Exploration of Wireless Sensor Network Based on Cloud Computing

https://doi.org/10.3991/ijoe.v14i11.9501

Qing Wan(✉), Ying Wang
Chongqing Vocational Institute of Engineering, Chongqing, China
qingwan38192@163.com

Abstract—To realize the exploration of wireless sensor network (WSN) based on cloud computing, the application service of WSN is taken as the starting point, the resource advantage of the cloud platform is used, and a WSN service framework based on cloud environment is proposed. Based on this framework, the problems of data management and reconstruction, network coverage optimization and monitoring, and edge recognition of holes are solved. In view of the node deployment of WSN and coverage problem of operation and maintenance optimization, the genetic algorithm is used to adjust the dormancy and energy state of nodes, and a parallel genetic algorithm for covering optimization in the cloud environment is proposed. For the operation and maintenance requirements of WSN, a parallel data statistics method for network monitoring is proposed. The experimental results show that the parallel algorithm is greatly improved in terms of the accuracy and time efficiency.

Keywords—cloud environment, wireless sensor, covering

1 Introduction

Cloud computing is a new computing mode that provides services, data and resources through the network. From a business perspective, cloud computing can simplify the management process and quickly respond to market changes; from the operational requirements, the process can be more standardized, cost reduced, and energy saving; more data and more users can be met from the demand of computing. The virtualization, multi core, automation and technology brought about by cloud computing will promote technology progress. Cloud computing will change the previous storage and computing patterns and provide different levels of computing services from infrastructure to computing platforms and then to process services. Mode and technology are interrelated, which makes the service mode need new technology implementation, and some technologies will also produce cloud computing use mode.

In the researches on wireless sensor network (WSN), it is also necessary to introduce the concept of cloud computing and build a new computing model to improve the performance and service quality of the network and system. For example, the data are stored and reconstructed in the cloud storage environment, and the data resources are shared. In the cloud, the data is analyzed by using parallel technology and the data required can be quickly extracted from the mass data, so as to provide information resources for network system monitoring. In the cloud end, intelligent technology can also be used to design various data mining algorithms for long-term accumulated data, and pattern recognition tools can be applied to find the internal rules in the application of WSN.

Large scale sensor networks are applied to various fields to monitor and manage the physical world, but the network itself has not been monitored or managed. The monitoring of large-scale network systems is also difficult to manage, and there are many challenges in the process of management. The great number of network equipment, the variety of sensors, the diversification of network deployment, the different installation environment of the network, the limited equipment resources and the instability of the equipment transmission all affect the precision monitoring. Therefore, in large-scale network management, a new type of large-scale network monitoring system needs to be built to realize the functions of distributed collection, high-performance parallel processing and real-time monitoring and alarm for WSN.

Based on the above background, cloud computing is combined with WSN, and WSN based on cloud computing is mainly studied.

2 Literature review

The Internet of things (IoT) is the network with connection of objects. Meddeb et al. (2018) studied the IoT through a large number of scattered radio frequency identification (sensors, laser scanners and other small devices). The information sensed is transmitted to the designated processing facilities for intelligent processing through the Internet, so as to complete identification, positioning, tracking, monitoring and management work [1]. WSN is one of the most important sensing networks on the basic level of IoT, which is flexible and widely used.

Shahzad et al. (2017) pointed out that WSN is a multi-hop self-organized network system composed of a large number of cheap micro sensor nodes deployed in the monitoring area and formed through wireless communication. The purpose of WSN is to collaboratively perceive, collect and handle the information of the objects sensed in the coverage area and send it to the observer. Sensors, sensing objects and observers constitute the three element of WSN [2].

WSN has extensive research in all countries all over the world and WSN research originated in the military field. As early as in 1978, a distributed sensor network working group was set up by the Carnegie Mellon University in the support of the United States Department of Defense. Singh et al. (2017) specially studied the WSN-based military surveillance system [3]. The United States, the European Union, Japan, South Korea and other countries attach great importance to theoretical research and applied research in various fields. China's 2010 vision plan and "fifteen plan" list sensors as one of the key industries. Many domestic and foreign scientific research institutions and major companies have also carried out WSN research, and can provide users with WSN solutions.

For the applications of WSN, it allows a single sensor node not reliable, and the normal work of the network is guaranteed by a large number of node redundancy. This is an ideal state, and in the actual application, redundancy is limited by funds and resources, and the memory, communication and computing power of the node itself is limited. Therefore, no matter which application needs the network management strategy to guarantee the quality of the network. For example, the sensing nodes, relay nodes and gateways of WSNs studied by Chitrakar et al. (2017), need to cooperate with each other to achieve communication. The reliability of wireless transmission is affected by environmental factors. Once the network has holes or breaks, the information cannot be transmitted [4]. Kordafshari et al. (2017) discussed that, to save energy and prolong the network life cycle, the sensing node used rotation working mode, and then the node's residual energy and working state would change dynamically, and the network topology changed. Then, the unsupervised network structure is bound to cause the inaccurate coverage of the perceived network, and it is difficult to achieve high coverage and low energy consumption without monitoring and optimization [5]. Moreover, Sun et al. (2017) pointed out that the problem of "big data" in WSNs is also a problem must face in various applications [6].

Hirai et al. (2017) stated that cloud computing is a business computing model. It distributed computing tasks on the resource pool composed of a large number of computers, so that users can obtain computing power, storage space and information service on demand [7]. Subsequently, many IT enterprises, such as Amazon and Facebook, have published their own cloud computing plans and schemes, and various research institutions have also carried out a large number of related research, forming a series of key technologies from system architecture to resource management.

The development of IoT is inseparable from the support of cloud computing. Chang et al. (2017) suggested that the driving forces for cloud computing services become the IoT include the following aspects. Demand driving: the pressure for traditional technology to deal with massive information cost is great, but cloud computing can fully and reasonably use resources to reduce operating costs. Technology driving: it promotes the upgrading of the architecture combined with technology and impels the rapid development of cloud computing standards. Policy driving: the government pays great attention to IoT, cloud computing and other infrastructure development strategy [8].

To sum up, the above researches are mainly focused on cloud computing and WSNs, and lack research on the integration of them. Therefore, based on the above research situation, WSNs based on cloud computing are attached importance to. Using the computing, storage and information service functions of the cloud platform, the WSN is integrated into the cloud environment, and the problem of WSN is efficiently solved by using the large storage capacity and parallel computing technology of the cloud platform.

3 Method

3.1 An overview of WSN

Wikipedia's definition of WSN: WSN is a self-organizing network system formed by wireless communication between sensor nodes. The purpose of WSN is to perceive, collect and process the information of the objects in the area they can perceive and send them to network monitors. Therefore, the three elements of WSN are sensor nodes, sensed objects and monitors.

A typical wireless sensor network system includes wireless sensor nodes, receiving and sending sink nodes, base stations, Internet or satellite, and user terminal.

The basic components of a single sensor node are shown in Figure 1. It is generally composed of sensor units (composed of sensors and analog digital conversion modules), processing units (composed of embedded systems, including central processing unit, memory, and embedded operating system), wireless communication units and power units. In addition, location system, power regeneration unit and mobile unit may be configured according to the needs of specific applications.

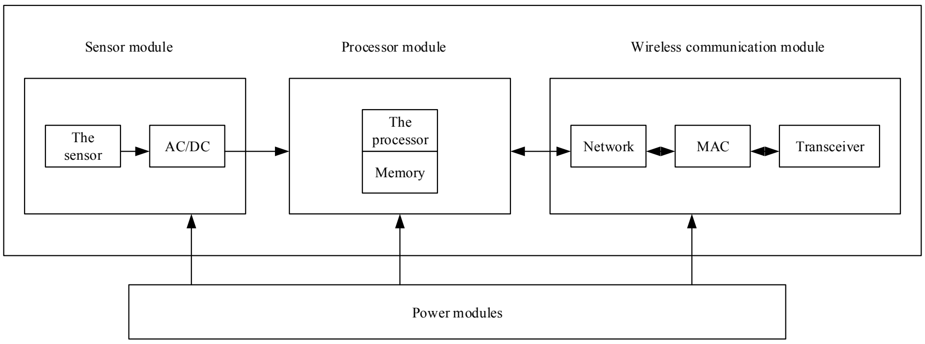

Fig. 1. The basic components of a single sensor node

3.2 Architecture of WSN

WSN is a kind of large-scale self-organizing network. It has the following important features: large scale nodes, self-organizing network, high dynamic network, data-centric monitoring network, and application related. Therefore, its network architecture is not only reflected in ensuring reliable network communication, but also related to other aspects such as network operation and maintenance.

As shown in Figure 2, the structure of WSN is divided into two levels: application service and network management. The application service is based on the application of sensor network, which is divided into the physical layer, the data link layer, the network layer, the transmission layer and the application layer in accordance with the architecture of the general computer network, so as to ensure the normal communication of the network. The distinguishing feature that network management is different from application service is that it aims at a certain management element and ensures the realization of management objectives through collaboration at different levels. For example, the realization of topology management is mainly link layer and network layer, and quality assurance, besides link and network, also contains the transmission layer. The architecture of WSN also includes a large number of alternative communication protocols, network management platforms and application support platforms.

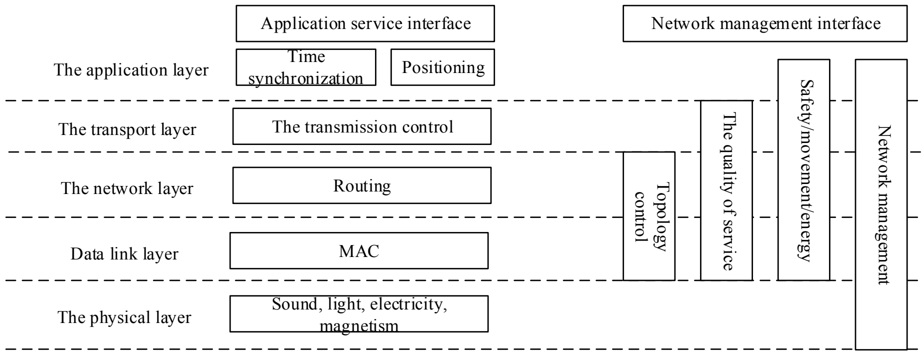

Fig. 2. Node structure of WSN

3.3 WSN management based on cloud environment

The main content of this project is to use the storage and computing environment provided by the cloud platform, take the quality of service in the network management of WSN as the research point, and use a small amount of data collection experiment and simulation to build a WSN service platform based on cloud environment.

The data management and network management functions shown in Figure 3 are implemented through distributed parallel computing using data centrally stored in the cloud.

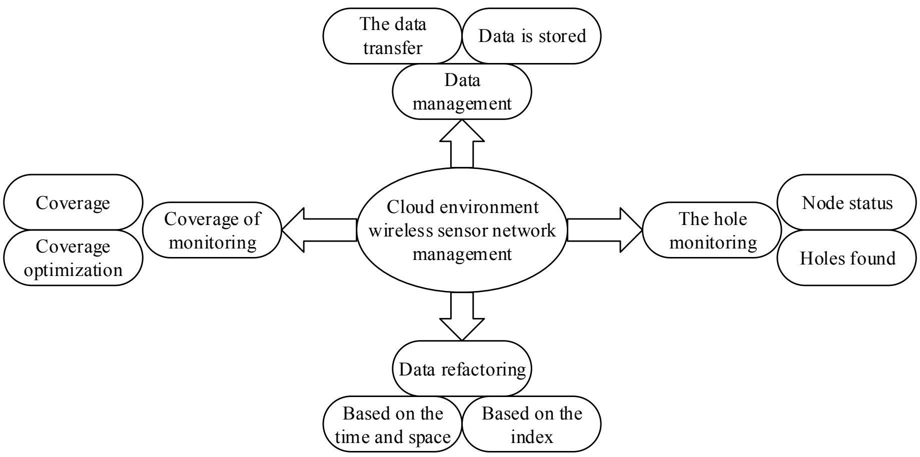

Fig. 3. The data management and network management functions

The main research problems can be summarized in three aspects: data management, network coverage and hole monitoring. Data management mainly includes the interface design of WSN network data to upload data for cloud environment and data reconstruction strategy after disordered data entering the cloud environment. Network coverage mainly solves the deployment, optimization and modeling problem of network, and adjusts the working state of nodes in time according to the history and current data monitoring. Hole monitoring calculates the node status in the network through data, and finds holes and the edge of the report hole in time by data analysis.

3.4 WSN application architecture based on cloud environment

The architecture of the cloud environment is the software architecture designed for this project, or it is understood as the software level based on the cloud platform. In the architecture, the abstract components and their relationships that constitute the system directly are briefly described. The architecture reflects the central and main levels of the system. More detailed implementation components and modules will be introduced in subsequent parts.

The architecture is mainly divided into three layers: resource layer, platform layer and application layer. The resource layer is a basic computing and storage unit with physical hardware devices, and the physical components of the resource layer are transformed into storage space and computing nodes by virtualization technology. The platform layer is a distributed software framework based on Hadoop architecture. The Hadoop architecture implements the distributed file management system (HDFS). HDFS is a master-slave structure with management nodes (NameNode) and data nodes, and create, retrieve, update and delete operations are performed on the file through the directory path. Clients interactively access files through NameNode and DataNode, NameNode manages the metadata of the file system, and DataNoode stores actual data. Each file is divided into different data blocks with data blocks (typical 64MB) and stored in different DataNode for use and fault tolerance. The Hadoop architecture implements the programming model and framework of MapReduce, which divides applications into many small working units and executes them on any cluster nodes. The application layer is the cloud service function implemented by this project. The main service content is data acquisition, storage and management of WSN. The data service realizes the storage of the perceived data stream in HDFS file, and reconstructs the disordered data in time and space relations to achieve the quick query when using the data. The management service mainly includes network deployment, coverage monitoring, node state statistics and hole discovery.

3.5 Data management in WSN based on cloud environment

WSN sensing data collection and storage process is the continuous accumulation process of streaming data. Data needs to go through gateway agent from WSN to cloud and gateway has two functions: converging WSN data and accessing the cloud end. The gateway receives the data uploaded by the perceived node, preprocesses and clean the data, and then stores it in the way of data flow. At this time, the gateway starts to access the cloud end and stores the accumulated data in the form of file. In the process of saving the flow data to the cloud end, it involves the problems of the network traffic consumption, the rational use of storage space and the data preprocessing during the data transmission.

The different application occasions of WSN can cause different requirements for the system to perceive data. If it is a security system, 1000 nodes are designed, each node takes data once per minute, then the amount of data is 1.44 million information per day; if the 500 nodes are designed to monitor the forest fire, then the monitoring information is scarce in no fire situation, and once a fire breaks out, it must respond immediately; if it is an intelligent traffic monitoring system, then there will be peak and trough periods. Therefore, the data upload mode is restricted by the frequency of event occurrence, real-time perception and periodic monitoring, so the design of data upload model should take into account the above cases.

In order to organize and manage data more reasonably, gateway nodes are divided into different types, and corresponding data upload models are set up for different types of gateway nodes. Three transmission models: real-time uploading, timing uploading and quantitative uploading are designed, and transferring data from gateway to cloud is called uploading. Real-time uploading is designed for real-time applications. The perception data in the gateway node cache area is uploaded to HDFS platform directly. The uploading interface of the gateway directly connects to the gateway data buffer and keeps the HDFS data connection. The data collected by the gateway is transmitted to the HDFS in real time through the network. Timing uploading is designed for periodic applications. The gateway designs the timer program, and the clock triggers the gateway to transmit data to HDFS. Read the gateway buffer every six hours, start the HDFS connection program, and upload data. Quantitative uploading is designed for the applications with small data collection. When the buffer size meets the preset size, it triggers the gateway to transmit data. At the gateway, the data buffer monitor program is designed, and the buffer size is monitored in real time. If the size of the data buffer meets the preset size, the connection program is started immediately and the data is uploaded.

4 Results

4.1 Data reconstruction based on cloud computing

Data reconstruction uses the MapReduce parallel programming framework of Hadoop cloud platform, and reconstructs the original WSN data files that have been saved on the HDFS, so that the newly constructed output files meet the requirements of the application.

The Map function is used to process key value pairs for different refactoring methods and output the results to HDFS, and the HDFS files generated are reconstructed data. The four reconstruction methods mainly use the process to reconstruct the file: the node reconstruction. The Map phase completes the classification according to the identification (ID), and the node ID is used as the key to judge the ID bit of each data information. If the ID of the record conforms to a certain key value, then this record is used as the value of the key. After a round of Map, the file will be divided into file blocks based on the node ID. In addition, the reconstruction scheme, such as the sorting of the nodes can be conducted. For the time reconstruction, the Map phase complete the classification by time, and the time is used as key to determine the time bits of each data information. If the time of the record is consistent with a certain key value, then this record is taken as a value of the key. Space reconstruction: because the data uploading to the cloud platform is carried out through the gateway, the space reconstruction is also the reconstruction based on the gateway. The Map phase completes the classification by gateway, and gateway ID is used as key to judge the gateway ID bit for each data information. If the ID of the recorded gateway corresponds to a certain key value, then this record is used as the value of the key. After a round of Map, the file will be divided into blocks based on the gateway. Mixed reconstruction: if it is a mixed reconstruction, the results of the first ten Map can be restructured using the Map continuously.

4.2 Region parallel SCPGA algorithm and principle component analysis PCPGA algorithm

According to the experimental results of status control genetic algorithm (SCGA) and power control genetic algorithm (PCGA), the performance of the algorithm cannot be increased by increasing the number of iterations when the population is certain, but it can be improved by increasing the population size. For large-scale population, the status control parallel genetic algorithm (SCPGA) and power control parallel genetic algorithm (PCPGA) based on parallel genetic algorithm are designed. The experiment shows that these two algorithms can efficiently solve the large-scale population genetic algorithm.

When the wireless sensor coverage optimization problem is solved by SCPGA algorithm and PCPGA algorithm, the gene length of the two algorithms will increase when the wireless sensor network is large. If genetic operation is conducted on all these genes, the parallel efficiency advantage will decrease, and the convergence speed and effect of the algorithm are difficult to be determined. Because of the high correlation between the sensing nodes and the region in the WSN, region division can reduce the complexity of genetic operation. In order to accelerate the iteration of the genetic algorithm and improve the efficiency, based on the convergence of the genetic algorithm of the distributed population, it is proposed to divide the problem and seek the optimal region. The concrete realization scheme is to divide the nodes into blocks according to the region. The nodes in each block are divided into a single population, the region is divided by Map, the area number is key, and the node information is used as the values. In other words, a Map operation is added before the PGA algorithm, so that each division area is iterated separately and parallel, which can improve time efficiency.

4.3 Design of genetic operator for coverage optimization

Two algorithms are designed to reduce network consumption on the premise of satisfying the coverage rate: controlling sleep and controlling perceived power.

The state control refers to the adjustment of the working state of the WSN node. The algorithm takes the coverage rate and the dormancy rate as the optimization goal, and selects the set of work nodes by genetic optimization to meet the requirements of coverage and energy consumption. The main steps of the SCGA genetic algorithm are described as follows:

First, because the energy and coverage rate is the optimization goal, fitness function uses the weighted sum of coverage and dormancy rate as the basis of evaluation. The function description is shown in Formula (1).

$$f = \alpha p_{\mathrm{cov}} + \beta p_{dor}. \tag{1}$$

In (1), α and β are adjustable parameters whose values depend on the comprehensive requirements of network applications for network performance indicators.

Second, binary coding is used, each gene bit corresponds to a node, and the same gene bits of different individuals correspond to the same node. 1 indicates that the node is in the working state, and 0 indicates that the node is in a dormant state.

Third, the selection operation. The tournament method is to rank the fitness values of individuals and select the high fitness value as the parent. The tournament method can identify the parent more quickly, which is used here to select the parent individuals.

Multi-point crossover. Crossover operation is mainly to preserve the good genes in the parent's generation. Through crossover operation, the genetic structure of the individual is changed, and multi-point crossover is the way to produce better new individuals. A random multi-point crossover method is used to generate a string of binary codes randomly according to the number of genes. If the corresponding gene bit is 1, then the gene bit participates in the crossover; if the corresponding gene bit is 0, then crossover is not conducted, and the parents are produced by random selection. Supposing that there are 4 genes, the parents are shown as Formula (2).

$$\begin{cases} X = \left(x_1, x_2, x_3, x_4\right) \\ Y = \left(y_1, y_2, y_3, y_4\right) \end{cases}. \tag{2}$$

Mutation operation. Mutation operation is to replace some of the gene values in a chromosome coded string on an individual by other alleles, thus forming a new individual, improving the local search ability of the genetic algorithm and maintaining the diversity of the population by the mutation operator, so as to prevent the occurrence of early mature. In this algorithm, the mutation gene bits are randomly generated.

Power control means that the perceptive power of the node can be adjusted to control the node perception radius, so as to optimize the network coverage and reduce the energy consumption of the network. The implementation steps of the PCGA are described as follows:

First, the algorithm aims at reducing energy consumption and increasing coverage, and the fitness function uses the weight of coverage and energy consumption as the evaluation basis. Because the fitness function is generally not negative, Formula (3) is used as the fitness function.

$$f = \alpha p_{\text{cov}} + \beta\left(1 - p_{cmc}\right). \tag{3}$$

In (3), α and β are adjustable parameters whose values depend on the comprehensive requirements of network applications for network performance indicators.

Second, the algorithm uses decimal coding, uses the power transmission level of the node as the coding, and uses 0~9 to code, of which the 0 level indicates the node dormancy, the 9 level indicates that the node senses with the maximum power, and the different power corresponds to the different sensing radius. Each individual's gene bit corresponds to different nodes, and the same gene bit of different individuals corresponds to the same node.

Third, the selection operation. Selection operation refers to how to select the parent individual, which commonly include roulette and tournament. Roulette method is used to evaluate genes by probability, and selection is closely related to probability. The probability of roulette is described as follows: assuming that the number of individuals in a population is N and the fitness of individual i is f_i, the probability of the individual being selected is shown in Formula (4):

$$p_{choose} = \frac{f_i}{\sum_{i=1}^{N} f_i}. \tag{4}$$

Obviously, the greater the adaptability of the roulette is, the greater the probability being selected is, which can make the genetic algorithm evolve in a good direction and greatly help to improve the quality of the solution.

Multi-point crossover. The crossover method adopts the random multi-point crossover method mentioned above.

Mutation operation. The mutation gene is randomly produced in this algorithm. The mutation gene accounts for 1% of the total number of genes. The mutation is based on the suburban mutation principle. If a gene is selected to mutate and if the gene is the minimum, the mutation is performed; if the maximum value is conducted with reduced by 1, if it is between the two values, then a number is randomly generated from 0 and 1. If the mutation of addition of 1 is conducted to 0, then the reduction of 1 is conducted, as shown in Formula (5).

$$x_k^{'} = \begin{cases} x_k - 1, x_k = 9 \\ x_k + 1, x_k = 0 \\ x_k + 1, 0 < x_k < 9 and \lambda = 0 \\ x_k + 1, 0 < x_k < 9 and \lambda = 1 \end{cases}. \tag{5}$$

λ refers to 0 and 1 these two random numbers.

Other operations. Since the coding sequence is arranged according to the geographic information of the node, the perception range of the adjacent nodes is overlapped and the neighborhood optimization strategy can be implemented after the genetic algorithm is executed. The optimal individual gene is alternated with operations of addition of 1 and reduction of 1, and then the target function is calculated to see if the results are improved, which is more conducive to the optimization of energy. Finally, combined with energy balance formula, the optimal results are selected.

The description of SCGA and PCGA algorithm:

Step 1: generate the initial population randomly, set the population algebra t=0, and produce the initialization group satisfying the fitness requirements. Step 2: calculate the evaluation function values of all individuals, and sort the individual according to the fitness value, then conduct the selection operation by the tournament or roulette way, and regard the selected individual as the new parent generation. Step 3: select the individual for crossover operation according to the set cross probability. Step 4: carry out mutation operation according to the mutation probability set. Step 5: determine whether the optimization effect or the number of iterations are achieved. If it is, turn to step 6; if it is not, turn to step 2. Step 6: output the best result. Step 7: optimize the field and select the best result.

5 Conclusion

As the core platform of data processing in WSNs, cloud computing is suitable for dealing with the regionally dispersed, massive data, dynamic and virtual application scenarios in WSNs. It can promote the sharing of sensor data at the bottom of the WSN, and provide super computing power for analysis and optimization, so as to provide reliable service more efficiently. The platform of WSN management platform based on cloud environment is constructed. Based on the analysis of the characteristics of the cloud platform and WSN, the services provided by the cloud environment for WSNs and the advantages of the cloud environment in solving the WSNs are summed up. The cloud environment can store and deal with the large amount of data in the WSN, and the cloud environment can efficiently solve the WSN management and supervision problems. Subsequently, the physical structure and service architecture of WSN in cloud environment is proposed, and a cloud management platform for WSN is created.

6 Acknowledgment

New Technology Promotion Project of Higher Vocational and Technical Colleges in Chongqing "Development and Application of Smart Campus New Technology Based on the Mobile Internet and Big Data Support" (Grant No.: GZTG201609); Science and Technology Research Project of Chongqing Municipal Education Commission "University Information Innovation Teaching Integrated Platform Research and Practice Based on Cloud Technology" (Grant No.: KJ1732434)

7 References

[1] Meddeb, M., Dhraief, A., Belghith, A., Monteil, T., Drira, K., & Al-Ahmadi, S. (2018). Named data networking: a promising architecture for the internet of things (iot). International Journal on Semantic Web & Information Systems, 14(2): 86-112. https://doi.org/10.4018/IJSWIS.2018040105

[2] Shahzad, M. K., & Cho, T. H. (2017). A network density-adaptive improved ccef scheme for enhanced network lifetime, energy efficiency, and filtering in wsns. Ad Hoc & Sensor Wireless Networks, 35(1): 129-149.

[3] Singh, A. K., Kumar, S., & Choudhary, A. (2017). An energy efficient architecture for border area surveillance system using wsns. International Journal of Control Theory & Applications, 10(13): 233-239.

[4] Chitrakar, R., & Huang, C. (2017). Selection of candidate support vectors in incremen-tal svm for network intrusion detection. Computers & Security, 45(3): 231-241.

[5] Kordafshari, M. S., Movaghar, A., & Meybodi, M. R. (2017). A joint duty cycle sched-uling and energy aware routing approach based on evolutionary game for wireless sen-sor networks. Iranian Journal of Fuzzy Systems, 14(2): 23-44.

[6] Sun, Z., Ji, X., & Alkhatib, G. I. (2017). Hdac high-dimensional data aggregation con-trol algorithm for big data in wireless sensor networks. International Journal of In-formation Technology & Web Engineering, 12(4): 72-86. https://doi.org/10.4018/IJITWE.2017100105

[7] Hirai, T., Masuyama, H., Kasahara, S., & Takahashi, Y. (2017). Performance analysis of large-scale parallel-distributed processing with backup tasks for cloud computing. Journal of Industrial & Management Optimization, 10(1): 113-129. https://doi.org/10.3934/jimo.2014.10.113

[8] Chang, B. J., Tsai, Y. L., & Liang, Y. H. (2017). Platoon-based cooperative adaptive cruise control for achieving active safe driving through mobile vehicular cloud com-puting. Wireless Personal Communications, 97(4): 5455-5481. https://doi.org/10.1007/s11277-017-4789-8

8 Authors

Qing Wan works as associate professor at Chongqing Vocational Institute of Engineering, Chongqing 402260, China.

Ying Wang works as associate professor at Chongqing Vocational Institute of Engineering, Chongqing 402260, China.

Article submitted 06 September 2018. Resubmitted 21 October 2018. Final acceptance 30 October 2018. Final version published as submitted by the authors.