Information Systems 58 (2016) 43–55

Contents lists available at ScienceDirect

Information Systems

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

A secure and quality-aware prototypical architecture for the Internet of Things

Sabrina Sicari [a,*], Alessandra Rizzardi [a], Daniele Miorandi [b], Cinzia Cappiello [c], Alberto Coen-Porisini [a]

[a] *Dipartimento di Scienze Teoriche e Applicate, Università degli Studi dell'Insubria, via Mazzini 5, 21100 Varese, Italy*
[b] *U-Hopper, via A. da Trento 8/2, 38122 Trento, Italy*
[c] *Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy*

ARTICLE INFO

Article history:
Received 14 January 2016
Received in revised form
4 February 2016
Accepted 10 February 2016
Recommended by: D. Shasha
Available online 18 February 2016

Keywords:
Internet of Things
Security
Quality
Middleware
Prototype

ABSTRACT

The increasing diffusion of services enabled by Internet of Things (IoT) technologies raises several risks associated to security and data quality. Together with the high number of heterogeneous interconnected devices, this creates scalability issues, thereby calling for a flexible middleware platform able to deal with both security threats and data quality issues in a dynamic IoT environment. In this paper a lightweight and cross-domain prototype of a distributed architecture for IoT is presented, providing minimum data caching functionality and in-memory data processing. A number of supporting algorithms for the assessment of data quality and security are presented and discussed. In the presented system, users can request services on the basis of a publish/subscribe mechanism, data from IoT devices being filtered according to users requirements in terms of security and quality. The prototype is validated in an experimental setting characterized by the usage of real-time open data feeds presenting different levels of reliability, quality and security.

1. Introduction

The Internet of Things (IoT) revolution is turning everyday objects into *smart* ones. Such *smart things* are able to interact among themselves and with the environment they are in order to fulfill a given goal [1]. As a result, a global network infrastructure is being created, supporting the provisioning of innovative and customized services to individuals and businesses in different application domains. The resulting system may include an extremely large number of heterogeneous devices, raising integration and scalability challenges to be addressed.

Security & privacy are widely acknowledged to represent critical issues in such a context [2]. On the one hand, the confidentiality and the integrity of the transmitted and stored information have to be guaranteed, and authentication and authorization mechanisms have to be provided to prevent unauthorized users or devices to improperly access the system. On the other hand, privacy of users, understood as ability to support data protection and users anonymity, has to be ensured, a critical aspect in particular in the presence of personal and/or sensitive information [3]. Beyond security, also data quality represents an essential requirement for the adoption at scale of IoT services. The provided information should be accurate, timely and complete, since, in some scenarios, errors or missing values might have critical impact on actions or decisions of

[*] Corresponding author.
E-mail addresses: sabrina.sicari@uninsubria.it (S. Sicari), alessandra.rizzardi@uninsubria.it (A. Rizzardi), daniele.miorandi@u-hopper.com (D. Miorandi), cinzia.cappiello@polimi.it (C. Cappiello), alberto.coenporisini@uninsubria.it (A. Coen-Porisini).

http://dx.doi.org/10.1016/j.is.2016.02.003
0306-4379/© 2016 Elsevier Ltd. All rights reserved.

the IoT system itself [4]. Indeed, as in IoT-enabled services and applications may make use of different data sources, the user (or the application itself) has to be aware of the security and quality level of the data being accessed, in order to take informed decisions about their usage.

As a result, what emerges is the need of a system able to deal with heterogeneous data sources and to evaluate the security and the quality of the information being collected, processed and transmitted, possibly in real-time and in an automatic manner. Furthermore, such a system shall be able to work in the absence of a priori complete knowledge of the sources themselves, since IoT environments are highly dynamic and different kinds of attack may occur. In fact, in such a scenario, data may be compromised by rogue devices, hindering the correctness and confidentiality of the information (e.g., data integrity violation, man-in-the-middle attacks, packet sniffing). Source authentication issues (e.g., compromised keys, session violation) shall also be accounted for. In order to deal with such issues, a mechanism for the assessment of data quality and security is proposed in this paper, supported by a number of novel algorithms aimed at analyzing the data sources as well as the data they generate over time. As far as security assessment is concerned, the presented algorithm is new and, according to the authors' knowledge, the first of its kind. Its aim is to assign a level of robustness to each data source according the following security features: integrity, confidentiality, authentication system, privacy. Therefore, the proposed solution does not tackle the security attacks directly, but aims at minimizing the associated risks by letting users and applications be aware of the security and data quality level.

The proposed solution is integrated in an existing IoT middleware, named *NetwOrked Smart objects* (*NOS*) [5]. NOS are conceived as computationally powerful devices connected to create a distributed processing and storage layer able to process the data acquired from large-scale IoT deployments close to the actual data sources. NOSs collect the data generated by nearby IoT devices, process them and finally transmit the processed data on a publish/subscribe broker. Such a middleware includes provisionings for users and applications to dynamically specify the level of security and data quality suitable for their own purpose. The distributed architecture automates the deployment of adequate filters for ensuring that only qualified data is being used by the actual service. This represents a clear innovation over conventional one-size-fits-all approaches, which provide the same information to all consumers, often without considering his/her requirements in terms pf security, privacy and data quality.

The algorithms integrated within the prototype are validated in an experimental setup characterized by the usage of real-time open data feeds and, contextually, assessing the resulting security and quality level. With respect to the original highly modular and lightweight prototype presented in [5], NOS functionalities are significantly improved in this work, enhancements including: (i) a set of methods for the distributed and autonomic management and run-time optimization of the middleware platform; (ii) a set of secure and privacy-aware mechanisms and their integration in the proposed platform; (iii) a set of data quality assessment methods that can be used for different types of data and in different scenarios; (iv) standardized interfaces and data models for applications/services to access qualified IoT information following a publish/subscribe architecture, where raw data are enriched with metadata specifying their security and quality levels.

The paper is organized as follows. Section 2 reviews the relevant literature and state of the art. Section 3 presents the system architecture and explains the storage and the data management aspects, as well as all the operating modules and their functions. Sections 4 and 5 present the prototype and its validation scenario, in order to demonstrate its effectiveness in a real-world IoT context. Section 6 concludes the paper providing directions for future research.

2. Related work

One of the main factors limiting the growth and take-up of IoT is the lack of a set of standardized tools, platforms and interfaces able to provide interoperability across different vendors of hardware and software solutions as well as across diverse vertical domains.

In the last few years, many initiatives tried to bridge this gap, reusing concepts, techniques and protocols from the Internet domain. For example, in recent years, the widespread adoption of web services has provided a standard framework to enable systems' interoperability according to the principles of Service Oriented Architectures (SOA). Service-oriented Communications (SOC) technologies manage web services by creating a virtual network and adapting applications to the specific needs of users rather than users being forced to adapt to the available functionality of applications [6,7]. Although the trend towards the adoption of SOA architectural principles in the IoT domain is shared by the majority of the scientific community, at the moment the state of the art in this area is still somehow limited [8,9].

Due to the very large number of heterogeneous technologies being used in IoT systems, several middleware layers have been proposed to enforce the integration and the security of devices and data within the same information network. Typically, they enforce data to be exchanged according to strict protection constraints. Heterogeneity of devices and communication technologies in IoT has to be accounted in the design of such middleware architecture. Indeed, while many smart devices can natively support IPv6 communications [10,11], existing deployments might not support the IP protocol within the local area scope, thus requiring ad hoc gateways and middlewares [12]. Recent works on IoT middlewares include VIRTUS [13], which relies on the open eXtensible Messaging and Presence Protocol (XMPP) to provide secure event-driven communications; Otsopack [14] and Naming, Addressing and Profile Server (NAPS) [15], which are data-centric frameworks based on the usage of HTTP and REpresentational State Transfer (REST) interfaces.

An attempt to provide a lightweight and flexible middleware for IoT applications is at the heart of the work reported in [5] where, starting from a general UML