Web Semantics: Science, Services and Agents on the World Wide Web 42 (2017) 19–37

Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

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

Real-time data analytics and event detection for IoT-enabled communication systems[☆]

Muhammad Intizar Ali[a,*], Naomi Ono[a], Mahedi Kaysar[a], Zia Ush Shamszaman[a], Thu-Le Pham[a], Feng Gao[a], Keith Griffin[b], Alessandra Mileo[a]

[a] Insight Centre for Data Analytics, National University of Ireland, Galway, Ireland

[b] Cisco Systems Galway, Ireland

ARTICLE INFO

Article history:
Received 16 July 2015
Received in revised form 29 June 2016
Accepted 21 July 2016
Available online 28 July 2016

Keywords:
RDF stream processing
Stream federation
Internet of Things (IoT)
Communication systems
Linked data
Stream reasoning

ABSTRACT

Enterprise Communication Systems are designed in such a way to maximise the efficiency of communication and collaboration within the enterprise. With users becoming mobile, the Internet of Things (IoT) can play a crucial role in this process, but is far from being seamlessly integrated into modern online communications. In this paper, we present a semantic infrastructure for gathering, integrating and reasoning upon heterogeneous, distributed and continuously changing data streams by means of semantic technologies and rule-based inference. Our solution exploits semantics to go beyond today's ad-hoc integration and processing of heterogeneous data sources for static and streaming data. It provides flexible and efficient processing techniques that can transform low-level data into high-level abstractions and actionable knowledge, bridging the gap between IoT and online Enterprise Communication Systems. We document the technologies used for acquisition and semantic enrichment of sensor data, continuous semantic query processing for integration and filtering, as well as stream reasoning for decision support. Our main contributions are the following, (i) we define and deploy a semantic processing pipeline for IoT-enabled Communication Systems, which builds upon existing systems for semantic data acquisition, continuous query processing and stream reasoning, detailing the implementation of each component of our framework; (ii) we present a rich semantic information model for representing and linking IoT data, social data and personal data in the Enterprise Communication scenario, by reusing and extending existing standard semantic models; (iii) we define and develop an expressive stream reasoning component as part of our framework, based on continuous query processing and non-monotonic reasoning for semantic streams, (iv) we conduct experiments to comparatively evaluate the performance of our data acquisition and semantic annotation layer based on OpenIoT, and the performance of our expressive reasoning layer in the scenario of Enterprise Communication.

1. Introduction

Enterprise communication systems have historically been primarily aimed at person-to-person communication. Users of such systems typically interact with an endpoint such as a phone, video system or unified communications software client capable of multi-modal communications. Communication modes typically consist of instant messaging, voice, video and voicemail to allow individuals or groups to communicate in real time. A few solutions providing support for machine-to-machine communication are currently on the market. These solutions help users in their search for information, but they are mainly based on Natural Language Processing techniques to interpret verbal requests and search-and-index strategies to provide fast answers.

Advances in the Internet of Things (IoT) have great potential in enabling communication between sensory devices and communication systems using open interfaces, but this potential is under-investigated and few solutions have existed in isolation. As a result, the flexible integration of a large amount of multi-modal

☆ This research is sponsored by Science Foundation Ireland (SFI) Grant No. SFI/12/RC/2289 and Cisco Systems (Grant No. RIB1144).

* Corresponding author.
E-mail addresses: ali.intizar@insight-centre.org (M.I. Ali), naomi.ono@insight-centre.org (N. Ono), mahedi.kaysar@insight-centre.org (M. Kaysar), zia.shamszaman@insight-centre.org (Z.U. Shamszaman), thule.pham@insight-centre.org (T.-L. Pham), feng.gao@insight-centre.org (F. Gao), kegriffi@cisco.com (K. Griffin), alessandra.mileo@insight-centre.org (A. Mileo).

http://dx.doi.org/10.1016/j.websem.2016.07.001
1570-8268/

data streams from diverse application domains is still one of the key challenges in developing IoT-enabled communication systems.

In addition, limited interoperability results in the inability of such systems to integrate information from external sources in an easy and cost-effective way. This issue becomes more evident if we consider advances in the IoT space, which demands a dynamic and flexible exchange of information between IoT sources. To overcome these interoperability issues in communication systems and across smart enterprise applications towards IoT-enabled solutions, we developed a Linked Data infrastructure for networking, managing and analysing streaming information. In order to ensure high reusability, we leveraged existing semantic models for the annotation of sensor data (e.g. SSN), social web (e.g. FOAF) and personal information (e.g. PIMO), and extended the ontological model to incorporate personal, business and online communication concepts.

In order to set the basis for our evaluation, we have identified our target usecase scenarios in the enterprise communication space, to illustrate the potentials of IoT-enabled Communication Systems, and we have designed, developed and tested a semantic processing pipeline from IoT sources to stream processing and complex reasoning, which is seamlessly integrated into our framework. Our main contributions in this paper include:

- design and development of a Linked Data framework for IoT-enabled smart enterprise applications, which connects heterogeneous streaming sources and enables scalable stream processing and reasoning;
- specification of a rich semantic information model which enhances interoperability;
- demonstration of the effectiveness of the entire processing pipeline in our framework, in particular highlighting the novelty of our declarative stream reasoning component aimed at decision support;
- experimental evaluation of our framework, with details related to the performance and scalability of data acquisition, semantic annotation and stream reasoning component, based on a concrete instance of OpenIoT and Apache Open Meetings.

The remainder of this paper is organised as follows: Section 2 details our scenarios and motivation; state of the art in Enterprise Communication Systems in general, as well as related work on acquisition and processing of semantic streams is presented in Section 3; Section 4 presents our IoT-enabled Linked Data infrastructure and its components, providing details of each processing layer; implementation details are presented in Section 5; we evaluate our framework in Section 6 and conclude with a list of interesting directions for future work in Section 7.

2. Motivation and usecase scenarios

Sensor technologies and sensory devices are nowadays part of our everyday lives. The Internet of Things (IoT) not only provides an infrastructure for sensor deployment, but also a mechanism for better communication among connected sensors. Data generated by these sensors is huge in size and continuously produced at a high rate. This requires mechanisms for continuous analysis in real-time in order to build better applications and services. Data streams produced by various sensors can be classified into three different categories, namely, (i) *Physical (static) Sensors*, (ii) *Mobile* & *Wearable Sensors*, and (iii) *Virtual Sensors* & *Social Media Streams*.

Among the above three categories, mobile sensors are harder to integrate within enterprise communication systems. This is not only due to technical integration issues and interoperability but also due to their dynamic nature and constantly changing context. Mobility and location-based sensory input, for example, result in a higher level of unpredictability and a lower level of control over the distributed infrastructure that characterises enterprise communication systems. These challenges are matched by new opportunities for IoT-enabled collaboration and communication systems to be designed in order to sense the context of a mobile user and take decisions according to dynamic sensory input. In the domain of enterprise communication systems, mobile users have the potential to produce a lot of dynamic sensory input that can be used for the next generation of mobile enterprise collaboration, with great potentials for better user experience. In this paper, we propose a framework and a set of software components for IoT-enabled online meeting management which combines semantic technologies and declarative rule-based reasoning in a scalable processing infrastructure.

2.1. Motivating scenario: IoT-enabled meeting management system

Alice is hosting an online meeting for her company *FictionDynamic*. The meeting is planned to be held in Meeting Room B at 11:00 am. Bob and Charlie will be attending the meeting while they are on the move, thus their availability and ability to participate in the meeting in various ways are dynamically changing. The IoT-enabled Meeting Management System (IoT-MMS) enables (i) automatic on-the-fly semantic enrichment of IoT information related to the meeting attendees, (ii) communication of such richer information to the participants via their IoT-MMS client through a panel showing IoT values and related user capabilities (e.g. ability to hear properly, share a screen, type, talk), (iii) use of such rich information to improve user experience and optimise meeting management on-the-fly. The integration of a web-based MMS with sensory input and enterprise data such as attendees details, calendars and agenda items makes it possible to characterise and manage the following aspects in a flexible and interoperable way:

- updating (enabling or disabling) users capabilities based on IoT input (via sensors' virtualization and interpretation, semantic integration and stream query processing);
- managing agenda items, including users assigned to a particular item and capability requirements for that item via declarative logic rules;
- dynamically verifying privacy constraints on agenda items based on location and context.

Contextual information can be explicitly available as the user specifies whether he/she is in a public or private place, but they can also be detected by specific simple event detection logic via query processing in our framework. However, the ability to reason about constraints and planning (e.g. for re-scheduling agenda items) needs to be handled by more expressive rules, and are handled by our stream reasoning component. In order to further illustrate the multiplicity of situations that can occur, we characterise a few instances of the motivating scenario in what follows.

Capability-aware participation. Charles is on the move and gets notified about a last-minute meeting to be held with a customer on a specific product with the development team. As the head of the team, Charles needs to be mainly a listener and intervene only if needed. Information about Charles capabilities as a listener (e.g. level of attention while driving or while in a noisy area) are collected and interpreted through inertial and environmental sensors via his Android phone. Such capabilities are continuously monitored and updated in the IoT-MMS web clients for the meeting host to see, so that when the customer is addressing Charles directly, he knows whether the issue should be answered right away or later. Charles is also notified when his intervention is requested and what capabilities are needed to participate (e.g. quieter area, ability to type and so on). All this information is stored and associated to a synthetic representation of the meeting minutes for future reference.