Advanced Engineering Informatics 30 (2016) 537–552

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Hybrid approach for selective delivery of information streams in data-intensive monitoring systems

Yulia Evchina *, Jose Luis Martinez Lastra

FAST Laboratory, Tampere University of Technology, P.O. Box 600, FIN-33101 Tampere, Finland

ARTICLE INFO

Article history:
Received 4 February 2016
Received in revised form 21 June 2016
Accepted 20 July 2016

Keywords:
Complex Event Processing (CEP)
Information filtering
Monitoring systems
Semantic Web (SW) technologies

ABSTRACT

Background: Present day control and monitoring systems are equipped with a large number of heterogeneous devices and are operated by many users with different roles and responsibilities. The information generated by these devices, although preprocessed and filtered, is usually delivered to users regardless of their actual information needs, thus overwhelming cognitive capacities and potentially affecting safety of the system.
Objectives: This work aims to reduce information load to the users of the data-intensive monitoring systems by delivering selected information to each user based on his/her roles in the system and responsibilities.
Methods: The proposed approach combines Semantic Web (SW) technologies and Complex Event Processing (CEP) for configuration purposes and run-time analyzing. The approach is exemplified with implemented tools and feasibility study based on the performance tests. The paper describes principles of proposed approach, demonstrates illustrative scenarios from building automation domain, gives description of implemented tools, and presents results of the initial performance tests.
Results: The combination of SW and CEP brings two major advantages: (1) the behavior of the system could be easily changed by configuring only underlying ontology and (2) utilization of CEP at runtime makes system event-driven and reactive to frequent changes in the environment. The performance tests demonstrated the response time of implemented tools within one second for 1000 updates per second (which corresponded to 10,000 devices in the performed experiments).
Conclusions: It is expected that the proposed approach is able to make monitoring systems personal oriented and thus safer during the operation. The results of the performance tests suggest feasibility of the approach for such systems as building and home automation, and non-critical industrial automation.

1. Introduction

Present day control and monitoring systems are equipped with multifarious devices, providing information about different aspects of the environment. The problem arises, when this information, although preprocessed and filtered, is fully delivered to users of the system, overwhelming cognitive capacities and potentially affecting safety of the system. The information overload is relevant in several domains of the monitoring systems, including process automation, discrete manufacturing and building automation (BA) systems. In the process automation industry around 30–90% of system failures are attributed to operator errors, most of which appear due to information overload [1–3]; while discrete manufacturing and BA systems operators suffer from the need of monitoring tens of thousands devices and addressing arising problems in the timely manner [4,5]. This paper addresses the problem of selective information delivery based on user role in the system and their information needs and proposes hybrid approach combining Semantic Web (SW) technologies and Complex Event Processing (CEP) for solving it.

The evolution of ICT has opened new possibilities for control and monitoring systems. Ethernet and IP-based networks have enabled easier integration of various devices, thus, systems are capable to deliver almost any information about environment, process of interest or even operators and other actors of the system [6]. The availability and diversity of information have made systems attractive to more users and user groups, who now have the possibility to find specific information assisting in their daily tasks. However, most of the monitoring systems are not capable of tailoring information flows to user needs and requirements.

* Corresponding author.
E-mail addresses: yulia.evchina@tut.fi (Y. Evchina), jose.lastra@tut.fi (J.L. Martinez Lastra).

http://dx.doi.org/10.1016/j.aei.2016.07.006
1474-0346/

Instead, almost all available information, although preprocessed and filtered on a field level, is delivered to each and every user regardless of their roles and responsibilities. In the best case scenario, monitoring systems have configuration abilities to control information flow for different user groups, but they do not consider individual information needs and are not able to address some short term requirements in information delivery, such as temporal reallocation of duties (e.g. due to someone's sick leave) or delivering of specific information to visiting workers (e.g. maintenance provided by outside companies). As a result, such systems create high information loads to the users, overwhelming cognitive capacities and complicating daily routine [6,7]. Therefore, monitoring systems require new ways of handling information streams, which would enable selective information delivery with (re)configuration capabilities.

This paper deals with selective information delivery in data-intensive monitoring systems. In this paper, the term "data-intensive" refers to the systems with large number of heterogeneous devices each producing at least one information stream, rather than devices producing highly intensive information streams. Due to large number of devices and their variety, such systems have additional properties related to information streams: (1) the variety of information about various aspects of the environment makes these systems potentially useful for different users and user groups, therefore the systems have multiple users with various information needs and (2) the large numbers of devices and users create more potential situations, when something has to be changed in the system (i.e. new requirements from existing or new users, new devices, etc.), therefore the requirements in delivery of information streams might change over time reflecting needs of the users or changed properties of devices. Thus, the problem of selective information delivery in data-intensive monitoring system could be formulated as delivery of multiple information streams to multiple users with respect to individual information needs and possible changes within these environments and/or user information needs.

This work presents a middleware capable to deliver information tailored to user needs based on his/her role and responsibilities. The approach combines SW technologies and CEP. The combination of mentioned technologies brings two major advantages: (1) the behavior of the middleware could be easily changed by configuring only underlying ontology and (2) utilization of CEP at runtime makes middleware event-driven and reactive to changes in the environment. The original contributions of the paper could be summarized as follows: (1) the proposed methodology for selective information delivery, featuring easier system reconfiguration, is described in detail with the guidelines for adopting it for related domains and applications; (2) the architecture and corresponding software implementation supporting proposed methodology are presented; and (3) the implemented software is evaluated with performance tests, demonstrating feasibility of the proposed methodology for the environments with numerous devices.

The rest of the paper is organized as follows. Section 2 presents a brief overview of related work and justifies taken approach. Section 3 presents detailed principles of combination of SW technologies and CEP. Section 4 describes functional architecture of the middleware, which is followed by Section 5 describing implementation. Section 6 presents examples of reconfiguration scenarios from a building automation (BA) domain. Section 7 presents the results of empirical evaluation of suggested approach. Finally, Section 8 concludes the paper and highlights directions for future work.

2. Related work

The problem of selective information delivery, when selection should be done out of many options, is relevant in many application domains. These examples include searching information in the Internet, medical diagnostic applications, delivery of information and/or services depending on context, etc. Thus, the approaches to solve the problem are numerous, and the right choice depends on application requirements and context of use. The specific features of data-intensive monitoring systems are (1) the large number and variety of devices in the system and its operators, (2) the run-time flow of information from the environment devices, which should be delivered in timely manner and (3) the possible alterations in the system due to changing rights or duties of users, adding/configuring new devices, etc. Based on these criteria we have highlighted five main approaches, which might be used in selective information delivery: SW technologies, publish/-subscribe middleware, CEP, Linked Stream Data, and hybrid solutions.

2.1. Semantic Web technologies

Semantic Web technologies are the family of W3C recommendation standards for describing and relating data on the Web and inside enterprises. Among these standards, the most popular are OWL[1] ontologies and SPARQL[2] queries. Although SW technologies were originally developed for describing and relating static information on the web, it is possible to use it within dynamic systems with additional technologies such as SPARQL/Update[3] allowing ontologies to reflect present state of the environment, and SWRL[4] rules for expressing conditional rules within OWL languages. The advantages of using Semantic Web technologies are in expressivity of knowledge models, which are human and machine readable. This prompts easier reconfiguration process, when changes made by human could be instantly interpreted by machine. Usually, systems based on Semantic Web technologies model situations of interest into an ontology, which describes and links various concepts together, and query this model for specific information, which is used for changing behavior of the system. The examples of such applications are numerous: home health monitoring [8], industrial applications [9–12], road traffic management [13], and delivering personalized services [14–17] to name a few.

SW technologies is a powerful tool for expressing concepts and their relationships, however there are limitations in temporal reasoning support, especially within context of dynamic information flows. It is a challenge to express temporal relationships among events such as "*before*", "*after*", and "*overlaps*". There are attempts to overcome these challenges by using extensions of SWRL language as in [18] or by adding additional models or reasoners on top of the ontological model as in [13,19,20]; however, the application is still limited and there is no standardized approach to reason with time constraints.

2.2. Publish/subscribe middleware

The publish/subscribe middleware are solutions providing asynchronous communication between multiple entities of the system. The approach is especially popular in Wireless Sensor Networks, which consist of multiple communicating devices and require efficient ways of communication [21–23]. The publishers produce information to the consumers (subscribers). Usually, publishers and subscribers are unaware of each other, i.e. publisher does not know to whom it produces information, while consumer is unaware from where the piece of information came. The

[1] Web Ontology Language (OWL): http://www.w3.org/2004/OWL/.
[2] SPARQL Protocol and RDF Query Language: http://www.w3.org/wiki/SPARQL.
[3] SPARQL 1.1 Update, update language for RDF graphs: http://www.w3.org/TR/sparql11-update/.
[4] Semantic Web Rule Language: http://www.w3.org/Submission/SWRL/.