Information Fusion 21 (2015) 42–56

Contents lists available at SciVerse ScienceDirect

Information Fusion

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

# High-performance scheduling model for multisensor gateway of cloud sensor system-based smart-living

Yongqiang Lyu [a,*], Fanxi Yan [b], Yu Chen [a], Dan Wang [b], Yuanchun Shi [a], Nazim Agoulmine [c]

[a] OS Lab, Research Institute of Information Technology, Tsinghua University, No. 1 Tsinghua Yuan, Haidian District, Beijing 10084, China
[b] Department of Computer Science and Technology, Beijing University of Technology, No. 100 Pingleyuan, Chaoyang District, Beijing 100022, China
[c] IBISC Laboratory, University of Evry, 1 Rue Ambroise Croizat, 91000 Évry, Paris, France

ARTICLE INFO

*Article history:*
Available online 28 April 2013

*Keywords:*
Cyber physical system
Internet of things
Cloud sensor system
Multisensor fusion
Periodic application

ABSTRACT

In cloud sensor system-based smart-living applications, large-scale distributed sensors may be deployed to collect information and report to the manipulator and the cloud. A gateway is often employed as infrastructure in this scenario, acting as the data collector, the relay and the agency of the initial multisensor fusion, and thus must be able to handle as many concurrent requests as possible from diverse sensors of different vendors. This study proposes a high-performance scheduling model with a cloud-supported caching mechanism for the gateway of the cloud sensor system-based smart-living. Scheduling and caching optimization are performed by 0–1 programming combined with the periodic task models. Correlation analyses of the simulated results determine the most effective factors to the performance, and the performance tests with the selected factors show that a gateway with 2.4 G-uniCPU/4 G-memory/300 G-harddisk can support the system with one million sensors registered in the cloud and 5000 concurrent live sensors through it, with a 25× gain throughput improvement compared to the traditional application-type scheduling.



## 1. Introduction

Ever-increasing multisensor data fusion technologies have been promoting advanced data manipulation, processing and analysis in areas of versatility and diverse applications. There have been fruitful algorithms and methodologies proposed in the past decades [1]. As support to information fusion, there are also several well-known and frequently used architectures proposed in the domain of the system architecture of sensor networks, e.g., star networks, multi-hop networks [2,3], mobile sensor networks [4], ad hoc networks [5,6], multilevel/hybrid networks [7] and distributed sensor networks [8,9].

Smart-living is an emerging application of sensor networks and information fusion. It often involves the services of automatic control, smart healthcare, comfortable environment, green energy, intelligent security and safety, etc. Based upon the low-level sensor and device networks, upper-level applications through information fusion, context awareness and reasoning can deliver smart services to end users to validate the smart-living. For example, in a scenario of smart healthcare, the user with cardiovascular disease may wear some biological sensors to monitor the vital signs such as the pulse rate and blood pressure; when the blood pressure and pulse rate are detected unstable, the smart healthcare service may notify and stop the user when it happens to detect heavy physical activities (such as doing exercise) via the activity sensors or the noisy environment via the environmental sensors in the meanwhile.

Moreover, emerging cloud systems are also providing more intelligent, flexible and scalable services to incorporate into the smart-living, especially when using sensor networks in the concept of the internet of things (IoT) or cyber physical systems (CPS). More often, business opportunities are believed to be created under a scenario of IoT/CPS plus the cloud. Many of these opportunities are found in end user-oriented cloud applications such as smart living (e.g., smart home/city and smart healthcare), vehicular networks [10] and customized location based services (LBS) [11]. From the perspective of system architecture, in most cases service providers must deploy their full sensor networks, including communication devices/modules and data collection adapters, to validate the data paths for their sensors, also providing corresponding automatic sensor data backend services on the cloud.

A *gateway* is often considered a good infrastructure model for implementing stable data outlets and flexible organizing nodes in sensor networks. From a networking perspective, a gateway is often used to interconnect different networks with different proto-

* Corresponding author. Tel.: +86 10 62793738.
*E-mail addresses:* luyq@tsinghua.edu.cn (Y. Lyu), fancylanlan@emails.bjut.edu.cn (F. Yan), chyyuu@gmail.com (Y. Chen), wangdan@bjut.edu.cn (D. Wang), shiyc@tsinghua.edu.cn (Y. Shi), nazim.agoulmine@ufrst.univ-evry.fr (N. Agoulmine).

1566-2535/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.inffus.2013.04.004

cols, or setup/maintain the data links as a router. From the system perspective, the gateway is often regarded as a relatively strong computing node for data manipulation and application management. Particularly in sensor-based systems (SBS), the gateway plays a very important role in communication, data manipulation and interaction with other networks and systems; it has been widely used in some emerging applications, e.g., the smart-living, vehicular networks, mobile networks, and ubiquitous healthcare.

Recently, an architectural change using gateway technology for cloud-supported sensor networks has emerged. The applications for different sensors are stored on the cloud, and the gateway downloads and runs them dynamically at the sensors' request [12,13]. The applications are offered by the sensor vendors and designed to handle the sensor data between the sensors and the cloud. Such cloud-based sensor systems show promising scalability and flexibility, which can greatly improve the system capacity, lower the power and reduce the cost, especially in the infrastructure for smart-living, smart healthcare or smart transportation. It is also possible to use the diverse sensors in "plug-and-play" mode, anywhere, anytime. However, the key issue of the gateway-based cloud sensor systems is the performance constrained by the scalable throughput requirement.

The gateway will miss any requests from the new sensors if the memory or CPU is busy processing the tasks (sensor applications) from the current sensors. For traditional real-time scheduling problems of operating systems and distributed systems, the periodic task model [14] and periodic resource model [15] are often employed; the earliest deadline first (EDF) based algorithms [14,16] and rate monotonic (RM) based algorithms [14,17] are also well-formulated and verified in real-time task scheduling. Hierarchical scheduling [15] combined with the periodic resource model is also well-studied to handle the problems of multiple service providers and complex application development. Various scheduling algorithms [18–21] have also been proposed for one processor, multiple processors, distributed computing and cloud computing. However, a dedicated scheduling model is still absent for multisensor-fusion architecture especially when there is an infrastructural fusion-gateway to coordinate both the sensor connections and applications in a common environment, e.g. a smart-living space.

Dedicated research, focused on the high-performance multisensor gateways, is needed, to validate the powerful sensor fusion applications in smart-living spaces using the architectures driven by the cloud, especially with respect to the following issues:

- The sensors in smart-living spaces may be offered by different vendors and employed by multiple users through diverse services offered by other service vendors, i.e., at least three parties are involved in the scenario.
- Sensors need their corresponding applications on the gateway to support the data link and initial manipulation. However, the work load of the sensor application (task) may be transient and not determined either before or after scheduling. In addition, it is difficult to formulate since the applications may be offered by different vendors and some may even be protected.
- The sensor applications to be scheduled may be continuous and non-ending in some special scenarios, e.g., health monitoring.
- The gateway is open to transient sensors, and scalability of the throughput is required to handle more concurrent sensors to guarantee the usability.
- The gateway is supported by the cloud and cannot load all the applications locally. An efficient caching scheme is required to improve the performance.
- The general resource requirement of such a gateway must be considered as a constraint to eliminate the impact to the cost while keeping the performance.

This study proposes a periodic-task based high-performance scheduling and caching architecture using 0–1 programming for the multisensor gateway, addressing the above issues and contributing in:

- The gateway can realize the infrastructural support to the multisensor-fusion scenarios with multiple sensor vendors, multiple users and multiple service providers through application scheduling in an open smart-living space.
- The periodic model is organized in the granularity of the sensor applications. Supported by I/O buffering, the sensor applications can function periodically, making the throughput scalable for arbitrary sensors despite many applications.
- The scheduling and caching problems formulated as minimizing the waiting time of tasks are simplified by 0–1 programming, which has been well-studied in related fields and can be considered mathematically scalable.
- The overhead of the task preempt is modeled by the setup/release time of the working waveform and considered in the formulation.
- The general resource requirement of the gateway is characterized via the simulation for the scale of the cloud-sensor system.

The experimental results show that the proposed model can achieve a 25× throughput improvement compared to the application-type flat scheduling without the model. A general gateway with 2.4 G uniCPU, 4 G memory and 300 G hard disk can support up to one million registered sensors in the cloud and 5000 concurrent sensors through it. The remainder of this paper is organized as follows. Section 2 introduces the related work and Section 3 gives the application scenarios and the gateway architecture proposed in this study; Section 4 gives the scheduling and cache formulation of the proposed architecture; Section 5 describes the high-performance scheduling and caching solver; Section 6 presents the experimental results by simulation and Section 7 gives the conclusions and future work.

## 2. Related work

### 2.1. Gateway

The gateway is often employed in both wireless and wired networks. Its main roles lie in: (1) the communication node including organizing and routing the networks and (2) the computing node for local task processing. In the architectures of real systems, the gateway may take the responsibilities of the both two or either of them.

#### 2.1.1. Communication node

From the perspective of communication node, the gateway often plays two roles: (1) the network organizer and (2) the router. As the network organizer, it is often used to organize a local network or interconnect different networks with different protocols. As the router, it is often used to setup/maintain the data links especially between the hierarchical networks. In this domain, the primary hot topics include the wireless-spectrum coverage [22] which optimizes the wireless coverage of the gateway to support wider bands of the radio frequencies, the multi-domain interconnection and optimization [23,24] which studies the networking over multiple-protocol networks through the gateway, the smart routing [6,25] which studies the high-performance routing strategies for both the topological wireless networks and ad hoc networks, and the gateway placement optimization [26,27] which optimizes the gateway nodes for the ad hoc networks.