2015 IEEE/CIC International Conference on Communications in China - Workshops (CIC/ICCC)

An Insight into Cloud-enabled Complex Industrial Applications

Zhaogang Shu
College of Computer and Information,
Fujian Agriculture and Forestry University,
Fuzhou, China
e-mail: zhaogang.shu@gmail.com

Jiafu Wan*
School of Mechanical & Automotive Engineering,
South China University of Technology,
Guangzhou, China
*Corresponding author, e-mail: jiafuwan_76@163.com

***Abstract*—The cloud-based systems can provide massive storage resources and low-cost computing as well as the flexibility of customizing the operating environment to Complex Industrial Applications (CIA). Also the Cloud-integrated CPS (CCPS) will bring hope to allow previously unachievable application scenarios to be built, deployed, managed and controlled effectively. In this paper, we propose a cloud-integrated CPS architecture and outline the enabling technologies for CIA. Then, we dissect three potential challenges and solutions from the perspective of CIA, including the virtualized resource management techniques, the scheduling of cloud resources, and life cycle management. We hope this paper can provide the insight and the roadmap for future research efforts in the emerging field of CCPS.**

***Index Terms*-Complex Industrial Applications; Cloud Computing; Cyber-Physical Systems; Industry 4.0**

I. INTRODUCTION

As we know, Cyber-Physical Systems (CPS) are the reliable and evolvable networked time-sensitive computational systems integrated with physical processes, which is being widely used in many critical areas, such as manufacturing, and traffic control and safety [1, 2]. Recently, CPS has made great strides in some aspects (e.g., safety and security [3, 4], abstraction and verification [5], modeling [6], energy efficiency [7], and control [8]), but seem to require more attention in others. For example, it is very necessary that how to design an innovative methodology for closely integrating the CPS with cloud computing meets the some specific domains, such the Complex Industrial Applications (CIA).

Since cloud computing has the ability to provide a flexible stack of massive computing, storage and software services in a scalable and low-cost manner, it has been widely used [9, 10]. However, it is generally known that the majority of current cloud systems and the corresponding techniques primarily focus their attentions on the Internet-based applications. The CIA brings in grand challenges to cloud computing since they are significantly distinguishable from those service-oriented Internet-based applications due to their inherent characteristics (e.g., workload variations, process control, Life Cycle Management (LCM), resource requirements, and environment configurations).

With the advances in embedded design, wireless sensor networks, mobile computing and big data, it is an inevitable trend for designing the large-scale complicated system by integrating CPS with cloud computing, and CIA is not an exception. In my view, Cloud-integrated CPS (CCPS) refers to virtually representing physical system components such as sensors, actuators, robots and other devices in clouds, accessing (e. g., monitoring, actuating and navigating) those physical components through their virtual representations, and processing/managing/controlling the large amount of data collected from physical components in clouds in a scalable, real-time, efficient, and reliable manner. Particularly, integrating cloud computing techniques (e.g., virtualization, and elastic re-configuration) with CPS techniques (e.g., real-time scheduling, and adaptive resource management and control) will allow previously unachievable systems such as Cloud-integrated Manufacturing CPS (CMCPS) and cloud-integrated internet of vehicles to be deployed effectively.

This research work is partially supported by the National Natural Science Foundation of China (Nos. 61262013, 61572220), the Fundamental Research Funds for the Central Universities (No. x2jq-D2154120), and Fujian Education Department Foundation for Youth Teachers (No. K8014026A)

In this paper, we explore the simplified architecture and the enabling technologies in the design of CCPS for CIA, and review the several challenges and give the potential solutions to improve the QoS of the related CCPS. We highlight the insights and contributions as follows:

- The simplified architecture of cloud and CPS: By incorporating the dynamic interactions between cloud and CPS, we propose a CCPS Architecture (termed CCPSA) to provide flexible services and applications for CIA.
- The challenges and solutions of CCPS for CIA: We carefully choose the challenges from the perspective of CIA and give possible solutions, including the virtualized resource management techniques, the scheduling of cloud resources, and LCM.

The remainder of the paper is organized as follows. In Section II, we described the proposed CCPSA, analyze the architecture and enabling technologies. Section III concentrates upon how the key challenges have been handled, and Section IV concludes this paper and gives the outlook.

II. CCPS FOR CIA: AN INEVITABLE TREND

In this section, we will give a cloud-enabled CPS Architecture (termed CCPSA) and review the key enabling technologies for CIA.

A. Architecture Description

As the key technological enabler, CCPS can be divided into three different domains: network-centric CPS, cloud-centric CPS, and data-centric CPS, corresponding to communications,

978-1-4673-9701-8/15/$31.00 ©2015 IEEE

management/control, and computation requirements of industrial development and deployment. Figure 1 shows the proposed CCPSA for CIA.

The key to the success of this new mode is to integrate effectively social demands and production capacity in real time. Therefore, it is imperative to combine the two emerging fields of social computing and industrial production, and to connect seamlessly Internet and logistic networks with industrial robot and 3D printing based manufacturing networks, so that customers can also participate fully in the whole life cycle of production processes. In Figure 1, the purple dotted line shows the process from social requirement towards product.

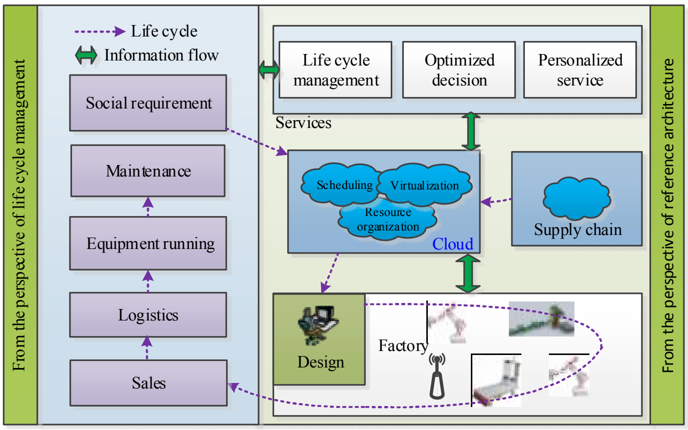

Figure 1. Proposed CCPSA for CIA

From the perspective of architecture, intelligent factory can autonomously implement the productive process with the support of new devices and technologies, such as industrial robot, Industrial Wireless Networks (IWN), and advanced coordination mechanisms. For example, in order to reduce the energy consumed by a vehicle body assembly line, the idle robots can be powered down during breaks in production by sensing the status each other. Also, the flexible assembly line can meet the individual customer requirements. In the manufacturing process, all the data (e.g., device status, product information, and raw material information) can be forwarded to industrial cloud for carrying out data analysis and then providing a personalized service.

As mentioned, since CIA has the especially obvious features (e.g., workload variations, process control, LCM, and reliability maintenance), the current cloud systems are very difficult to meet the application requirement of CIA. To address these problems, we should break through the efficient virtualization techniques to improve the resource management, feature-aware multi- dimensional resource scheduling, etc. By achieving a major breakthrough in theories, methods and techniques, we should establish the resource management and scheduling systems targeting CIA in clouds. Further, this paper adopts a typical CIA (e.g., intelligent manufacturing) to effectively evaluate our research results. In the application layer, some innovative services (e.g., LCM, see Figure 3) and decision (e.g., production guidance) are also provided by designing the big data-based applications.

B. The Enabling Technologies for CIA

As pointed out in [10], the revolution of CIA is not necessarily the technical realization but the new horizon of business models, services, and individualized products, so the enabling technologies of CIA are very important to strengthen the functionality and performance of the whole CIA system. The novel designs of the CIA system are challenging, which require the significant optimization, control, or even reconstruction of all layers and components in the system, as stated in [11, 12]. Therefore, the various research studies focus on the specific aspects of the CIA, which are classified by the following seven components, including Embedded Control Systems (ECS), IWN, AGV & 3D printing, industrial cloud, industrial big data, social computing, and system integration and optimization decision, as illustrated in Table 1 (since we carry out the research in this paper from the perspective of information technology, the AGV & 3D printing are not included in the table).

TABLE 1. THE SUMMARIZATION ON THE ENABLING TECHNIQUES OF ALL COMPONENTS OF THE CIA SYSTEM [13, 14,15,16]

Components	Topics	Future research targeting on CCPS for CIA
Embedded control systems	Model-based design	Interactions and co-design of cyber- and physical sub-systems
	Resource management	Collaborative energy-saving design of hardware and software
	Integrated control & scheduling	Efficient integration among control, communication, and computing
Industrial wireless network	Standards	Emerging standards and protocols (e.g. IEEE 802.15.4, 6LoWPAN)
	MAC protocols	Energy efficiency of IWN
	Routing protocols	Optimization and control of IWN
Industrial cloud	Scheduling & management	Scheduling of the cloud resources for industrial applications
	Virtualization	Virtualized resource management techniques for industrial applications
	Security	Data security (e.g., transmission, processing and storage)
Industrial big data	Methodologies	Data-driven control, data visualization, and real-time data retrieval methods
	Platforms & applications	Case studies, such as prototype platform
Social computing	Theories	Data clustering and classification targeting on CIA
	Typical platforms	Case studies, such as prototype platform
System integration & optimization decision	Coordination mechanisms	Formal methods for interaction behaviors among devices
	Distributed network control	Coordination control of multi-agent systems with various constraints

III. CHALLENGES AND POSSIBLE SOLUTIONS FOR CIA

Though other issues such as security and privacy are equally important and need to be addressed separately, we focus on three aspects (virtualized resource management techniques, scheduling of cloud resources for CIA) to state the challenges and possible solutions.

A. Virtualized Resource Management Techniques

Compared to the traditional High Performance Computing (HPC), cloud computing has the unique advantage on satisfying the requirements for CIA in cost reductions, high maintainability and so on, but it is still hard to get ideal performance because of the particularity of the CIA.

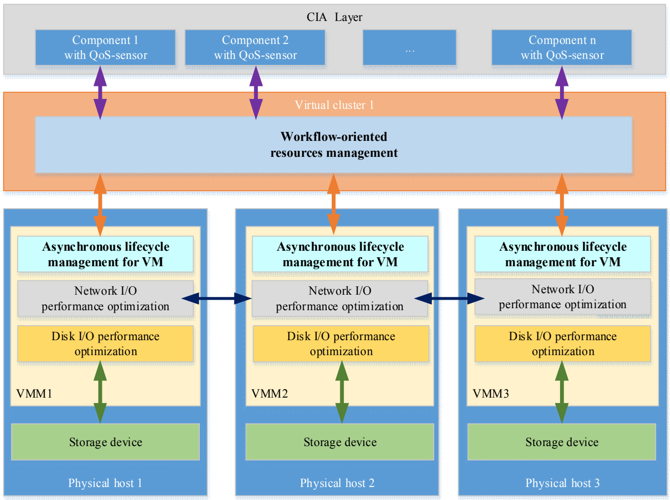

Figure 2. Virtualized resources management framework for CIA in cloud environment

To satisfy the new requirements (e.g., heterogeneous computing, large amounts of synergy nodes, and low communication latency between virtual nodes) for CIA in cloud environment, we proposed a virtualized resource management framework, which adopts a set of vector to describe computing performance, memory capacity, communication bandwidth and I/O performance, instead of the traditional parameters, such as CPU frequency to mark computing power. With the support of vector, the framework can conduct a comprehensive evaluation for all virtualized resources. In my view, the virtualized resource management technologies targeting on CIA include four aspects: workflow-oriented virtual cluster management technology, disk I/O optimization techniques in data-intensive environment, network I/O optimization techniques in communication-intensive environment, and virtual machine asynchronous lifecycle management mechanism. Figure 2 shows the virtualized resources management framework for CIA in cloud environment. In the following, we will analyze how to carry out virtualized resource management for CIA.

1) Workflow-oriented virtual cluster resources management

The resource requirements of components of CIA may be analyzed in the workflow diagram according to the characteristics of each component, fully considering performance loss across physical host, and setting the scale of required resources through heuristic optimization algorithm. When the resources scale is small, the optimal configuration can be got through genetic algorithm, simulated annealing, or ant colony algorithm. For larger resources scale, the affinity and dependence of all components must be analyzed in the workflow diagram according to the resource feature vector. For example, in order to get the mapping from Virtual CPU to Physical CPU, fuzzy mathematics may be used to get the fuzzy definition of computing capacity requirements, then take the efficient heuristic algorithm, and finally produce the reasonable computing resources scale. Based on the scale, the collaborative scheduling mechanism can be designed for the virtual CPU across physical host. The general steps are as follows: *1) analyze the source of communication latency in virtual cluster; 2) prove the correctness of collaborative scheduling in theory; and 3) implement the collaborative scheduling algorithm.*

2) Asynchronous lifecycle management of virtual machine resources

The virtual machine lifecycle management includes the creation, startup, configuration, optimization, preservation, restoration, cloning, close and other activities of virtual machines. Combined with the virtual machine template technology, the virtual machine lifecycle management can cooperate with the efficient shared storage or store a snapshot in the cloud environment, which can effectively create virtual machine quickly. Through virtual machine clone technology and virtual machine memory recovery technology, the problem of launching a large number of virtual machines in a short time can be solved. The problem for destruction of virtual machine is relatively simple, but it is necessary to guarantee that virtual machine persistent storage must be protected.

3) Disk I/O resources optimization in data-intensive environment

Usually, there are many intermediate files to be produced in the workflow of data-intensive CIA, and these intermediate files may be stored on disk through virtual machine I/O interface. In order to improve the I/O performance, it is necessary to establish virtual machine I/O access prediction model through analyzing the characteristic of virtual machine disk access, but the premise is not to change the transparency and encapsulation characters of virtual machines. Based on the predict model, disk I/O scheduling algorithm can be optimized. Therefore, architecture of disk I/O optimization can be divided into two parts: virtual machines I/O prediction, and scheduling of multidimensional disk resources. The prediction model of virtual machines I/O determines the service status transformation and waiting time of disk head according to the virtual machine information. Considering the frequent change of the load for data-intensive application, the nonlinear feedback model can be adopted to guarantee I/O model prediction accuracy, which can adjust the time window of prediction model quickly. The multidimensional disk resource scheduling can be divided into three modules: 1) the first module is the virtual machine disk interface converter; 2) the second module is dynamic virtual machine disk controller; and 3) the third module is the virtual machine delay controller. The multidimensional disk resource allocation algorithm should be implemented with the help of QoS-sensor of each virtual machine.

4) Network I/O resources optimization in communication-intensive environment

In the CPPS environment for CIA, network I/O communications between virtual nodes frequently happen. There is an alternative to establish performance optimization strategy for parallel application network I/O communication. The key to improve the communication performance of parallel applications is to avoid the unnecessary message polling. To this end, the following three approaches may be considered: 1) using blocking polling mechanism instead of busy polling mechanism; 2) improving scheduling algorithm of VMM and publishing scheduling information of the guest OS to the VMM; and 3) publishing the scheduling information of VMM to the guest OS. For the first approach, regarding the communication scenario between different virtual machines, the blocking operation should be performed on file descriptor set, instead of non-blocking operations implemented on the MPI library. This method can avoid to waste CPU resources on the unnecessary message polling. However, this solution will cause frequent and expensive VCPU context switch, so it may reduce the performance and also make the implementation of block polling operation in the shared memory more difficult. In order to solve this problem, the third method is proposed to improve performance of communication for the parallel I/O application. When the process that is receiving a message is blocked and eady to give up the ownership of the processor, it will inform the VMM to reschedule at appropriate time.

B. Scheduling of Cloud Resources for CIA

As we know, the existing cloud resource scheduling strategy is mainly designed for Internet application, but CIA has a completely different application characteristics. First of all, the traditional CIA generally runs on a dedicated system or High Performance Computing (HPC) system, so its execution environment is a dedicated or pre-configured with less dynamic resource scheduling. In cloud environment of CIA, each layer has complex functions. For example, the IaaS layer is the virtual cluster with a group specific constraints, the PaaS layer is a processing framework and computing scripts under its control, the SaaS layer is a particular application service, and there probably exists the combination module of above three layers, which is more complex. Therefore, CIA has to face more complex resource scheduling problem. Second, the scheduling goal of cloud resources for CIA is to map the virtual machines in the cluster to the appropriate physical servers according to the index constraints, and enable multi-level synergy scheduling between physical machine, multi-core processors and virtual machines. At the same time, the cloud resource scheduling management needs to consider the lifecycle management of virtual machines, as well as resources reserve or queue sharing scheduling policy of virtual machines, which should be more complex than traditional resources scheduling. Thirdly, another goal of the cloud resource scheduling for CIA is to assign processing tasks to the appropriate computing nodes according to various characteristics of CIA's component, which also increases the difficulty of resource scheduling. Finally, CIA has enormous components and complex process flow, the single reliability safeguard mechanism cannot adapt to application characteristics of CIA. So, it is necessary to establish the better reliability safeguard mechanism to ensure the high reliability of CIA. To solve the resource schedule problems in cloud environment for CIA, we can consider from the following four aspects respectively: the resource mapping strategy, multi-layer scheduling model, the CPU/GPU hybrid scheduling algorithm, and reliability guarantee mechanism. Figure 3 shows the whole technical framework.

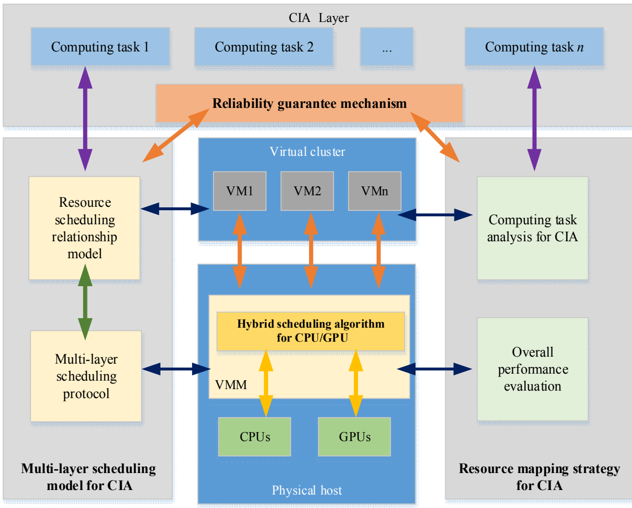

Figure 3. Resources scheduling framework for CIA in cloud environment

1) Multi-layer scheduling model for CIA

The key of multi-layer scheduling for CIA is how to solve the coordination scheduling problem between multi-layer resources. For example, the virtual cluster may require execution rates of all virtual machines are coordinated to be roughly the same. This scheduling demand is not only the problem of virtual cluster scheduling, but also is the problem of physical resources scheduling, which must be guaranteed by the coordinated scheduling between virtual nodes and physical nodes. We suggest establishing the resource scheduling relationship model between virtual layer and physical layer, and design multi-layer scheduling protocol to ensure the harmonization of multi-layer scheduling.

2) Resource mapping strategy for CIA.

The problem of resource mapping in cloud computing environment for CIA covers three aspects: the overall performance optimization, resources utilization maximization, and dynamic resource mapping to adapt the process complexity of CIA. In order to completely solve resource mapping problem and consider above three aspects, we may adopt the following tree steps. The first step is to analyze and refine the calculation and process features of CIA, then evaluate the physical hardware computing resources. The second step is to establish the relationship model for different characteristics of computing tasks and computing environment. The third step is to create the resource mapping strategy for performance optimization of CIA. If cloud computing resources are abundant, in order to get best overall performance of CIA, it is necessary to maximize the communication ability between virtual machines, especially to ensure low delay communication between virtual nodes. In a cloud environment, there are many complex engineering tasks. If we just meet the best overall performance of each CIA one by one, the resources of cloud computing will be wasted greatly. The best way is to calculate the cost of each CIA, then adjust the overall performance of the single CIA through appropriate resource sharing and reuse, which may result in the overall maximum effectiveness of cloud computing resources.

3) Hybrid scheduling algorithm for CPU/GPU

In order to implement a fair and effective shared accelerator between multiple virtual machines, a feasible solution is to add GPU in the VMM virtualization layer, which may provide the unified virtual accelerator interface for application layer. When application layer tries to invoke the GPU interface, VMM will take over the control for the execution, which can enable multiple applications execute in GPU in parallel. This method can make full use of the GPU parallel execution ability and implement hybrid scheduling that combines CPU and GPU. Through shared memory technology, the communication between virtual machines and VMM also can be reduced, which will improve the performance of CPU/GPU greatly.

4) Reliability guarantee mechanism for CIA

CIA has complex process flow, and the reliability of the whole process is necessary. An effective method to guarantee CIA's reliability is redundancy control service, also called synchronous hot standby. The various process stages of CIA involve various cloud services, such as IaaS, PaaS and SaaS. The reliability of PaaS and SaaS can be guaranteed by traditional parallel and distributed computing technology, but the reliability of IaaS is not so simple, because it cannot simply be guaranteed by the reliability of each independent virtual machine. Redundancy control service creates synchronous checkpoints of all virtual machines of CIA, used to find fault recovery points within the scope of the cluster. When a failure occurs, all of the virtual machines will be restored from the checkpoint, thus ensure the subsequent calculation correct. The difficulty is how to determine the synchronization checkpoints of all virtual machines in virtual cluster. A solution is to let virtual cluster enter a specific status, called half synchronous consistent coordinate status when it begins to synchronize checkpoints. The virtual machines that have entered this status cannot send communication data, but can receive communication data. Until all of the virtual machines have entered this status, they start to create consistent synchronous checkpoints. This solution can guarantee that the checkpoints of all virtual machines are consistent within the scope of virtual cluster. The synchronization process may bring extra time overhead, but the influence on overall performance of CIA is very slight.

IV. CONCLUSIONS

With the advances in emerging technologies, it is feasible to seamlessly integrate CPS with cloud computing, which provides tremendous opportunities for CIA. In this article, we give a brief review and outlook of the related background, and discuss the CCPS architecture and the enabling technologies for CIA. In particular, we dissect three key challenges (virtualized resource management techniques, scheduling of cloud resources for CIA, and LCM) and give the possible solutions to improve the performance and QoS of CCPS. We believe CCPS for CIA will attract enormous attention and research effort in the near future.

REFERENCES

[1] J. Wan, M. Chen, F. Xia, D. Li and K. Zhou, "From Machine-to-Machine Communications towards Cyber-Physical Systems," Computer Science and Information Systems, vol. 10, no. 3, pp. 1105-1128, 2013.

[2] J. Wan, D. Zhang, S. Zhao, L. T, Yang and J. Lloret, "Context-aware Vehicular Cyber-Physical Systems with Cloud Support: Architecture, Challenges and Solutions," IEEE Communications Magazine, vol. 52, no. 8, pp. 106-113, 2014.

[3] S. Sridhar, A. Hahn, M. Govindarasu, "Cyber-Physical System Security for the Electric Power Grid," Proceedings of the IEEE, vol. 100, no. 1, pp. 210-224, Jan. 2012.

[4] A. Banerjee, K. K. Venkatasubramanian, T. Mukherjee and S. K. S. Gupta, "Ensuring Safety, Security, and Sustainability of Mission-Critical Cyber-Physical Systems," Proceedings of the IEEE, vol. 100 , no. 1 pp. 283-299, Jan. 2012.

[5] A. Rajhans, A. Bhave, I. Ruchkin, B. H. Krogh, D. Garlan, A. Platzer, B. Schmerl, "Supporting Heterogeneity in Cyber-Physical Systems Architectures," IEEE Transactions on Automatic Control, vol. 59, issue. 12, pp. 3178-3193, Dec. 2014.

[6] P. Derler, E. A. Lee and S. V. Alberto, "Modeling Cyber-Physical Systems," Proceedings of the IEEE, vol. 100, no. 1, pp. 13-28, Jan. 2012.

[7] L. Parolini, B. Sinopoli, B. H. Krogh and Z. Wang, "A Cyber-Physical Systems Approach to Data Center Modeling and Control for Energy Efficiency," Proceedings of the IEEE, vol. 100, no. 1, pp. 254-268, 2012.

[8] S. Caliskan, M. Rungger, R. Majumdar, "Towards Robustness for Cyber-Physical Systems," IEEE Transactions on Automatic Control, vol. 59, no. 12, pp. 3151-3163, 2014.

[9] J. Wan, D. Zhang, Y. Sun, K. Lin, C. Zou and H. Cai, "VCMIA: A Novel Architecture for Integrating Vehicular Cyber-Physical Systems and Mobile Cloud Computing," ACM/Springer Mobile Networks and Applications, vol. 19, no. 2, pp. 153-160, 2014.

[10] R. Drath and A. Horch, "Industrie 4. 0: Hit or Hype?" IEEE Industrial Electronics Magazine, vol. 8, no. 2, pp. 56-58, Jun. 2014.

[11] H. Suo, Z. Liu, J. Wan and K. Zhou, "Security and Privacy in Mobile Cloud Computing," Proc. of the 9th IEEE Int. Wireless Communications and Mobile Computing Conf., Cagliari, Italy, July, 2013.

[12] M. Chen, J. Wan, S. González, X. Liao and V. Leung, "A Survey of Recent Developments in Home M2M Networks," IEEE Communications Surveys and Tutorials, vol. 16, no. 1, pp. 98-114, 2014.

[13] J. Liu, Q. Wang, J. Wan and J. Xiong, "Towards Real-time Indoor Localization in Wireless Sensor Networks, " Proc. of the 12th IEEE Int. Conf. on Computer and Information Technology, Chengdu, China, October, 2012, pp. 877-884.

[14] D. Zhang, J. Wan, Q. Liu, X. Guan and X. Liang, "A Taxonomy of Agent Technologies for Ubiquitous Computing Environments," KSII Transactions on Internet and Information Systems, vol. 6, no. 2, pp. 547-565, 2012.

[15] J. Wan, D. Li, H. Yan and P. Zhang, "Fuzzy Feedback Scheduling Algorithm based on Central Processing Unit Utilization for a Software-based Computer Numerical Control System," Proc. of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, vol. 224, no. 7, pp. 1133-1143, 2010.

[16] M. Chen, S. Mao, Y. Liu, "Big Data: A Survey," ACM/Springer Mobile Networks and Applications, vol. 19, no. 2, pp. 171-209, Apr. 2014.