1176	6.29	6.61	980	126
\mathcal{M}_2 (301)	2050	7.76	7.74	1709	341
\mathcal{M}_3 (485)	3756	8.80	8.54	3130	626
\mathcal{M}_4 (778)	5866	8.89	9.11	4895	971
\mathcal{M}_5 (946)	7624	8.64	8.84	6422	1202
\mathcal{M}_6 (1256)	9860	7.87	7.98	7917	1943
\mathcal{M}_7 (1725)	12,263	7.94	8.18	9964	2299
\mathcal{M}_8 (2028)	15,568	8.22	8.38	12,857	2711
\mathcal{M}_9 (3544)	26,428	8.36	8.42	22,718	3710
\mathcal{M}_{10} (5024)	38,642	8.44	8.54	33,724	4918

We carried out all the tests presented in this section on a server equipped with an Intel I7 Quad Core 7700 HQ processor and 16 GB of RAM with Ubuntu 16.04 operating system.

To implement our approaches we adopted:

- Python, powered with the NetworkX library, as programming language;
- Neo4J (Version 3.4.5) as underlying DBMS; we also exploited some plugins of Neo4J to perform community detection and to compute clustering coefficients.

6.2. Cohesion of the Obtained Topic-Guided Virtual IoTs

Our first test started from the idea that if our approach aims at extracting virtual thematic IoTs, they should present both a structural and a semantic cohesion higher than the corresponding ones characterizing the original IoTs of the MIoT. This experiment was devoted to evaluate if this assumption is verified. We considered two well known structural cohesion parameters used in network analysis literature, namely *clustering coefficient* and *density* [61]. Both of them range in the real interval $[0, 1]$; the higher their value the higher the corresponding network cohesion. In the following, first we test the supervised approach and, then, we consider the unsupervised one.

6.2.1. Supervised Approach

In this test, we run our supervised approach on ten MIOts, $\mathcal{M}_1, \dots, \mathcal{M}_{10}$, consisting of 176, 301, 485, 778, 946, 1256, 1725, 2028, 3544 and 5024 nodes. Clearly, the number of IoTs for each MIOt was equal to six, one for each area of the city of Porto that we have defined. For each MIOt, we submitted a set of 10 queries consisting of 1 (resp., 2, 4, 6, 8 and 10) word(s).

Each query returned a virtual thematic IoT for which we computed the corresponding clustering coefficient and density. Finally, we averaged the obtained results for each MIOt and for each set of queries, and we compared them with the average clustering coefficient and the average density of the corresponding real IoTs. The obtained results are reported in Tables 2 and 3.

Table 2. Values of the clustering coefficient for real and virtual IoTs against the size of MIOts and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Avg. Clustering Coeff. (Real IoTs)	Avg. Clustering Coeff. (Virtual IoTs)					
		$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.230	0.318	0.368	0.389	0.394	0.401	0.408
\mathcal{M}_2 (301)	0.272	0.343	0.388	0.419	0.424	0.434	0.446
\mathcal{M}_3 (485)	0.293	0.396	0.437	0.477	0.482	0.488	0.497
\mathcal{M}_4 (778)	0.353	0.447	0.478	0.503	0.508	0.511	0.517
\mathcal{M}_5 (946)	0.371	0.452	0.492	0.512	0.522	0.524	0.526
\mathcal{M}_6 (1256)	0.385	0.486	0.511	0.529	0.530	0.532	0.535
\mathcal{M}_7 (1725)	0.386	0.501	0.524	0.536	0.537	0.538	0.539
\mathcal{M}_8 (2028)	0.388	0.519	0.536	0.541	0.541	0.542	0.543
\mathcal{M}_9 (3544)	0.392	0.522	0.540	0.544	0.544	0.545	0.546
\mathcal{M}_{10} (5024)	0.395	0.534	0.546	0.546	0.546	0.547	0.548

Table 3. Values of the density for real and virtual IoTs against the size of MIOts and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average Density (Real IoTs)	Average Density (Virtual IoTs)					
		$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.348	0.260	0.264	0.280	0.289	0.296	0.301
\mathcal{M}_2 (301)	0.262	0.292	0.303	0.309	0.315	0.320	0.324
\mathcal{M}_3 (485)	0.274	0.390	0.395	0.400	0.402	0.405	0.408
\mathcal{M}_4 (778)	0.269	0.476	0.483	0.490	0.501	0.509	0.514
\mathcal{M}_5 (946)	0.276	0.492	0.509	0.521	0.536	0.534	0.556
\mathcal{M}_6 (1256)	0.284	0.547	0.556	0.567	0.572	0.576	0.581
\mathcal{M}_7 (1725)	0.278	0.582	0.582	0.594	0.598	0.598	0.601
\mathcal{M}_8 (2028)	0.273	0.609	0.610	0.620	0.626	0.630	0.639
\mathcal{M}_9 (3544)	0.269	0.626	0.628	0.630	0.634	0.636	0.637
\mathcal{M}_{10} (5024)	0.262	0.636	0.636	0.638	0.638	0.640	0.642

From the analysis of these tables, we can observe that, in almost all circumstances, the values of both clustering coefficient and density are higher or much higher for the virtual thematic IoTs than for the real ones. This is clearly a confirmation of the goodness of our supervised approach, which returns topic-guided IoTs more cohesive than the original ones. We also observe that when $|Q|$ increases, the values of both clustering coefficient and density increases. This can be explained by observing that, in processing Q , our approach takes the portions of networks containing at least one keyword of Q . When $|Q|$ increases, the portion of networks selected by our approach increases too, and the probability of selecting a very high number of edges (i.e., a number so high to lead to an increase of clustering coefficient and density) increases as well.

6.2.2. Unsupervised Approach

In this test, we run our unsupervised approach, powered with the Louvain graph clustering algorithm [62] as underlying engine, on the same MIoT described in Section 6.2.1. For each MIoT, we computed the average clustering coefficient and the average density of real and virtual IoTs. The obtained results are reported in Table 4.

Table 4. Values of both clustering coefficient and density of real and virtual IoTs against the size of MIoT (unsupervised approach).

MIoT (Size)	Average Clustering Coefficient		Average Density	
	Real IoTs	Virtual IoTs	Real IoTs	Virtual IoTs
\mathcal{M}_1 (176)	0.230	0.473	0.348	0.315
\mathcal{M}_2 (301)	0.272	0.499	0.262	0.350
\mathcal{M}_3 (485)	0.293	0.500	0.274	0.375
\mathcal{M}_4 (778)	0.353	0.511	0.269	0.318
\mathcal{M}_5 (946)	0.372	0.509	0.276	0.316
\mathcal{M}_6 (1256)	0.385	0.506	0.284	0.314
\mathcal{M}_7 (1725)	0.386	0.522	0.280	0.328
\mathcal{M}_8 (2028)	0.388	0.535	0.273	0.360
\mathcal{M}_9 (3544)	0.394	0.547	0.271	0.364
\mathcal{M}_{10} (5024)	0.398	0.562	0.269	0.368

From the analysis of this table we can observe that, in this case, analogously to what happened for the supervised approach, the cohesion level of the virtual IoTs is higher or much higher than the corresponding ones of the real original IoTs. Interestingly, both clustering coefficient and density values obtained by the unsupervised approach are generally higher than those returned by the supervised one, at least when the MIoT size is small. Instead, when the MIoT size is large, they become lower than the ones of the supervised approach. Actually, the increase of both clustering coefficient and density when the MIoT size increases is significant for the supervised approach, whereas it is more limited for the unsupervised one.

6.3. Average Fraction of Merged C-Nodes and Analysis of Node Distribution in Virtual IoTs

Another quality parameter for virtual IoTs returned by our approach regards the average number of merged c-nodes present in each of them. Indeed, the presence of merged c-nodes in an IoT is an indicator of the fact that this IoT is capable of connecting concepts coming from different real IoTs, and, therefore, from concepts whose relationships would have been uncaptured otherwise, or, in other words, that the knowledge it is presenting is new and did not exist previously. Clearly, the higher the fraction of merged c-nodes and the higher the fraction of different original IoTs they belong to, the higher the connecting capability of virtual IoTs.

Also for this experiment, we considered the ten MIoT described in Section 6.2 and performed the same tasks illustrated therein for both the supervised and the unsupervised approaches. The obtained results are reported in Tables 5–7.

Table 5. Average fraction of merged c-nodes against the size of MIoTs and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average Fraction of Merged C-Nodes					
	$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.304	0.455	0.517	0.532	0.554	0.572
\mathcal{M}_2 (301)	0.380	0.515	0.608	0.627	0.652	0.679
\mathcal{M}_3 (485)	0.539	0.661	0.782	0.798	0.813	0.823
\mathcal{M}_4 (778)	0.690	0.786	0.860	0.874	0.883	0.892
\mathcal{M}_5 (946)	0.724	0.812	0.884	0.898	0.916	0.924
\mathcal{M}_6 (1256)	0.808	0.883	0.939	0.943	0.946	0.948
\mathcal{M}_7 (1725)	0.862	0.908	0.952	0.961	0.961	0.963
\mathcal{M}_8 (2028)	0.908	0.959	0.974	0.975	0.976	0.977
\mathcal{M}_9 (3544)	0.928	0.963	0.976	0.977	0.977	0.978
\mathcal{M}_{10} (5024)	0.936	0.968	0.978	0.979	0.980	0.981

Table 6. Average fraction of real IoTs involved in a virtual IoT against the size of MIoTs and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average Fraction of Involved Real IoTs					
	$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.373	0.467	0.488	0.476	0.452	0.448
\mathcal{M}_2 (301)	0.365	0.469	0.525	0.501	0.488	0.480
\mathcal{M}_3 (485)	0.482	0.477	0.448	0.442	0.435	0.432
\mathcal{M}_4 (778)	0.457	0.432	0.418	0.415	0.413	0.411
\mathcal{M}_5 (946)	0.455	0.482	0.624	0.628	0.647	0.644
\mathcal{M}_6 (1256)	0.453	0.514	0.805	0.864	0.917	0.924
\mathcal{M}_7 (1725)	0.482	0.577	0.815	0.872	0.917	0.924
\mathcal{M}_8 (2028)	0.514	0.672	0.833	0.898	0.917	0.924
\mathcal{M}_9 (3544)	0.584	0.704	0.844	0.905	0.924	0.926
\mathcal{M}_{10} (5024)	0.624	0.727	0.888	0.911	0.928	0.934

Table 7. Average fraction of merged c-nodes and average fraction of real IoTs involved in a virtual IoT against the size of MIoTs (unsupervised approach).

MIoT (Size)	Average Fraction of Merged C-Nodes	Average Fraction of Involved Real IoTs
\mathcal{M}_1 (176)	0.227	0.361
\mathcal{M}_2 (301)	0.306	0.353
\mathcal{M}_3 (485)	0.309	0.357
\mathcal{M}_4 (778)	0.342	0.356
\mathcal{M}_5 (946)	0.334	0.359
\mathcal{M}_6 (1256)	0.326	0.361
\mathcal{M}_7 (778)	0.332	0.360
\mathcal{M}_8 (2028)	0.335	0.358
\mathcal{M}_9 (3544)	0.341	0.371
\mathcal{M}_{10} (5024)	0.344	0.378

From the analysis of these tables, we observe that both the supervised and the unsupervised approaches return satisfying results. As for the supervised approach, we can observe that the fraction of merged c-nodes increases when the size of MIoT increases. Furthermore, we can also observe a slight increase of this fraction when $|Q|$ increases. The same trends can be observed for the average fraction of involved real IoTs, even if, for this parameter, its increase against the increase of $|Q|$ is more pronounced. As for the unsupervised approach, we can observe that the average fraction of merged nodes is always very high, independently of the MIoT size. By contrast, in this case, the fraction of involved real IoTs is quite high even if lower than the ones generally observed for the supervised approach. Furthermore, its value does not significantly change when the MIoT size increases.

In order to deepen this investigation, for each virtual IoT, we compared the distribution of its nodes against the real IoTs they belong to. Indeed, if almost all the nodes of a virtual IoT derive from only one real IoT, the information contribution provided by the virtual IoT would be very small because it would be analogous to the one provided by the corresponding real IoT. By contrast, if the nodes of a virtual IoT homogeneously derive from several real IoTs, then the knowledge it provides is really new, and this knowledge would be uncaptured and lost if the new IoT had not been extracted. On the basis of this reasoning, we evaluated the heterogeneity of the provenance of the various nodes of each virtual IoT (see below). For this purpose, we adapted the Herfindahl Index [63] to our context. This index is very used in several research fields of Economics from several decades; for instance, it is exploited to evaluate the concentration degree in an industry.

In order to adapt the Herfindahl Index to our scenario, consider a MIoT \mathcal{M} consisting of s real IoTs ($\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_s$). Consider, also, a virtual IoT \mathcal{V}_j derived by either the supervised or the unsupervised approach. Let n_j be the number of nodes of \mathcal{V}_j and let $\frac{n_{jk}}{n_j}$, $1 \leq k \leq s$, be the fraction of the nodes of \mathcal{V}_j belonging to \mathcal{R}_k (i.e., the k^{th} real IoT of the MIoT). The Herfindahl Index H_j of \mathcal{V}_j is defined as $\sum_{k=1}^s \left(\frac{n_{jk}}{n_j}\right)^2$. H_j ranges in the real interval $\left[\frac{1}{s}, 1\right]$; the higher its value, the higher the concentration degree of the nodes of \mathcal{R}_k in \mathcal{V}_j . Clearly, as previously pointed out, one property desired for our approach is the ability to construct virtual IoTs connecting nodes that belong to different real IoTs in such a way as to extract knowledge that would be lost otherwise. If we report this property to the Herfindahl Index, this implies to obtain a value of this index as lower as possible (Consider that, since we have six real IoTs in our MIoTs, the minimum value of the Herfindahl Index is $\frac{1}{6} = 0.167$).

We computed the average Herfindahl Index of the thematic IoTs returned by both the supervised and the unsupervised approaches by considering the ten MIoTs described in Section 6.2 and performing the same tasks illustrated therein. The obtained results are reported in Tables 8 and 9.

Table 8. Average Herfindahl Index of virtual IoTs against the size of MIoTs and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average Herfindahl Index					
	$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.207	0.186	0.177	0.175	0.173	0.172
\mathcal{M}_2 (301)	0.204	0.183	0.174	0.173	0.172	0.171
\mathcal{M}_3 (485)	0.178	0.173	0.170	0.170	0.169	0.168
\mathcal{M}_4 (778)	0.172	0.172	0.170	0.170	0.169	0.168
\mathcal{M}_5 (946)	0.172	0.170	0.169	0.169	0.169	0.168
\mathcal{M}_6 (1256)	0.173	0.168	0.167	0.169	0.168	0.167
\mathcal{M}_7 (1725)	0.170	0.168	0.167	0.169	0.168	0.167
\mathcal{M}_8 (2028)	0.168	0.167	0.167	0.167	0.167	0.167
\mathcal{M}_9 (3544)	0.168	0.167	0.167	0.167	0.167	0.167
\mathcal{M}_{10} (5024)	0.167	0.167	0.167	0.167	0.167	0.167

Table 9. Average Herfindahl Index of virtual IoTs against the size of MIoTs (unsupervised approach).

MIoT (Size)	Average Herfindahl Index
\mathcal{M}_1 (176)	0.658
\mathcal{M}_2 (301)	0.543
\mathcal{M}_3 (485)	0.658
\mathcal{M}_4 (778)	0.636
\mathcal{M}_5 (946)	0.654
\mathcal{M}_6 (1256)	0.694
\mathcal{M}_7 (1725)	0.656
\mathcal{M}_8 (2028)	0.635
\mathcal{M}_9 (3544)	0.664
\mathcal{M}_{10} (5024)	0.686

These tables evidence that also the analysis based on object distribution and Herfindahl Index returns very satisfying results that confirm and strengthen those obtained by examining the average fraction of merged nodes involved in a virtual IoT. Interestingly, as for this parameter, we observe that the supervised approach returns excellent results, very close to the best ones. By contrast, the unsupervised approach returns good results, even if those returned by the supervised approach are better.

6.4. Computation Time

In this experiment, we aimed at evaluating the variation of the computation time of both the supervised and the unsupervised approaches against the variation of the size of the involved MIoT. Furthermore, as for the supervised approach, we also evaluated the variation of the computation time against the variation of the size of queries.

To perform this task, we considered the ten MIoTs described in Section 6.2 and carried out the same tasks illustrated therein. Finally, we measured the corresponding average computation times. The obtained results are reported in Figures 1–3.

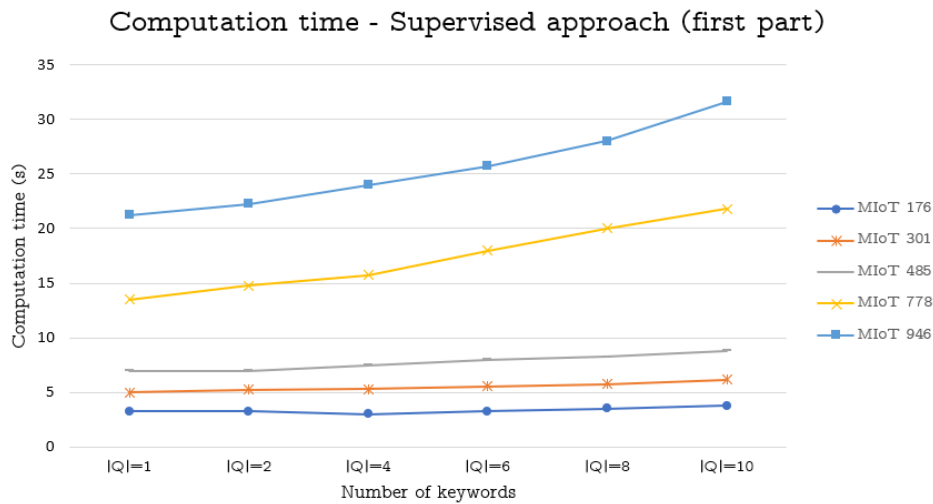


Figure 1. Computation time (in seconds) against the size of MIoTs and queries used to generate the virtual IoTs (supervised approach)—first part.

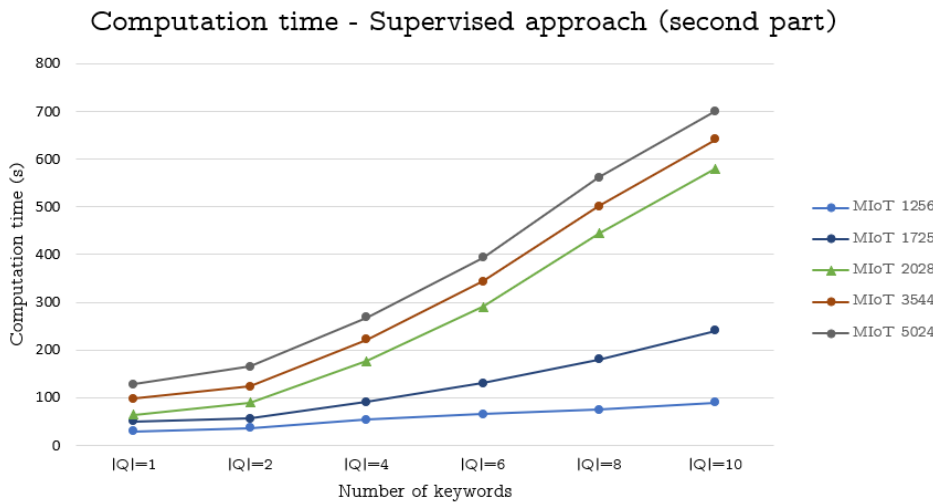


Figure 2. Computation time (in seconds) against the size of MIoTs and queries used to generate the virtual IoTs (supervised approach)—second part

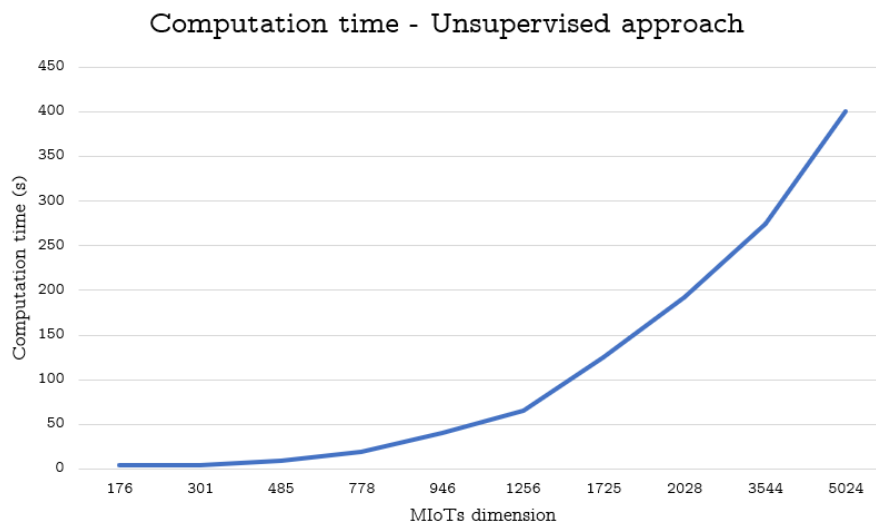


Figure 3. Computation time (in seconds) against the size of MIoTs (unsupervised approach)

From the analysis of these figures, we can observe that our approaches obtain satisfying results. Specifically, as for the supervised approach, the computation time is always very low for MIoTs having at most 1256 nodes. Instead, for MIoTs with more than 2028 nodes, the computation time is low for $|Q| = 1$ or $|Q| = 2$. Then, it increases, even if it remains acceptable for $|Q| = 4$ and $|Q| = 6$, whereas it becomes excessive for $|Q| = 8$ and $|Q| = 10$. However, with regard to this fact, we must point out that queries consisting of 8 or 10 keywords are very uncommon (It is worth pointing out that the topics considered by our approach for constructing a thing's profile are extremely generic and heterogeneous. As a consequence, in our scenario, a query with 8 or 10 keywords would encompass a great number of different topics and, as such, it would not be generally able to capture a clear and specific desire of a user.).

As for the unsupervised approach, its computation time is still acceptable also for 2028 nodes. It starts to become excessive with MIoTs consisting of at least 10,000 nodes.

6.5. Our Approaches' Capability of Improving the Efficiency of Information Dissemination

This experiment was devoted to measure the efficiency of both supervised and unsupervised approaches. The rationale underlying this experiment is that if some information must be transferred from a source object o_s to a target one o_t , the number of objects to be contacted for this task should be minimized. At the same time, if an object is involved in an information dissemination task, it would be desirable that the information it is transmitting is also useful for it (which, in our case, means that it is in line with the interests of its profile).

In order to perform this experiment, we randomly selected some pairs of (source, target) nodes from our MIoT. Let (n_s, n_t) be one of these pairs. We verified if there existed at least one virtual IoT comprising both n_s and n_t (This is always true for the unsupervised approach, whereas it could not happen for the supervised one.). In the negative case, we discarded that pair. Let \mathcal{V} be a virtual IoT comprising both n_s and n_t .

After this, we computed the number $num_{st}^{\mathcal{V}}$ (resp., $\widehat{num}_{st}^{\mathcal{V}}$) of MIoT nodes involved in the dissemination of information in presence (resp., absence) of the virtual IoT \mathcal{V} . Specifically, we computed $num_{st}^{\mathcal{V}}$ by performing the information dissemination task only through its nodes; instead, we obtained $\widehat{num}_{st}^{\mathcal{V}}$ by performing the same task on the whole MIoT. Finally, we computed: $f_{st} = \frac{num_{st}^{\mathcal{V}}}{\widehat{num}_{st}^{\mathcal{V}}}$. Clearly, the lower f_{st} , the higher the contribution of the virtual IoTs in reducing the number of nodes necessary for the information dissemination task and, consequently, the higher the contribution that our virtual IoT detection approach can provide to information dissemination.

We computed the average values of f_{st} by operating on the ten MIoT introduced in Section 6.2 and by performing the same tasks described therein for both the supervised and the unsupervised approaches. The obtained results are reported in Tables 10 and 11.

Table 10. Average values of f_{st} against the size of MIoT and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average f_{st}					
	$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	0.144	0.220	0.290	0.304	0.336	0.347
\mathcal{M}_2 (301)	0.126	0.170	0.177	0.175	0.178	0.179
\mathcal{M}_3 (485)	0.104	0.112	0.074	0.052	0.041	0.037
\mathcal{M}_4 (778)	0.057	0.051	0.028	0.038	0.047	0.049
\mathcal{M}_5 (946)	0.048	0.034	0.022	0.028	0.032	0.024
\mathcal{M}_6 (1256)	0.031	0.015	0.017	0.011	0.007	0.007
\mathcal{M}_7 (1725)	0.026	0.014	0.011	0.010	0.008	0.008
\mathcal{M}_8 (2028)	0.016	0.010	0.009	0.009	0.009	0.009
\mathcal{M}_9 (3544)	0.012	0.009	0.009	0.009	0.009	0.009
\mathcal{M}_{10} (5024)	0.011	0.008	0.007	0.007	0.007	0.007

Table 11. Average values of f_{st} against the size of MIoT (unsupervised approach).

MIoT (Size)	Average f_{st}
\mathcal{M}_1 (176)	0.904
\mathcal{M}_2 (301)	0.722
\mathcal{M}_3 (485)	0.635
\mathcal{M}_4 (778)	0.584
\mathcal{M}_5 (946)	0.580
\mathcal{M}_6 (1256)	0.576
\mathcal{M}_7 (1725)	0.516
\mathcal{M}_8 (2028)	0.477
\mathcal{M}_9 (3544)	0.452
\mathcal{M}_{10} (5024)	0.426

From the analysis of these tables we can observe that both the supervised and the unsupervised approaches really contribute to decrease the number of the nodes of a MIoT involved in the information dissemination, and, therefore, to increase the efficiency of this task. As for the supervised approach, we observe that the decrease of the number of involved nodes is always high. It becomes very high as the MIoT size and the number of keywords composing the query increase. As for the unsupervised approach, we observe that it leads to a decrease of the number of the MIoT nodes involved in the dissemination task. However, this decrease is minimum for small MIoT, whereas it becomes significant for large ones (i.e., for MIoT with a number of nodes higher than 1256).

We performed a second experiment in this direction. Specifically, given a pair (n_s, n_t) of a MIoT such that information must be disseminated from n_s to n_t and there exists at least one virtual IoT \mathcal{V} comprising both n_s and n_t , we computed the fraction $g_{st}^{\mathcal{V}}$ (resp., $\widehat{g_{st}^{\mathcal{V}}}$) of the nodes of the MIoT involved in the diffusion of information from n_s to n_t and having at least one content of the disseminated information registered in their profile (which implies that, in principle, they could benefit from the information they are required to disseminate). As in the previous experiment, we computed $g_{st}^{\mathcal{V}}$ by assuming the existence of \mathcal{V} and, hence, by performing the information dissemination task through it; by contrast, we computed $\widehat{g_{st}^{\mathcal{V}}}$ by carrying out the information dissemination task through the whole MIoT. Finally, we computed $g_{st} = \frac{g_{st}^{\mathcal{V}}}{\widehat{g_{st}^{\mathcal{V}}}}$. Roughly speaking, it denotes how much the presence of the virtual IoT \mathcal{V} can contribute to require information dissemination tasks only to nodes possibly benefiting of it. A value of this coefficient higher than 1 denotes a positive contribution of \mathcal{V} ; the higher this value the higher the contribution. As in the previous experiment, we computed the average

values of g_{st} by operating on the ten MIoT introduced in Section 6.2 and by performing the same tasks described therein for both the supervised and the unsupervised approaches. The obtained results are reported in Tables 12 and 13.

Table 12. Average values of g_{st} against the size of MIoT and queries used to generate the virtual IoTs (supervised approach).

MIoT (Size)	Average g_{st}					
	$ Q = 1$	$ Q = 2$	$ Q = 4$	$ Q = 6$	$ Q = 8$	$ Q = 10$
\mathcal{M}_1 (176)	4.018	2.792	2.223	1.918	1.331	1.321
\mathcal{M}_2 (301)	3.563	2.619	2.445	2.009	1.683	1.664
\mathcal{M}_3 (485)	3.269	2.370	1.426	1.528	1.626	1.674
\mathcal{M}_4 (778)	3.130	2.168	2.367	1.916	1.494	1.325
\mathcal{M}_5 (946)	3.232	2.102	1.864	1.712	1.461	1.391
\mathcal{M}_6 (1256)	3.467	1.979	1.378	1.412	1.438	1.452
\mathcal{M}_7 (1725)	3.476	2.224	1.414	1.444	1.494	1.492
\mathcal{M}_8 (2028)	3.496	2.669	1.489	1.491	1.521	1.545
\mathcal{M}_9 (3544)	3.507	2.712	1.612	1.624	1.631	1.632
\mathcal{M}_{10} (5024)	3.517	2.926	1.783	1.841	1.864	1.874

Table 13. Average values of g_{st} against the size of MIoT (unsupervised approach).

MIoT (Size)	Average g_{st}
\mathcal{M}_1 (176)	1.341
\mathcal{M}_2 (301)	1.269
\mathcal{M}_3 (485)	1.211
\mathcal{M}_4 (778)	1.177
\mathcal{M}_5 (946)	1.173
\mathcal{M}_6 (1256)	1.171
\mathcal{M}_7 (1725)	1.194
\mathcal{M}_8 (2028)	1.273
\mathcal{M}_9 (3544)	1.281
\mathcal{M}_{10} (5024)	1.301

The analysis of these tables is a further confirmation of the efficiency of our approach. Indeed, thanks to the presence of virtual IoTs, the fraction of nodes participating to the spreading of information that can also benefit from this task increases remarkably.

The results of Tables 10 and 11, along with the ones of Tables 12 and 13, agree to evidence that the discovery of virtual IoTs is highly beneficial in terms of efficiency for the information dissemination task in a MIoT. In this case, the contribution of \mathcal{V} in increasing the efficiency of the spreading task, by limiting it mainly to nodes that could benefit from the information they are disseminating, is very high for the supervised approach when $|Q| = 1$ or $|Q| = 2$. When $|Q|$ increases, this contribution decreases, even if it remains still significant. As for the unsupervised approach, the contribution of \mathcal{V} can be always observed even if it is less evident than the one characterizing the supervised approach.

6.6. Number and Size of Returned Virtual IoTs

This last experiment makes sense only for the unsupervised approach. Through it we aimed at investigating how the number and the size of returned virtual IoTs (and, therefore, the number and the size of returned clusters) vary when the MIoT size increases. To make this experiment significant, we maintained constant all the parameters of the adopted clustering algorithm. We considered the MIoT $\mathcal{M}_1 \cdots \mathcal{M}_{10}$ used in the previous experiments because, in this way, we had the possibility to investigate MIoT sizes ranging from 176 to 5024 nodes. We report the obtained results in Table 14.

From the analysis of this table we can observe that the average size of virtual IoTs:

- increases when the MIoT size ranges from 176 to 946;

- slightly increases when the MIoT size ranges from 946 to 2028;
- remains essentially constant when the MIoT size is higher than 2028.

In the meantime, the number of clusters:

- slightly increases when the MIoT size ranges from 176 to 946;
- increases when the MIoT size ranges from 946 to 2028;
- highly increases when the MIoT size is higher than 2028.

Table 14. Average size and number of virtual IoTs against the increase of the MIoT size (unsupervised approach).

MIoT (Size)	Average Size of Virtual IoTs	Number of Virtual IoTs
\mathcal{M}_1 (176)	22.44	10
\mathcal{M}_2 (301)	28.21	13
\mathcal{M}_3 (485)	36.64	16
\mathcal{M}_4 (778)	40.82	22
\mathcal{M}_5 (946)	44.66	24
\mathcal{M}_6 (1256)	46.74	30
\mathcal{M}_7 (1725)	48.12	39
\mathcal{M}_8 (2028)	50.24	45
\mathcal{M}_9 (3544)	50.46	78
\mathcal{M}_{10} (5024)	50.64	105

The obtained results are extremely interesting because they confirm the soundness of the reasoning made in Section 5.3. In particular, this experiment confirms the scalability of our approach. As a matter of fact, after the virtual IoTs have been constructed offline, their usage for querying and for the other tasks of interest for the user can be performed online. Now, we observed that the number of available virtual IoTs highly increases when the MIoT size increases. However, because the size of each virtual IoT is only slightly impacted by the growth of the corresponding MIoT, and because user tasks generally involve one or at most a few of available virtual IoTs, we can conclude that our approach is scalable with respect to the size variation of the MIoT.

7. Conclusions

In this paper, we have discussed about the attempt of “humanizing” things. We have seen that this trend will become increasingly challenging in the future because things are becoming more and more smart and social. As a consequence, it appears natural to apply concepts typical of social networking to the Internet of Things. Actually, as things are becoming increasingly heterogeneous in their formats, semantics and behaviors, it appears even better to apply social internetworking ideas and concepts to this scenario.

For this reason, we have decided to adopt the MIoT paradigm as the reference one for our proposal. With the support of this paradigm, we have proposed a rich and high-level abstraction profile of a thing, taking into account the content that it exchanged with the other things in the past. Then, we have introduced the concept of topic-guided virtual IoT and we have proposed a supervised and an unsupervised approach to constructing topic-guided virtual IoTs in a Multi-IoT scenario.

This paper must not be considered as an ending point; on the contrary, it is a starting point for future research efforts. Indeed, the definition of a thing’s profile and the usage of paradigms, like MIoT, allowing multiple IoTs to be modeled in a way analogous to how multiple social networks interacting with each other are modeled, allow us to investigate the possible extension to the IoT context of many research themes already analyzed for social networks. For instance, it could be possible: (i) to model the concepts of trust and reputation of a thing in the IoTs it belongs to; (ii) to develop “team building” approaches aiming at constructing teams of things to perform a certain activity; (iii) to investigate new

forms of centrality of a thing in a MIIoT based on both its position and its profile. Actually, these are just three of the many possible future developments of our research in such a rapidly evolving and very promising scenario.

Author Contributions: Conceptualization, D.U. and A.N.; methodology, D.U. and L.V.; software, L.V. and P.L.G.; validation, L.V. and P.L.G.; formal analysis, D.U. and A.N.; investigation, A.N. and L.V.; resources, A.N.; data curation, D.U. and L.V.; writing—original draft preparation, D.U., A.N., P.L.G. and L.V.; writing—review and editing, D.U. and A.N.; visualization, L.V.; supervision, D.U.; project administration, Domenico Ursino.; funding acquisition, D.U.

Funding: This work was partially funded by the Department of Information Engineering at the Polytechnic University of Marche under the project “A network-based approach to uniformly extract knowledge and support decision making in heterogeneous application contexts” (RSAB 2018).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guinard, D.; Trifa, V.; Mattern, F.; Wilde, E. From the internet of things to the web of things: Resource-oriented architecture and best practices. In *Architecting the Internet of Things*; Springer: Berlin, Germany, 2011; pp. 97–129.
2. Santos, J.; Wauters, T.; Volckaert, B.; Turck, F.D. Resource Provisioning in Fog Computing: From Theory to Practice. *Sensors* **2019**, *19*, 2238. [[CrossRef](#)] [[PubMed](#)]
3. Distefano, S.; Merlino, G.; Puliafito, A. Enabling the cloud of things. In Proceedings of the International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS'2012), Palermo, Italy, 4–6 July 2012; pp. 858–863.
4. Park, M.; Oh, H.; Lee, K. Security Risk Measurement for Information Leakage in IoT-Based Smart Homes from a Situational Awareness Perspective. *Sensors* **2019**, *19*, 2148. [[CrossRef](#)] [[PubMed](#)]
5. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A survey. In *Computer Networks*; Elsevier: Amsterdam, The Netherlands, 2010; Volume 54, pp. 2787–2805.
6. Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. In *Future Generation Computer Systems*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 29, pp. 1645–1660.
7. Lee, J.; Yu, S.; Park, K.; Park, Y.; Park, Y. Secure Three-Factor Authentication Protocol for Multi-Gateway IoT Environments. *Sensors* **2019**, *19*, 2358. [[CrossRef](#)] [[PubMed](#)]
8. Araniti, G.; Orsino, A.; Militano, L.; Wang, L.; Iera, A. Context-aware information diffusion for alerting messages in 5G mobile social networks. *IEEE Int. Things J.* **2017**, *4*, 427–436. [[CrossRef](#)]
9. Atzori, L.; Iera, A.; Morabito, G. SIoT: Giving a social structure to the Internet of Things. *IEEE Commun. Lett.* **2011**, *15*, 1193–1195. [[CrossRef](#)]
10. Farris, I.; Girau, R.; Militano, L.; Nitti, M.; Atzori, L.; Iera, A.; Morabito, G. Social virtual objects in the edge cloud. *IEEE Cloud Comput.* **2015**, *2*, 20–28. [[CrossRef](#)]
11. Atzori, L.; Iera, A.; Morabito, G. From “smart objects” to “social objects”: The next evolutionary step of the Internet of Things. *IEEE Commun. Mag.* **2014**, *52*, 97–105. [[CrossRef](#)]
12. Saleem, Y.; Crespi, N.; Rehmani, M.; Copeland, R.; Hussein, D.; Bertin, E. Exploitation of social IoT for recommendation services. In Proceedings of the World Forum on Internet of Things (WF-IoT'16), Reston, VA, USA, 12–14 December 2016; pp. 359–364.
13. Baldassarre, G.; Lo Giudice, P.; Musarella, L.; Ursino, D. A paradigm for the cooperation of objects belonging to different IoTs. In Proceedings of the International Database Engineering & Applications Symposium (IDEAS 2018), Villa San Giovanni, Italy, 18–20 June 2018; pp. 157–164.
14. Baldassarre, G.; Lo Giudice, P.; Musarella, L.; Ursino, D. The MIIoT paradigm: main features and an “ad-hoc” crawler. *Future Gener. Comput. Syst.* **2019**, *92*, 29–42. [[CrossRef](#)]
15. Buccafurri, F.; Foti, V.; Lax, G.; Nocera, A.; Ursino, D. Bridge Analysis in a Social Internetworking Scenario. *Inf. Sci.* **2013**, *224*, 1–18. [[CrossRef](#)]
16. McPherson, M.; Smith-Lovin, L.; Cook, J. Birds of a feather: Homophily in social networks. *Annu. Rev. Soc.* **2001**, *27*, 415–444. [[CrossRef](#)]

17. Şimşek, Ö.; Jensen, D. Navigating networks by using homophily and degree. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 12758–12762. [[CrossRef](#)] [[PubMed](#)]
18. Schiaffino, S.; Amandi, A. Intelligent user profiling. In *Artificial Intelligence An International Perspective*; Springer: Berlin, Germany, 2009; pp. 193–216.
19. Plantié, M.; Crampes, M. Survey on social community detection. In *Social Media Retrieval*; Springer: Berlin, Germany, 2013; pp. 65–85.
20. Atzori, L.; Iera, A.; Morabito, G. Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm. *Ad Hoc Netw.* **2017**, *56*, 122–140. [[CrossRef](#)]
21. Miorandi, D.; Sicari, S.; Pellegrini, F.D.; Chlamtac, I. Internet of things: Vision, applications and research challenges. *Ad Hoc Netw.* **2012**, *10*, 1497–1516. [[CrossRef](#)]
22. Perera, C.; Zaslavsky, A.; Christen, P.; Georgakopoulos, D. Context aware computing for the Internet of Things: A survey. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 414–454. [[CrossRef](#)]
23. Tsai, C.; Lai, C.; Vasilakos, A. Future Internet of Things: open issues and challenges. *Wirel. Netw.* **2014**, *20*, 2201–2217. [[CrossRef](#)]
24. Stojmenovic, I.; Olariu, S. Data-centric protocols for wireless sensor networks. In *Handbook of Sensor Networks: Algorithms and Architectures*; Wiley: Hoboken, NJ, USA, 2005; pp. 417–456.
25. Zhang, Y.; Raychadhuri, D.; Grieco, L.; Baccelli, E.; Burke, J.; Ravindran, R.; Wang, G.; Lindgren, A.; Ahlgren, B.; Schelen, O. Requirements and Challenges for IoT over ICN. 2015. Available online: <https://tools.ietf.org/html/draft-zhang-icnrg-icniot-requirements-00> (accessed on 30 March 2019).
26. Zhang, Y.; Raychadhuri, D.; Ravindran, R.; Wang, G. ICN Based Architecture for IoT. 2013. Available online: <https://tools.ietf.org/html/draft-zhang-iot-icn-challenges-02> (accessed on 30 March 2019)
27. Amadeo, M.; Campolo, C.; Iera, A.; Molinaro, A. Named data networking for IoT: An architectural perspective. In Proceedings of the European Conference on Networks and Communications (EuCNC'2014), Bologna, Italy, 23–26 June 2014; pp. 1–5.
28. Qin, Y.; Sheng, Q.; Falkner, N.; Dustdar, S.; Wang, H.; Vasilakos, A. When things matter: A survey on data-centric internet of things. *J. Netw. Comput. Appl.* **2016**, *64*, 137–153. [[CrossRef](#)]
29. Amadeo, M.; Campolo, C.; Quevedo, J.; Corujo, D.; Molinaro, A.; Iera, A.; Aguiar, R.; Vasilakos, A. Information-centric networking for the internet of things: challenges and opportunities. *IEEE Netw.* **2016**, *30*, 92–100. [[CrossRef](#)]
30. Perera, C.; Qin, Y.; Estrella, J.; Reiff-Marganiec, S.; Vasilakos, A. Fog computing for sustainable smart cities: A survey. *ACM Comput. Surv. (CSUR)* **2017**, *50*, 32. [[CrossRef](#)]
31. Tei, K.; Gurgun, L. ClouT: Cloud of things for empowering the citizen clout in smart cities. In Proceedings of the World Forum on Internet of Things (WF-IoT'2014), Seoul, Korea, 6–8 March 2014; pp. 369–370.
32. Karnouskos, S. Smart houses in the smart grid and the search for value-added services in the cloud of things era. In Proceedings of the International Conference on Industrial Technology (ICIT'2013), Cape Town, South Africa, 25–28 February 2013; pp. 2016–2021.
33. Xu, K.; Qu, Y.; Yang, K. A tutorial on the internet of things: From a heterogeneous network integration perspective. *IEEE Netw.* **2016**, *30*, 102–108. [[CrossRef](#)]
34. Wan, J.; Tang, S.; Shu, Z.; Li, D.; Wang, S.; Imran, M.; Vasilakos, A. Software-defined industrial internet of things in the context of industry 4.0. *IEEE Sens. J.* **2016**, *16*, 7373–7380. [[CrossRef](#)]
35. Wan, J.; Liu, J.; Shao, Z.; Vasilakos, A.; Imran, M.; Zhou, K. Mobile crowd sensing for traffic prediction in internet of vehicles. *Sensors* **2016**, *16*, 88. [[CrossRef](#)] [[PubMed](#)]
36. Guinard, D.; Fischer, M.; Trifa, V. Sharing using social networks in a composable web of things. In Proceedings of the International Conference on Pervasive Computing and Communications (PERCOM 2010), Mannheim, Germany, 29 March–2 April 2010; pp. 702–707.
37. Ning, H.; Wang, Z. Future internet of things architecture: like mankind neural system or social organization framework? *IEEE Commun. Lett.* **2011**, *15*, 461–463. [[CrossRef](#)]
38. Kranz, M.; Roalter, L.; Michahelles, F. Things that Twitter: social networks and the Internet of Things. In Proceedings of the International Workshop on Pervasive Computing (Pervasive 2010), Helsinki, Finland, 17–20 May 2010; pp. 1–10.
39. Holmquist, L.; Mattern, F.; Schiele, B.; Alahuhta, P.; Beigl, M.; Gellersen, H. Smart-its friends: A technique for users to easily establish connections between smart artefacts. In Proceedings of the International Conference on Ubiquitous Computing (Ubicomp'2001), Atlanta, GA, USA, 30 September–2 October 2001; pp. 116–122.

40. Atzori, L.; Iera, A.; Morabito, G.; Nitti, M. The Social Internet of Things (SIoT)– when social networks meet the Internet of Things: Concept, architecture and network characterization. *Comput. Netw.* **2012**, *56*, 3594–3608. [[CrossRef](#)]
41. Ortiz, A.; Hussein, D.; Park, S.; Han, S.; Crespi, N. The cluster between internet of things and social networks: Review and research challenges. *IEEE Int. Things J.* **2014**, *1*, 206–215. [[CrossRef](#)]
42. Oh, S.; Kim, Y.; Cho, S. An Interoperable Access Control Framework for Diverse IoT Platforms Based on OAuth and Role. *Sensors* **2019**, *19*, 1884. [[CrossRef](#)] [[PubMed](#)]
43. Sobral, J.; Rodrigues, J.; Rabêlo, R.; Al-Muhtadi, J.; Korotaev, V. Routing Protocols for Low Power and Lossy Networks in Internet of Things Applications. *Sensors* **2019**, *19*, 2144. [[CrossRef](#)] [[PubMed](#)]
44. Ganzha, M.; Paprzycki, M.; Pawłowski, W.; Szmeja, P.; Wasielewska, K. Semantic interoperability in the Internet of Things: An overview from the INTER-IoT perspective. *J. Netw. Comput. Appl.* **2017**, *81*, 111–124. [[CrossRef](#)]
45. Li, J.; Wang, H.; Khan, S. A semantics-based approach to large-scale mobile social networking. *Mob. Netw. Appl.* **2012**, *17*, 192–205. [[CrossRef](#)]
46. Suciu, G.; Halunga, S.; Vulpe, A.; Suciu, V. Generic platform for IoT and cloud computing interoperability study. In Proceedings of the International Symposium on Signals, Circuits and Systems (ISSCS'13), Iasi, Romania, 11–12 July 2013; pp. 1–4.
47. Zhao, F.; Sun, Z.; Jin, H. Topic-centric and semantic-aware retrieval system for internet of things. *Inf. Fusion* **2015**, *23*, 33–42. [[CrossRef](#)]
48. Nocera, A.; Ursino, D. PHIS: a system for scouting potential hubs and for favoring their “growth” in a Social Internetworking Scenario. *Knowl. Based Syst.* **2012**, *36*, 288–299. [[CrossRef](#)]
49. Diamantini, D.; Potena, D.; Storti, E. A virtual mart for knowledge discovery in databases. *Inf. Syst. Front.* **2013**, *15*, 447–463. [[CrossRef](#)]
50. Wang, L. Using the relationship of shared neighbors to find hierarchical overlapping communities for effective connectivity in IoT. In Proceedings of the International Conference on Pervasive Computing and Applications (ICPCA'11), Port Elizabeth, South Africa, 26–28 October 2011; pp. 400–406.
51. Karyotis, V.; Tsitsekis, K.; Sotiropoulos, K.; Papavassiliou, S. Big Data Clustering via Community Detection and Hyperbolic Network Embedding in IoT Applications. *Sensors* **2018**, *18*, 1205. [[CrossRef](#)] [[PubMed](#)]
52. Barthwal, R.; Misra, S.; Obaidat, M. Finding overlapping communities in a complex network of social linkages and Internet of things. *J. Supercomput.* **2013**, *66*, 1749–1772. [[CrossRef](#)]
53. Kowshalya, A.; Valarmathi, M. Community Detection in the Social Internet of Things Based on Movement, Preference and Social Similarity. *Stud. Inf. Control* **2016**, *25*, 499–506. [[CrossRef](#)]
54. Nitti, M.; Girau, R.; Atzori, L. Trustworthiness management in the social internet of things. *IEEE Trans. Knowl. Data Eng.* **2014**, *26*, 1253–1266. [[CrossRef](#)]
55. Misra, S.; Barthwal, R.; Obaidat, M.S. Community detection in an integrated Internet of Things and social network architecture. In Proceedings of the IEEE Global Communications Conference (GLOBECOM 2012), Anaheim, CA, USA, 3–7 December 2012; pp. 1647–1652.
56. IPSO Alliance. 2019. Available online: <https://www.ipso-alliance.org/> (accessed on 30 March 2019).
57. Bernstein, P.; Madhavan, J.; Rahm, E. Generic Schema Matching, Ten Years Later. *Proc. VLDB Endow.* **2011**, *4*, 695–701.
58. De Meo, P.; Quattrone, G.; Terracina, G.; Ursino, D. Integration of XML Schemas at various “severity” levels. *Inf. Syst.* **2006**, *31*, 397–434. [[CrossRef](#)]
59. Schaeffer, S. Graph clustering. *Comput. Sci. Rev.* **2007**, *1*, 27–64. [[CrossRef](#)]
60. Fahad, A.; Alshatri, N.; Tari, Z.; Alamri, A.; Khalil, I.; Zomaya, A.; Fofou, S.; Bouras, A. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE Trans. Emerg. Top. Comput.* **2014**, *2*, 267–279. [[CrossRef](#)]
61. Tsvetovat, M.; Kouznetsov, A. *Social Network Analysis for Startups: Finding Connections on the Social Web*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2011.

62. Blondel, V.; Guillaume, J.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [[CrossRef](#)]
63. Hirschman, A. The paternity of an index. *Am. Econ. Rev.* **1964**, *54*, 761–762.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).