Resource Discovery in Internet of Things: Current Trends and Future Standardization Aspects

Soumya Kanti Datta, Rui Pedro Ferreira Da Costa, Christian Bonnet
Mobile Communications Department, Eurecom
Biot, France
Emails – {dattas, ferreira, bonnet}@eurecom.fr

Abstract—To realize the vision of Internet of Things, there must be mechanisms to discover resources and their capabilities. Thus resource discovery becomes a fundamental requirement of any IoT platform. This paper provides a comprehensive categorization of the current technology landscape for discovery while pointing out their advantages and limitations. Then, a novel search engine based resource discovery framework is proposed. The framework comprises of a proxy layer which includes drivers for various communication technologies and protocols. There is a central registry to store the configurations of the resources and index them based on the configuration parameters. A “lifetime” attribute is introduced which denotes the period through which a resource is discoverable. The search engine ranks the resulting resources of a discovery request and returns an URI to directly access each resource. The functionalities of the proposed framework are exposed using RESTful web services. The framework allows discovery of both smart and legacy devices. We have performed gap analysis of current IoT standards for discovery mechanisms and provide suggestions to improve them to maintain interoperability. The outcomes are communicated to Standard Development Organizations like W3C (in Web of Things Interest Group) and oneM2M. The main contributions of the paper are – (i) discovery of resources, their capabilities & properties regardless of communication technologies and protocols used, (ii) integration of a search engine to provide indexing, look-up and ranking facilities into discovery framework, (iii) searching for both smart and legacy objects, (iv) a lifetime attribute through which resources remain discoverable, (v) flexibility in design which allows the framework to be integrated into a cloud based system, an M2M gateway or in a smartphone application and (vi) explaining the future standardization aspects.

I. INTRODUCTION

The Internet of Things (IoT) envisions seamlessly connecting the physical objects with the Internet. This in turn integrates the physical world into the digital world. This trend enables creating consumer-centric applications and services in various domains like intelligent home control (smart home), eHealth, intelligent transportation system, environmental monitoring etc. It is estimated that 50 Billions of smart objects will be connected to the Internet by 20201. To provide value-added services to the end users through IoT platforms, these devices must interact with the environment and among themselves. Such interaction in turn facilitates exchanging and processing metadata and reacting automatically to the environment. However, the diverse nature of smart objects, their capabilities & properties, communication technologies add to the complexity of effective realization of the IoT platforms. Therefore, to realize the vision of IoT, there must be mechanisms available for automatic discovery of resources, their properties and capabilities as well as the means to access them. Furthermore, such discovery mechanisms also depend on other services like configuration management, registration and un-registration of smart objects. This makes discovery a fundamental requirement for any IoT framework and platform.

However, traditional web based discovery services are not suitable for doing the same in IoT because of different requirements of IoT. These are outlined in [18]. In this paper, we have proposed a framework for automatic and efficient resource discovery in IoT. Depending on the use case, resource could mean physical thing(s) and/or associated metadata or the services provided by the thing(s). The framework incorporates a search engine which provides a “look-up” discovery feature. The scopes of the discovery mechanisms include both local and remote aspects in terms of location and network. In the local scope, discovery takes place within a gateway of an intelligent home environment in a local network. The remote scope takes care of discovery from smart city perspective and remote network. At the same time, the scopes also include one-time discovery (applicable to smart home) as well as long standing (pub-sub style discovery) mechanisms. The framework also supports multiple communication technologies and protocols through a proxy layer which basically includes the necessary drivers to provide binding to specific protocols. There is a central registry managing the registration and un-registration of resources for both smart and legacy devices. The central registry extends the M2M device management framework presented in [17]. Consumers can access the functionalities of the framework through RESTful web services shown in Figure 1. The main contributions of the paper are – (i) discovery of resources, their capabilities and properties regardless of communication technologies and protocols used, (ii) integration of a search engine to provide indexing, look-up and ranking facilities into discovery framework, (iii) searching for both smart and legacy objects, (iv) a lifetime attribute through which resources remain discoverable, (v) flexibility in design which allows the framework to be integrated into a cloud based system, an M2M gateway or in a smartphone application and (vi) explaining the future standardization aspects.

The rest of the paper is organized as follows. Section II surveys and categorizes the state-of-the-art and highlights the pros and cons of various approaches. Section III presents the framework and describes its components while highlighting the novel aspects of the work. Section IV outlines future standardization aspects. Finally, the paper concludes with some future directions.

II. RELATED WORKS

This section categorizes the related works in resource discovery and reports about the technologies used as well as their advantages and limitations.

A. Distributed and P2P discovery services

The authors of [1] reports a system for distributed discovery service. The philosophy behind such system is peer-to-peer (P2P) approach that adopts the distributed hash table (DHT) techniques. It supports multi-attribute and range queries. The authors also describe experiments on RFID based scenarios. Liu et al presents architecture for distributed resource discovery (DRD) aimed for Internet of Things [2]. The distributed resource peers communicate among each using P2P overlay protocol. The resource peers also handle the M2M device registration and assist in the overall discovery process. As for resource identification, CoAP based URI is suggested as it contains the resource path and the name of the necessary endpoint. The authors use a MAC-address hashing technique to generate unique names for the endpoints. But this suffers from a problem that the MAC address can be spoofed using software and it can potentially lead to duplicate names for the endpoints. The resource description registration stores several information (IP address, resource path, resource type, content type and endpoint name) into the P2P overlay. A resource discovery component (RDC) is used by the clients and the component locally checks for the requested resource. If found, the description is returned to the client or else the P2P overlay is used to search for it. When found it is sent to the clients as well as cached to the first resource peer.

Cirani et al reports about a scalable and self-configurable P2P architecture for service discovery (SD) [3]. Utilization of P2P technologies enables deployment of distributed and large scale infrastructure for SD. IoT gateway acts as a backbone of the SD architecture. The gateway keeps track of any things joining or leaving its network and updates the list maintained at its CoAP server. The SD is based on CoAP where a GET request is sent to /.well-known/core to retrieve the necessary information of the attached resources. In the distributed architecture several gateways are interlinked through two P2P overlays namely distributed local service (DLS) and distributed geographic table (DGT) to facilitate global service discovery.

B. Centralized architecture for resource discovery

Jara et al have presented a mechanism for the global resource discovery of devices and sensors across several scenarios [4]. An infrastructure called ‘digcovery’ is developed which allows sensors to be registered into a common centralized infrastructure. A mobile service is developed which allows the clients to discover and access the sensors. The architecture employs several ‘Digirectory’ to handle different resources. Each digirectory is attached to a particular domain and connected to the objects of the domain over NFC, 6LoWPAN, IPv6 etc. The mobile application takes advantage of geo-location and context awareness for discovery phase. The application offers several avenues for discovery. Users willing to provide services can register about their devices and sensors to the back-end of the architecture through RFID tags, NFC or QR codes which enables the mobile clients to discover. The architecture also implements digirectories which allow legacy objects and EPC based objects to be included in the common infrastructure. The paper describes several technologies (IPv6, 6LoWPAN, CoAP, Web Services) which integrate the real world devices into IoT systems. At the same time, they provide ways to interact with those physical objects using RFID tags, NFC, Bluetooth and QR Codes. The paper [5] proposes a service oriented discovery framework based on popular web standards (REST and JSON). The framework is integrated into a centralized architecture. A central registry is the backbone of the architecture and is responsible for indexing smart objects according to domains they belong. The searching of resources in a given domain can be done by simply connecting to the central registry which provides a direct reference of the objects to the clients. But such centralized architecture and indexing are done using domains. This does not consider the fact that same object can be shared among several domains.

C. CoAP based service discovery

CoAP includes a mechanism for service discovery [6]. The CoAP servers expose a RESTful web service at /.well-known/core which can reply to any CoAP client requesting service discovery. The client receives many information including a list of the available resources, an attribute specifying the format of metadata of the resource etc. Although useful in several scenarios, but the approach has several shortcomings – (i) CoAP does not specify how a thing should join the CoAP server first time and announce itself, (ii) there is no specification on how a remote client can look up into the resource directory (RD) and query for the resource of interest, (iii) a centralized approach using RD and CoAP suffers from scalability issue. Also it is relatively easier to perform DoS attack resulting in unavailability of discovery process and resources. Ishaq et al have identified several gaps in current IoT based systems [8] in terms of (i) automatic discovery of sensors, (ii) integration with DNS and (iii) user friendly integration and access of sensors from web browsers. To mitigate the challenges, the authors proposed a self-configuration and bootstrapping mechanism that enable sensor discovery. The system uses CoAP and DNS and provides protocol translation between CoAP and HTTP making any IPv6 enabled sensor to be discoverable. The CoAP client connects to a well-known entry point in the CoAP server (defined by CoRE resource discovery). Finally, the authors have illustrated the self-configuration process along with sensors discovery and resource access. This work utilizes RESTful interactions between sensors, gateways and servers as well as IETF RFCs which promote interoperability with existing solutions.

D. Semantic based discovery

Zhou and Ma present an ontology focused web service matching algorithm aimed at IoT systems [7]. As a proof-of-concept, they have portrayed an ontology concept for vehicular sensor. The algorithm calculates semantic similarity, reliability and combines them to work out the maximum value of the
required concepts of the web services. Then a matching degree is computed to find out the relevant web services. Authors Alam and Noll [11] have introduced a semantic based framework which uses the concept of service advertisement of a smart object. They argue that such mechanism makes the service registration easier which in turn facilitates discovery. The advertisement contains a service metadata including name, id, endpoint, location and semantic annotation link. Another semantic based service discovery is presented in [12]. It proposes a middleware which performs SD using semantic web technologies on the contextual information inferred from sensor data. Further semantic based discovery service can be found in [13]. It looks at the discovery from the Web of Things point of view and uses multiple mapping scheme called “Discovery Strategies” to semantically discover resources. It aims at discovering the functionalities provided by WoT devices. The proposed DiscoWoT extensively uses Microformats and Microdata along semantic web technologies over RESTful web services. The resources are represented semantically and implemented with JSON to preserve interoperability. But the main limitation here is that the network addresses of the resources have to be known and if the resource is not connected to the web, then that is not discoverable.

E. Search engine for resource discovery

A hybrid search engine (SE) is proposed in [16]. The authors have noted that there is very limited work on search engine for IoT and the existing ones does not support multimodal search like spatial-temporal, value-based and keyword-based criteria. The proposed SE takes into account these criteria and its architecture is composed of three layers. The sensor and device monitoring layer consists of the physical things. The data generated by these things are stored at the storage layer. It contains several Raw-Data Storages and each of such Storages manage huge volume of things. And there is an index layer on top of the storage layer managing three indices, one each for full-text keyword, spatial-temporal and value-symbolized keyword based searches. The paper then discusses storage method of data generated by the things. The main goal here has been developing a SE for effective multimodal query processing to obtain data generated by things in real time. The performance evaluation points out that with huge volume of sensors, keyword based searches take the minimum time to discover the sensor data. But the problem of IoT-SVK is that the lowest layer generates unstructured data and focused on retrieving the sensor data rather than things description etc.

F. Utilization of ONS and DNS

This paper [9] has utilized object name service (ONS) and information discovery as a part of a distributed information service system. The products used in the system are equipped with a tracking code. ONS provides the necessary mapping of the product code with an IoT resource address. Therefore, in principle, the lookup service is composed of a code resolver and DNS which stores all relevant information. Since agricultural products may go through different supply chain phases, additional discovery service (DS) is introduced in the system. DS maps an object ID or code to a list of IoT information service servers. Ishaq et al [10] have identified several gaps in current IoT based systems in terms of (i) automatic discovery of sensors, (ii) integration with DNS and (iii) user friendly integration and access of sensors from web browsers. To mitigate the challenges, the authors proposed a self-configuration and bootstrapping mechanism that enable sensor discovery. UPnP and ONS are further used to assist in integrated device and service discovery in [14]. This paper target the home automation devices.

Apart from the discussions above, an adaptive and context aware service discovery protocol is presented in [15]. The paper outlines several requirements for such protocol and the proposed architecture takes care of registration and grouping of objects. The grouping is done based on location. The discovery is based on a directory agent and contextual information. The system has unique characteristics in terms of adaptive reporting timer and optimal service selection.

III. PROPOSED DISCOVERY FRAMEWORK

This section presents the proposed discovery framework along with its building blocks and highlights the novel aspects. Figure 1 portrays the architecture of the framework which is composed of three core layers.

A. Proxy layer

One of the main goals of the proposed framework is to discover physical things regardless of communication technologies and protocols used by the things. This is accomplished by the proxy layer. The functionalities of the proxy-in and proxy-out are introduced in [19]. In this case, the layer includes the necessary drivers for the low power radio links and provides binding to appropriate protocols. It also facilitates the integration of legacy devices to be discoverable by the framework. This is done by managing the configurations of those legacy devices so that they appear to be smart devices from higher operational layers.
B. Discovery layer

The discovery layer constitutes of four main building blocks, namely, configuration registry, search engine, indexing API and lifetime. The configuration registry provides a database to store and manage the configuration parameters of the resources. The configuration registry API is accessed by the resources through the proxy layer RESTful web services. The API manages the registration and un-registration of devices where the configuration metadata are represented using CoRE Link Format while LwM2M is used for the management of resources [17]. Once resources are registered, they are indexed using the indexing API. The motivation behind the indexing is to expedite the search process and save time. The search engine receives the discovery request from the end user along with some input keywords or parameters. Using the same API, indices are extracted from the input request and then matched with the ones stored into the configuration registry. The matched results are then ranked based on relevance, availability, access control policies and scopes. The metadata for the discovery response contains an URI corresponding to each of the resources as well as the information about their capabilities and properties. The lifetime attribute is a period of time through which resources are discoverable.

C. Service enablement layer

This layer exposes the discovery mechanism as well as additional functionalities to the consumers through RESTful web services. An important function is the access control which restricts the search operation to the resources to which the consumer or another IoT application has access to. The discovery request is directed to the ‘search’ service which forwards it to the search engine at discovery layer. In return the engine provides the list of discovered resources back to requestor. Finally, the subscription and notification are used to subscribe to periodic discovery notifications. The beacon based advertisement corresponds to UriBeacon2 (formerly known as Google’s physical web) approach where resources periodically broadcast their URLs over Bluetooth Low Energy (BLE) radio. Since the proxy layer includes that technology, the service enablement layer includes a web service for beacon based discovery. In this exceptional case, there is no need to search for resources through the discovery layer.

D. Case Study with Intelligent Transport Systems

In this section we will take a brief overview of a case study application in which our discovery framework is integrated in ITS environment. Integration of our framework with the access layer is not limited by the boundaries of discovery itself, however to remain in the scope of this paper we shall remain focused on the discovery aspects.

Within the proposed architecture, a proxy layer is proposed with a two-fold purpose: to introduce a level of abstraction that would enable our solution to work across a multitude of deployed technologies (the role of the proxy manager); and to mask the complexity of such access technologies by providing proper interfacing mechanisms for each technology (the role of the technology specific proxies). As with any access technology, ITS has its own specific proxy that will enable us to recover information related to available devices within its domain.

Discovery is a primary service already embedded in ITS, as such is a basic need of mobile environments [20]. Despite the existing service, interfacing between the IoT approach and ITS is not a straightforward matter, as our IoT framework requires a unique way of addressing each and every device in the form of a URI, while in an ITS environment, addressing a single device can be done in various ways depending on the deployment. As a first taste of such, ITS deployments may use one or more networking and transport protocols, such as, TCP/UDP over IPv6 or a Geo networking stack [21], both of which use location and addressing in a different fashion.

Main proxy functions are implemented as an application within the ITS architecture, which are aware of available transportation methods of the current deployment. If, for instance an IP stack is present, the proxy can make use of IP addresses while if a Geo Networking stack is present, a Geo Network Identifier or location coordinates are a more suitable option.

![Fig. 2. Proxy within ITS stack.](Image)

Therefore, a mechanism must be instated in the proxy to provide a suitable alternative to the conventional URI required by the IoT framework while keeping it as simple and as consistent as possible with the ITS location formats. Taking advantage of the basic composition of an URI, we propose the following scheme to accommodate the location of an ITS device within a URI. Firstly, we consider the two-part nature of a URI: the URL and the URN. Leveraging on it we build a composite address for any device considering the first part (URL) as the address of the anchoring gateway, which in our case would be the correspondent ITS proxy, while on the second part we would define the location of the ITS device based on the available format. This approach would enable a certain degree of required abstraction to include different formats, as well as fulfilling the purpose of individually addressing each device.

Furthermore, it leaves the proxy with the freedom to choose which way to address each device, as each URI specifically designates it as the entity entrusted with locally discovering and naming the devices (i.e., generating its URN) under its domain.

General procedure would involve firstly the URN generation by the ITS proxy upon discovery of the device by ITS embedded mechanisms such as beaconing. The discovery and communication of its IoT capabilities would then be put in

---

2 http://uribeacon.io/
motion by the IoT framework with the support of the ITS proxy. The URI generation should be put into place by the IoT framework in its own naming service taking into account the ITS proxy recommendation for the URN.

![Diagram of URI composition under ITS domains.](image)

The main benefits of the proposed discovery framework are - (i) the consumer can perform resource discovery regardless of communication technologies and protocols used and learn their capabilities & properties, (ii) integration of a lightweight search engine for indexing, look-up and ranking facilities into discovery framework, (iii) searching for both smart and legacy objects, (iv) a lifetime attribute through which resources remain discoverable which addresses the sleeping time of certain resources, (v) ability to deploy the framework into a cloud based system, an M2M gateway or in a smartphone application.

IV. FUTURE STANDARDIZATION ASPECTS

This section presents a gap analysis between research directions and standardization efforts in oneM2M and W3C. Improving current standards in terms of discovery will preserve interoperability among different IoT platforms and resources. We demonstrate that the proposed approach could bridge the identified gaps.

A. Current status

The standard development organizations (SDOs) (e.g. ETSI, oneM2M, W3C, IETF) working on creating relevant IoT standards have released various specifications for discovery. In oneM2M standard release version 1.0, discovery is considered as a common service function (CSF), one of the common service entities (CSE) [22]. The CSE could be a part of physical resources like things, an M2M gateway and even a remote cloud platform. But detailed discussion on how the discovery procedure should function has not received much attention.

The ETSI M2M architecture also points out the necessity of discovery at the M2M service enablement layer [23]. The ongoing efforts in W3C Web of Things Interest Group Task Force on Discovery [24] are looking at discovery from a similar perspective. But there is no concrete effort to standardize the discovery procedure.

B. Gap analysis

Our gap analysis shows that the SDOs should pay attention to the following aspects. They lack –

- A common format or syntax to describe the resources, units and domains [25]. Synonyms are also often not recognized. Therefore, searching for a resource using a synonym might not discover the resource.
- The content of discovery result metadata, ranking of the discovered resources are not standardized.
- The IoT platforms implementing discovery do not adhere to any interoperability mechanisms. Thus metadata exchange among IoT platforms become difficult. This might increase the challenges faced in smart city scenarios.
- SDOs do not provide any guidelines how to integrate adequate security mechanisms and access control functions into the discovery framework.
- The semantic components for semantic based discovery are not well studied.
- Discovering vehicular resources have not been well investigated.

C. Improving standardization

The proposed discovery mechanism could mitigate the mentioned gaps between researches and IoT standards. A list of suggestions to the SDOs are mentioned below.

- **Uniform vocabulary for resources**: The resources, their capabilities and properties should be described following a catalogue of uniform vocabulary. The discovery framework proposes to utilize the catalogue presented in our previous work [25]. This will maintain interoperability in discovery mechanisms implemented across different IoT platforms.
- **Standardized service enablement layer**: The proposed service enablement layer (SEL) exposes the core discovery layer functions to consumers and other IoT applications over RESTful interactions. A standardized SEL would ease the exchange of discovery metadata among IoT platforms and consumer applications. This will further allow to decouple the relationships among the platforms and applications from discovery mechanisms.
- **Metadata of discovery result**: The metadata structure should be represented in a serialized format like JSON or JSON-LD\(^1\) to maintain interoperability among other IoT platform components. The structure of the metadata itself should include URIs of discovered resources, their capabilities and properties. The metadata content should be ranked based on relevance, availability, access control policies and scopes as proposed here.
- **Access control**: The discovery requests and responses should always pass through proper access control functions. This will limit the discovery to authorized resources only as shown in this framework.
- **Security**: Although it is out of the scope of the paper, but SDOs should provide guidelines about how to maintain end-to-end security during discovery requests and integrity of the response metadata.

\(^1\) http://www.w3.org/TR/json-ld/
V. CONCLUSION

In a nutshell, the paper studies the state-of-the-art in resource discovery, categorizes them along with their advantages and limitations. Then a novel discovery framework is presented which contains three layers namely proxy, discovery and service enablement. The core discovery mechanism depends on configuration registry of resources, indexing of resources, a search engine and the lifetime attribute. The service enablement layer incorporates access control policies to limit discovery to the authorized resources only. Then a case study is presented which discusses how to adopt the framework to search for vehicular resources over ITS systems. Then we focus on IoT standard aspects. Our gap analysis reveals the various limitations of current standardization efforts. Consequently, we provide several suggestions to the SDOs on improving the standards on discovery to preserve interoperability. Finally, as of future direction, we are developing a prototype of the framework as well as examining the search and ranking algorithms further.

ACKNOWLEDGMENT

This work is supported by French research project DataTweet (ANR-13-INFR-0008)4.

REFERENCES
