Expert Systems with Applications 42 (2015) 3528–3538

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An experimental study on geospatial indexing for sensor service discovery

Wei Wang [a], Suparna De [b,*], Gilbert Cassar [c], Klaus Moessner [b]

[a] Department of Computer Science and Software Engineering, Xi'an Jiaotong Liverpool University, China
[b] Institute for Communication Systems, University of Surrey, UK
[c] UReason, 1-13 York Road, Maidenhead SL6 1SQ, UK

ARTICLE INFO

Article history:
Available online 15 December 2014

Keywords:
Sensor services
Sensor service discovery
Geospatial indexing
R-Tree
Semantic search

ABSTRACT

The Internet of Things enables human beings to better interact with and understand their surrounding environments by extending computational capabilities to the physical world. A critical driving force behind this is the rapid development and wide deployment of wireless sensor networks, which continuously produce a large amount of real-world data for many application domains. Similar to many other large-scale distributed technologies, interoperability and scalability are the prominent and persistent challenges. The proposal of sensor-as-a-service aims to address these challenges; however, to our knowledge, there are no concrete implementations of techniques to support the idea, in particular, large-scale, distributed sensor service discovery. Based on the distinctive characteristics of the sensor services, we develop a scalable discovery architecture using geospatial indexing techniques and semantic service technologies. We perform extensive experimental studies to verify the performance of the proposed method and its applicability to large-scale, distributed sensor service discovery.

1. Introduction

The Internet of Things (IoT) is seen as the most promising multidisciplinary effort that will help human beings live better lives by enabling many intelligent applications, e.g., smart office/home/city, remote healthcare, autonomous traffic control, emergency response, sustainable urban planning, and effective environmental monitoring and protection, to name a few. One of the driving forces behind this is that embedded devices (e.g., wireless sensors) with ever-increasing computation power and communication capabilities can be manufactured and deployed in large-scale while at low cost. The explosive development of the IoT in recent years has the potential to add millions or even billions of sensors to the future Internet. According to Cisco IBSG, there will be 25 billion devices connected to the Internet by 2015 and 50 billion by 2020 (Evans, 2011). Similar to the large-scale deployment of many other distributed technologies, interoperability and scalability are the persistent challenges for wireless sensor networks, which inevitably make efficient sensor discovery and sensor data access difficult.

To enable interoperability, the IoT community proposes to use semantic technologies for describing and annotating sensors and entities of interest. This has the potential to enable representing, storing, interconnecting, searching and organising information related to or generated by heterogeneous things (Atzori, Iera, & Morabito, 2010) and to facilitate the creation of "Semantic Sensor Web" (Sheth, Henson, & Sahoo, 2008). Some of the notable works in this line include the Semantic Sensor Network (SSN) ontology for sensor knowledge representation (Compton et al., 2012) developed by the W3C Semantic Sensor Networks Incubator Group; the ontology framework developed in (Roda & Musulin, 2014) for supporting intelligent analysis of sensor measurement data; and the semantic service modelling for real world IoT resources (De, Elsaleh, Barnaghi, & Meissner, 2012). To enable scalability, researchers propose to apply service-oriented principles to the design of IoT infrastructure, which can facilitate the development of large-scale, loosely-coupled IoT based applications and services. The concept of sensor-as-a-service (Perera, Zaslavsky, Christen, & Georgakopoulos, 2014a) is important for the creation of a service-oriented sensor Web (Gibbons, Karp, Ke, Nath, & Seshan, 2003), or more generally, a service-oriented IoT (Barnaghi, Wang, Henson, & Taylor, 2012). The idea is to abstract sensor functionalities and capabilities in terms of standard service interfaces and to support uniform service operations.

Following the sensor-as-a-service paradigm, if all sensors expose their functionalities as services, there will be numerous real-world services generating a large amount of streaming data

* Corresponding author.
E-mail addresses: wei.wang03@xjtlu.edu.cn (W. Wang), s.de@surrey.ac.uk (S. De), gcassar@ureason.com (G. Cassar), k.moessner@surrey.ac.uk (K. Moessner).

http://dx.doi.org/10.1016/j.eswa.2014.11.058
0957-4174/

continuously. This highlights the significance of efficient sensor service discovery, which aims at locating relevant and quality sensor services according to users' requirements. It should be noted that in the discovery process, the users (e.g., human users or applications) usually do not have exact knowledge of the actual sensor services (e.g., where they are located, what functionalities they provide or how they work). To this end, we share similar perspectives with Delicato et al.'s view on a flexible architecture with which sensor network data can be accessed by users spread all over the world (Delicato, Pires, Pirmez, & Carmo, 2005). The discovered services can be accessed in real-time and used for many purposes, for example, reading and aggregating sensor data (Wu, Kung, Chen, & Kuo, 2014; Stavropoulos, Gottis, Vrakas, & Vlahavas, 2013), sensor data abstraction and analysis (Roda & Musulin, 2014), service composition and runtime adaptation (Coria, Castellanos-Garzón, & Corchado, 2014).

Sensor services are significantly different from standard Web services, e.g., extremely large in number, location dependent, dynamic and unreliable. Finding a particular sensor service(s) from billions according to the search criteria can be a challenging task. Most of the existing techniques developed for standard Web service discovery are not directly applicable to sensor services, for instance, it is not possible to build a centralised index or portal for all sensor services due to their large number and highly distributed nature; the location dependency implies that sensor services need to be organised in a way to allow efficient search based on spatial properties, e.g., region containment or overlap; the dynamic and unreliable nature of the sensor services implies that update operations may need to be performed frequently, which may introduce substantial overhead to the discovery platform.

Our work aims to overcome the identified difficulties and to design an efficient method for large-scale sensor service discovery based on geospatial indexing, semantic and service-oriented techniques. The design demonstrates the following features: (1) geospatial indexing to efficiently and effectively reduce search space; (2) resilience to dynamicity and reducing the number of expensive update operations; (3) more precise computation of the spatial properties of sensor service gateways using computational geometric techniques; and (4) distributed semantic repositories and semantic service matchmaking to provide more accurate results. The rest of the paper is organised as follows. In Section 2 we provide a review on some of the representative works for sensor service discovery. In Section 3 we first briefly present the background information relevant to sensor modelling and geospatial indexing, then we elaborate the design and implementation of the discovery platform, and how the identified problems are addressed. Section 4 presents the experiment and evaluation results, which are also compared to the ones generated by the benchmark methods. Section 5 concludes the paper and discusses the future work.

2. Related work

Service discovery has been extensively studied in the literature, especially for services on the Web (including semantic Web) (Coria et al., 2014; Garofalakis, Panagis, Sakkopoulos, & Tsakalidis, 2006; McIlraith, Son, & Zeng, 2001) and in pervasive and mobile environments (Chakraborty, Joshi, Yesha, & Finin, 2006; Nidd, 2001). The study in (Perera, Zaslavsky, Christen, Compton, & Georgakopoulos, 2013) identifies the similarities and differences between sensor search and Web service search, and argues that the scale of the IoT makes sensor discovery (as the precursor to sensor selection and ranking) much more challenging. Recent research on extending the service-oriented principles to the IoT domain also shows that existing methods cannot be directly applied to the vast number of distributed services exposed by the networked, real-world resources on the IoT (Barnaghi et al., 2012; Perera et al., 2014b; Wang, De, Cassar, & Moessner, 2013; Wei & Jin, 2012) Generally speaking, IoT services can be seen as a special class of Web services that have a number of distinctive characteristics, e.g., large-scale, highly distributed and dynamic, location-dependent and capability constrained.

Already there have been considerable research efforts from the IoT and service computing communities to address the identified issues. For example, the work in (Abangar et al., 2010; Evdokimov, Fabian, Kunz, & Schoenemann, 2010; Guinard, Trifa, Karnouskos, Spiess, & Savio, 2010) provides sensor based services using the Device Profile for Web Services (DPWS) (OASIS, 2009), a lightweight subset of Web service technologies that enables plug-and-play features for resource constrained devices. The work in (Guinard et al., 2010) develops a platform to facilitate discovery, selection and on-demand provisioning of real-world services for business applications. In addition to the DPWS based discovery method, the SOCRADES middleware in (Guinard et al., 2010) also implements a RESTful network discovery mechanism for devices. The work in the aWESoME middleware (Stavropoulos et al., 2013) also uses services to expose IoT devices and employs a service broker to implement the service discovery functionalities. The main limitation of the above discussed research is that with each service corresponding to a device type, the discovery methods require unique device IDs (typically MAC addresses) as input to distinguish between different devices of the same type. The authors in (Wei & Jin, 2012) highlight the resource-constrained and dynamic nature of the IoT and propose a context-aware framework for service discovery based on formal context modelling (Gu, Pung, & Zhang, 2004) and uncertain context modelling (Gu, Pung, & Zhang, 2005), in particular, their work considers the temporal dimension and applies probabilistic reasoning in dynamic Bayesian networks. However, the methods are mostly designed for environments with limited scopes (e.g., home or enterprise) and do not apply to the pervasive computing domains.

To support interoperability, recent research proposes the use of semantic Web technologies for discovery of sensor services and data in IoT environments (De et al., 2012; De, Christophe, & Moessner, 2014). The sensor data analysis framework in (Roda & Musulin, 2014) applies the Semantic Query Web Rule Language, an extension of the Semantic Web Rule Language, as the query engine language for knowledge extraction from sensor data. However, the work requires all sensor metadata to be available in a centralised registry and does not consider the location dependency characteristics of sensor deployment. The CASSARAM sensor search model (Perera et al., 2014b) utilises a weighted Euclidean distance based indexing technique to measure the similarity between the sensor description and the user requirements. A heuristic filtering and a relational expression based filtering methods are then applied to reduce the amount of metadata needed to be processed during the discovery. In contrast, our proposed method partitions the discovery space into much smaller ones based on geospatial index prior to query processing. This offline step allows the discovery to be performed in a more efficient way.

As sensor services are location dependent, geographical information plays a substantial role in discovery, especially in pervasive environments, e.g., wireless sensor networks or vehicular networks (Abrougui, Boukerche, & Ramadan, 2012; Niforatos, Karapanos, & Sioutas, 2012). The work in (Fredj, Boussard, Kofman, & Noirie, 2013) uses a hierarchy of nodes to represent indoor locations (e.g., room, building or floor). The nodes encapsulate semantic service descriptions of objects located within their geographic scope. However, the work does not consider the cost of maintaining the discovery platform which might be computationally expensive in a dynamic environment. The discovery approach in (Mayer,