A Real-Time Big Data Gathering Algorithm Based on Indoor Wireless Sensor Networks for Risk Analysis of Industrial Operations

Xuejun Ding, Yong Tian, and Yan Yu

***Abstract*—The era of big data has begun and an enormous amount of real-time data is used for the risk analysis of various industrial applications. However, a technical challenge exists in gathering real-time big data in a complex indoor industrial environment. Indoor wireless sensor networks (WSNs) technology can overcome this limitation by collecting the big data generated from source nodes and transmitting them to the data center in real time. In this study, typical residence, office, and manufacturing environments were chosen. The signal transmission characteristics of an indoor WSN were obtained by analyzing the test data. According to these characteristics, a real-time big data gathering (RTBDG) algorithm based on an indoor WSN is proposed for the risk analysis of industrial operations. In this algorithm, sensor nodes can screen the data collected from the environment and equipment according to the requirements of risk analysis. Clustering data transmission structure is then established on the basis of the received signal strength indicator (RSSI) and residual energy information. Experimental results show that RTBDG not only uses the limited energy of network nodes efficiently, but also balances the energy consumption of all nodes. In the near future, the algorithm will be widely applied to risk analysis in different industrial operations.**

***Index Terms*—Adaptive clustering, big data, risk analysis, received signal strength indicator (RSSI), wireless sensor networks (WSNs).**

I. Introduction

BIG DATA represent a new era in data exploration and utilization and cover various industrial applications, e.g., professionalizing business intelligence operation in the automobile industry [1], solving routing and scheduling problems in transportation systems [2], improving the performance of supply chains by minimizing the negative effect of demand uncertainties [3], providing security for buildings and physical infrastructure in home surveillance and security systems [4], and analyzing supply chains with radio-frequency identification technology from both the risk and benefit perspectives [5]. Big data can also be used to analyze risks in industrial operations, particularly in product manufacturing [6]. Many manufacturing enterprises have strict requirements on equipment working conditions and environment conditions for high-quality products, such as chip fabrication plants, pharmaceutical factories, and food factories. In the product manufacturing process of the aforementioned enterprises, the working condition parameters of manufacturing equipment and environment data need to be gathered in real time. Abnormal information is extracted from these data for the risk analysis of product manufacturing to ensure the normal operation of a production system [7]. However, a technical challenge exists in gathering real-time big data in various environments. This limitation on real-time data collection can be overcome by wireless sensor networks (WSNs) [8]. WSN has become an important technological support for gathering big data, such as temperature, humidity, equipment working condition, health information, and electricity consumption, particularly for data collection and transmission in indoor environments. Real-time data can be gathered using smart sensors, including atmospheric sensors, thermometric sensors, humidity sensors, and accelerometers. The volume of data gathered by these sensors may reach the order of petabytes according to a report of ORACLE [9]. In the industrial field, big data denote the enormous volume of various real-time data that are gathered, managed, processed, and analyzed for industrial operations.

However, a major challenge for WSN is ensuring that real-time data can be transmitted to the data center. Sensor nodes require enough energy to relay the data gathered by many surrounding sensors [10]. Therefore, energy is one of the most important indicators in WSN and energy consumption should be managed well to maximize network lifetime [11], [12]. To solve the aforementioned problems, an energy-efficient routing algorithm for WSN needs to be designed to gather big data in real time.

Many routing algorithms for WSN have been reported to prolong network lifetime. Some routing algorithms are proposed for the universal environment [13]–[19]. Heinzelman *et al.* [13]

Manuscript received December 29, 2014; revised March 17, 2015 and May 07, 2015; accepted May 08, 2015. Date of publication May 20, 2015; date of current version June 02, 2016. This work was supported in part by the Science Research Project of Liaoning under Grant L2013518, in part by the Social Science Planning Fund Project of Liaoning under Grant L14DGL045, in part by the National Natural Science Foundation of China under Grant 50921001, in part by the National Basic Research Program of China under Grant 2011CB013705, and in part by the Fundamental Research Funds for the Central Universities under Grant DUT15ZD117. Paper no. TII-15-0062. *(Corresponding authors: Yong Tian and Yan Yu.)*

X. Ding is with the School of Management Science and Engineering, Dongbei University of Finance and Economics, Dalian 116025, China (e-mail: dingxj812@163.com).

Y. Tian is with the School of Physics and Electronic Technology, Liaoning Normal University, Dalian 116029, China (e-mail: tianyong_9081@163.com).

Y. Yu is with the School of Electronic Science and Technology, Dalian University of Technology, Dalian 116024, China (e-mail: yuyan@dlut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2015.2436337

proposed a centralized clustering algorithm called the low-energy adaptive clustering hierarchy (LEACH). The disadvantage of centralized algorithms is that each sensor node must transmit its location and residual energy information to a base station (BS). Thus, additional energy overhead is needed. Younis and Fahmy [14] proposed a distributed clustering algorithm called hybrid energy-efficient distributed (HEED) clustering algorithm. The disadvantage of distributed algorithms is that clusters are built on the basis of the information of the neighbor nodes. Hence, the energy overheads are uneven. Other scholars have made some progresses toward the routing algorithm of indoor WSN [20]–[26]. Huang *et al.* [20] presented an indoor routing algorithm called weight coefficient adaptive-based indoor energy load-balanced routing (WAIER). This algorithm uses the multiple attribute decision method with objective and subjective weight coefficients to build the routing, as well as selects the optimal routing paths by comprehensively considering different attributes, such as the residual energy of a node. Daabaj *et al.* [21] presented a reliable load-balanced routing (RLBR) algorithm for indoor WSN. In this algorithm, the relaying workload and energy consumption are redistributed in accordance with the routing information of different layers. Tümer and Gündüz [22] presented randomly clustered energy-efficient routing protocol (R-EERP) and sequentially clustered energy-efficient routing protocol (S-EERP) algorithms based on LEACH. In R-EERP and S-EERP, cluster nodes do not change when cluster-head nodes are changed. Nevertheless, the R-EERP algorithm randomly deploys sensor nodes in a field, whereas S-EERP algorithm sequentially deploys sensor nodes in an indoor environment. Li *et al.* [23] studied the hop count and link quality of WSN and presented a routing algorithm that can adaptively choose an optimal routing in large-scale WSN. Some key technologies for industrial WSN have been investigated in [27]–[33]. Marchenko *et al.* [27] presented an experimental study of selective cooperative relaying protocols, including three practical relay update schemes, and evaluated the results in an industrial production plant. Abrishambaf *et al.* [28] presented an energy analysis of routing protocols in WSN for industrial applications. The analysis results show that the key factors of a routing algorithm include the frequency of transmission, the number of sensor nodes, the message length, and the distance between sensor nodes. Shen *et al.* [29] presented the first priority-enhanced medium access control (MAC) protocol that is compatible with industrial standards for industrial wireless sensors and actuator networks. The experimental results show that the proposed protocol achieves a higher performance than those obtained by current industrial standards. Villaverde *et al.* [30] presented a route selection algorithm. In the algorithm, local information is shared among neighboring nodes to meet the requirements of industrial applications. Simulation results show that the proposed algorithm satisfies a typical quality of service requirements with low-control overhead. The aforementioned routing protocols obtain good performance for different application environments. However, their performances are not ideal when gathering real-time big data for risk analysis of industrial operations because they did not comprehensively consider the signal transmission characteristics and routing requirements for industrial WSN [34].

In this study, three typical experimental scenarios are selected: 1) residence; 2) office; and 3) factory environments. The received signal strength indicator (RSSI) values of nodes are tested in different environments, different locations of the same environment, and dynamic environments. The signal transmission characteristics in an indoor environment are analyzed on the basis of the test results. According to the characteristics, an adaptive clustering routing algorithm based on RSSI called real-time big data gathering (RTBDG) is proposed to gather real-time big data for the risk analysis of industrial operations. In this algorithm, a cluster-head node and its relay node are chosen by a comprehensive consideration of the path loss and the residual energy of sensor nodes. The proposed RTBDG algorithm can adaptively build routing with changes in the indoor environment. The experimental results show that RTBDG can obtain a better performance in the aspect of energy consumption than the other three algorithms.

This paper is organized as follows. Section II describes the testing system. Section III shows the testing process in a residence environment and analyzes the test results. Section IV describes the testing process in an office environment and analyzes the test results. Section V shows the testing process in a chip manufacturing environment and analyzes the test results. Section VI gives the signal transmission characteristics. Section VII describes the proposed RTBDG algorithm. Section VIII presents the evaluation of RTBDG by an experiment and comparison with other algorithms. Finally, Section IX gives the conclusion.

II. Testing System

The experimental platform for the testing system is mainly composed of Pro2110 suites of the CrossBow Company. The testing system is designed using the following parameters: a signal center frequency of 2.4 GHz; a maximum data transfer rate of 250 kbps; and an output power of the transmitter node of 3.2 dBm. The wireless sensor nodes are in a stationary state, and no interference exists from moving object.

III. Residence Environment Testing

The RSSI values for residence environment are tested in a typical residence with two bedrooms, a living room, and a dining room to reflect the basic characteristics of a common residence environment. The testing environment is shown in Fig. 1. Fig. 1(a) shows the photo of the living and dining rooms. Fig. 1(b) shows the plan view of the residence environment, where room 1 is the living room, room 2 is the south bedroom, room 3 is the dining room, room 4 is the north bedroom, room 5 is the bathroom, room 6 is the kitchen, and room 7 is the wardrobe. In rooms 1, 2, 3, 4, and 7, the walls and ceilings are whitewashed with wall paints, the floors are laminate flooring, and the floor height is 2.7 m. In rooms 5 and 6, the walls and floors are made of tiles, the ceilings are made of alloy boards, and the floor height is 2.4 m.

Fig. 1. Residence environment. (a) Photo of the residence environment. (b) Plan view of the residence environment.

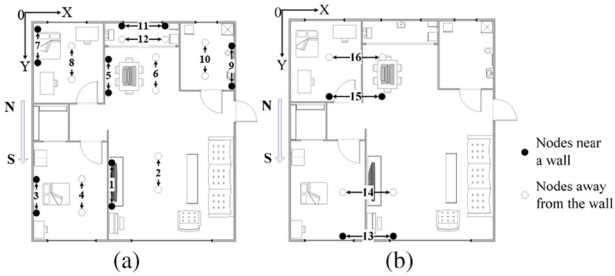


Fig. 2. Experimental deployment of the first test in a residence environment. (a) Deployment of nodes in the same room. (b) Deployment of nodes in different rooms.

A. Test Deployment and Test Method

The signal transmission path loss between nodes is generally smaller in the case of line-of-sight away from the floor, ceilings, and walls in the same room than in the case of nonline-of-sight close to them in the different room. Therefore, in the first test, data are measured in the same room and different rooms separated by one wall. In the aforementioned two scenarios, nodes are deployed near a wall and away from the wall, and the node heights from the floor include 0, 1.5, and 2.7 m (or 2.4 m). The distance is 2 m between the transmitting node and the receiving node, and 1000 sets of data are recorded for each test. Random noise is filtered out by a Kalman filter.

The experimental deployment of the first test is shown in Fig. 2. The filled circles represent the nodes near a wall and the hollow circles represent the nodes away from the wall. The digitals in solid line with double arrows are the testing sequence numbers. Fig. 2(a) and (b) shows the deployment of nodes in the same room and in different rooms, respectively.

When the positions of nodes are fixed, the signal transmission path loss between nodes will be changed if the deployment of furniture is changed or if personnel move continually. Therefore, in the second test, the distances between nodes include 2, 4, 6, and 8 m. In every scenario above, data are measured in the environment without any changes, with obstructions between nodes and with personnel movement between nodes, respectively. The node heights from the floor are 1 m and 1000 sets of data are recorded for each test. Random noise is filtered out by a Kalman filter.

The experimental deployment of the second test is shown in Fig. 3. The digitals in the figure are the testing labels. Tests 1–4 indicate that the distances between nodes are 2, 4, 6, and 8 m, respectively. For each scenario, data are measured in the aforementioned three environments and are shown in Fig. 3(a)–(c).

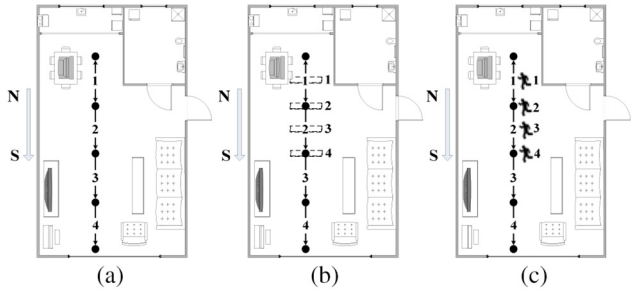


Fig. 3. Experimental deployment of the second test in a residence environment: (a) without any changes; (b) with obstructions between nodes; and (c) with personnel movement between nodes.

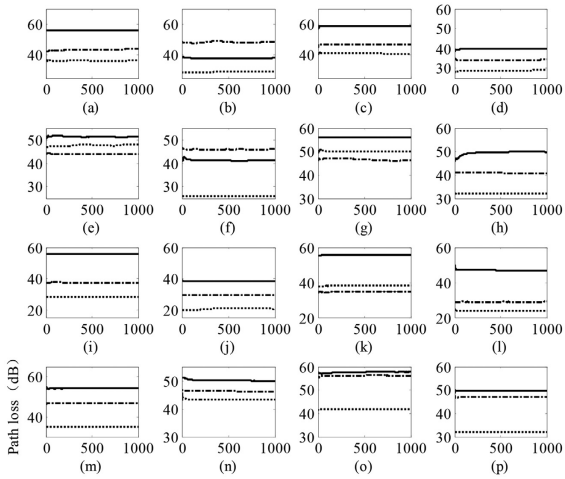


Fig. 4. Experimental results in the first test. Data of (a) test 1, (b) test 2, (c) test 3, (d) test 4, (e) test 5, (f) test 6, (g) test 7, (h) test 8, (i) test 9, (j) test 10, (k) test 11, (l) test 12, (m) test 13, (n) test 14, (o) test 15, and (p) test 16.

B. Test Results

Data are measured for every scenario in accordance with the test method described in Section III-A, and multiple sets of test data are obtained. Fig. 4 shows the experimental results in the first test, where Fig. 4(a)–(p) is the experimental results of test 1 and 2, test 3 and 4, test 5 and 6, test 7 and 8, test 9 and 10, test 11 and 12, test 13 and 14, and test 15 and 16 in the first test, respectively. The solid lines in Fig. 4 are the test data of the nodes placed on the floor, the dotted lines are the test data of the nodes that are 1.5-m high from the floor, and the dashed-dotted lines are the test data of the nodes placed on the ceiling.

Fig. 5 shows the experimental results of the second test, where Fig. 5(a)–(d) shows the original experimental results and Fig. 5(e)–(h) the experimental results filtered by a Kalman filter. In these figures, the solid lines are the test results in the environment without any changes, the dotted lines s are the test results in the environment with obstructions between nodes, and the dashed-dotted lines are the test results in the environment with personnel movement between nodes.

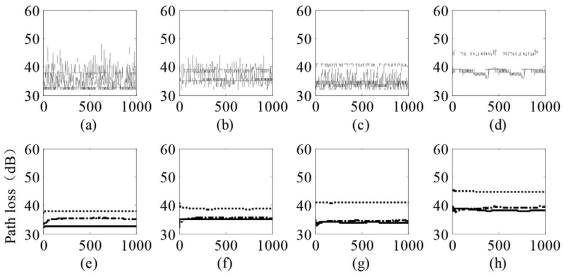


Fig. 5. Experimental results in the second test. Original data of (a) test 1, (b) test 2, (c) test 3, and (d) test 4. Filtered data of (e) test 1, (f) test 2, (g) test 3, and (h) test 4.

Fig. 6. Office environment. (a) Photo of the office environment. (b) Plan view of the office environment.

C. Analysis of Test Results

Two types of tests are conducted to fully test the signal transmission characteristics of WSN in a residence environment. The purpose of the first test is to analyze the signal transmission characteristics of WSN in a typical residence environment, when objects have large differences in properties and positions. The purpose of the second test is to analyze the characteristics when the testing environment changes dynamically.

From the experimental results of the first test in Section III-B, the following conclusion can be drawn.

1) The path loss of nodes near a wall is generally greater than that away from the wall.

   In Fig. 4(a), (b), (e), and (f), the path loss of nodes near a wall is slightly less than that away from the wall when the node height is 2.7 m. The reason for this condition is that the chandeliers exist between nodes in the living and dining rooms when the nodes away from the wall are placed on the ceiling. In Fig. 4(m) and (n), the same situation can be seen because the nodes are near the windows in test 13.

2) The path loss of nodes near the floor and ceiling is generally greater than that away from the floor and ceiling.

   In Fig. 4(e)–(h), (k), and (l), the path loss of nodes away from the floor and ceiling is greater than that near the ceiling when the node height near a wall is 1.5 m. The reason for this condition is that the frescos of the dining room, frescos of the north bedroom, and objects of the kitchen exist between nodes away from the floor and ceiling.

3) The path loss of nodes near the floor is generally greater than that near the ceiling.

   In Fig. 4(a), (b), (e), and (f), the path loss of nodes near the ceiling is greater than that near the floor when the nodes away from the wall are placed on the ceiling. The reason for this condition is that the chandeliers exist between nodes in the living and dining rooms.

4) The path loss of nonline-of-sight nodes is greater than line-of-sight nodes.

   For example, the aforementioned situation can be seen from Fig. 4(a), (b), (e), and (f) when the nodes away from the wall are placed on the ceiling.

5) Two sensor nodes with a fixed distance generate different path losses when they are in different indoor environments with different indoor materials, layouts, and sizes or in different locations of the same room.

For example, in Fig. 4(i)–(l), the path loss of nodes placed on the ceiling is smaller than that in other scenarios. The reason for this condition is that the materials of ceiling in Fig. 4(i)–(l) are alloys and wall paints in other scenarios. In Fig. 4(c), (d), (g), and (h), the path losses of nodes deployed in two different bedrooms are different because of the different indoor layouts. In Fig. 4(i)–(l), the path loss of nodes away from the wall is smaller than that in other scenarios when the node height is 1.5 m, because the areas of the kitchen and bathroom are smaller than other rooms.

From the experimental results of the second test in Section III-B, the following conclusion can be drawn.

1) Path loss increases with the increasing distances between nodes in a constant environment.
2) In all scenarios, the path loss between nodes significantly increases when obstacles are added between nodes.
3) Path loss will have a great fluctuation when a person is walking back and forth between nodes; however, after being filtered, the path loss will become close to that in the constant environment.
4) The obstacles or moving persons between nodes can increase path loss. Thus, the communication between nodes can be interrupted or data packets can be lost.

IV. Office Environment Testing

The RSSI values for an office environment are tested in a typical multiperson office with various office supplies to reflect the basic characteristics of a common office environment. The testing environment is shown in Fig. 6. Fig. 6(a) is the photo of the office. Fig. 6(b) is the plan view of the office environment, where rooms 1 and 2 are two offices with the same pattern and room 3 is the hallway. The two offices have foam ceilings, cement floors, and walls whitewashed with wall paints. Both the walls and ceilings of the hallway are whitewashed with wall paints, the floor of the hallway is made of floor tiles, and the ceiling height is 3.2 m.

A. Test Deployment and Test Method

Two types of tests similar to those used in the residence environment are conducted in an office environment. In the first test, data are measured in the same room and different rooms separated by one wall. In these two scenarios, nodes are deployed close to a wall and away from the wall, and the node heights

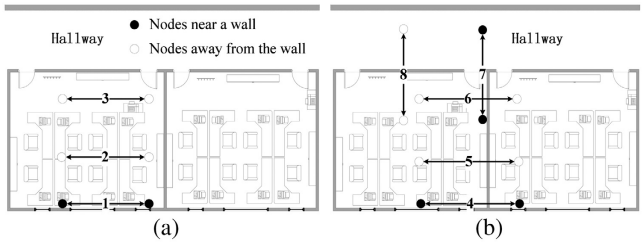


Fig. 7. Experimental deployment of the first test in an office environment. (a) Deployment of nodes in the same room. (b) Deployment of nodes in different rooms.

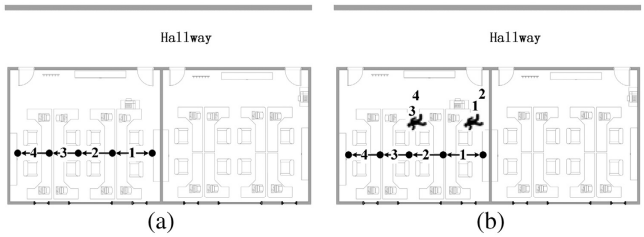


Fig. 8. Experimental deployment of the second test in an office environment (a) without any changes and (b) with personnel movement.

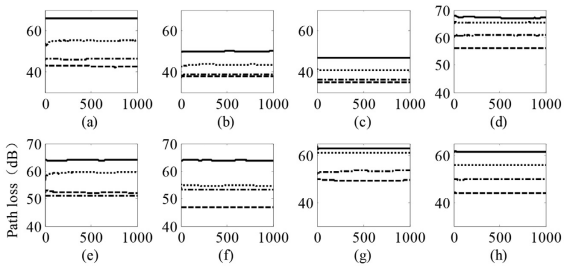


Fig. 9. Experimental results in the first test. Data of (a) test 1, (b) test 2, (c) test 3, (d) test 4, (e) test 5, (f) test 6, (g) test 7, and (h) test 8.

from the floor include 0, 1, 2, and 3.2 m. The distance is 5 m between the transmitting and receiving nodes, and 1000 sets of data are recorded for each test. Random noise is filtered out by a Kalman filter. The experimental deployment of the first test is shown in Fig. 7.

In the second test, the distances between nodes include 2, 4, 6, and 8 m. In every scenario, data are measured in the environment without any changes and with personnel movement between nodes. The node height from the floor is 1 m, and 1000 sets of data are recorded for each test. Random noise is filtered out by a Kalman filter. The experimental deployment of the second test is shown in Fig. 8.

B. Test Result

Fig. 9(a)–(h) shows the experimental results of tests 1–8 in the first test. The solid lines in these figures denote the test data from the nodes placed on the floor, the dotted lines are the test data from the nodes that are 1-m high from the floor, the dashed lines are the test data from the nodes that are 2-m high from the floor, and the dashed-dotted lines are the test data of the nodes placed on the ceiling.

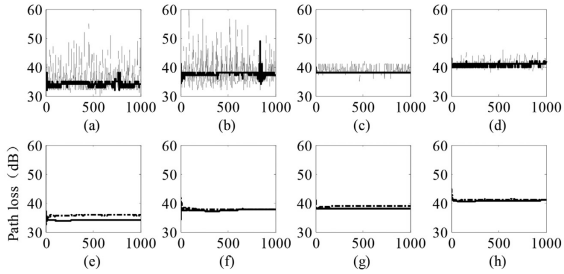


Fig. 10. Experimental results in the second test. Original data of (a) test 1, (b) test 2, (c) test 3, and (d) test 4. Filtered data of (e) test 1, (f) test 2, (g) test 3, and (h) test 4.

Fig. 10(a)–(h) is the experimental results of the second test, where Fig. 10(a)–(d) shows the original experimental results and Fig. 10(e)–(h) is the experimental results filtered by a Kalman filter. In these figures, the solid lines are the test results in the environment without any changes and the dashed-dotted lines are the test results in the environment with personnel movement between nodes.

C. Analysis of Test Result

From the experimental results of the first test in Section IV-B, the following conclusion can be drawn.

1) The path loss of nodes near a wall is generally greater than that away from the wall given the office partitions, various office supplies, and walls between nodes.
2) The path loss of nodes near the floors is generally greater than that away from the floors given the office partitions, various office supplies, and floors between nodes near the floor.
3) The path loss of nodes with 2-m height is generally smaller than nodes with other heights.

   The reason for this condition is that no office partitions, various office supplies, the floors, and ceilings between nodes exist when the node height is 2 m. In Fig. 9(e) and (f), the path loss of nodes with 2-m height is greater than nodes placed on the ceiling because file cabinets exist between nodes.
4) The path loss of nonline-of-sight nodes is generally greater than line-of-sight nodes.
5) Two sensor nodes with fixed distances generate different path losses when they are in different indoor environments or in different locations of the same room.

From the experimental results of the second test in Section IV-B, the following conclusion can be drawn.

1) Path loss increases with the increasing distances between nodes in a constant environment.
2) Path loss will have a great fluctuation when a person is walking back and forth between nodes, but it is close to that in the constant environment after it is filtered.
3) If a person moves between nodes, the communication between nodes can be interrupted, or data packets can be lost when the path loss is large enough.
4) If a person walks back and forth between nodes, the path loss will have a decreasing fluctuation with increasing distances between nodes.

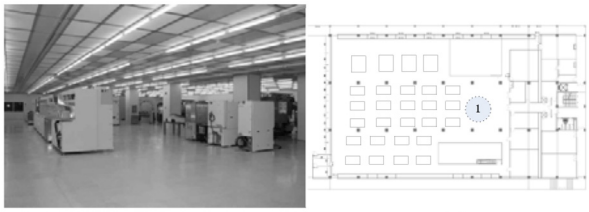


Fig. 11. Ultra-clean laboratory environment. (a) Photo of the ultra-clean laboratory. (b) Plan view of the ultra-clean laboratory.

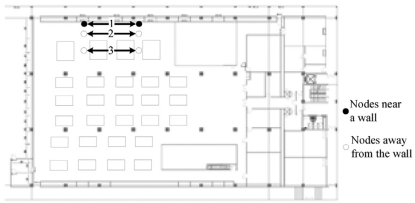


Fig. 12. Experimental deployment of the first test in the chip manufacturing environment.

V. Chip Manufacturing Environment Testing

The RSSI values for the chip manufacturing environment are tested in the 8 inch integrated circuit chip manufacturing laboratory of the Dalian Institute of Semiconductor Technology. The ultra-clean laboratory is a 1500 m^2 hundred grade clean room. This laboratory should have a temperature of 23 ± 1 °C, relative humidity of $45\% \pm 10\%$, and indoor air pressure of $+20$ Pa. The selected environment can reflect the basic characteristics of the factory manufacturing environment. The ultra-clean laboratory is shown in Fig. 11. Fig. 11(a) shows the photo of the laboratory, and Fig. 11(b) shows the plan view of the chip manufacturing environment, where room 1 is the ultra-clean laboratory and the rectangle is the schematic position of the chip production equipment. The ultra-clean laboratory is divided into two layers. The lower layer is built on different pipelines and cables, and different types of chip production equipment are placed in the upper layer. The baffles between the two layers are metal floors. Both the walls and ceilings of the laboratory are made of metal materials, and the floor height is 3.2 m.

A. Test Deployment and Test Method

Two types of tests are performed in the chip manufacturing environment. The first test includes tests 1–3. The nodes are deployed near a wall in test 1. The nodes in tests 2 and 3 are deployed away from a wall, and equipment is located between the nodes in test 3. The nodes in the above three tests are 0-, 1.5-, and 3.2-m high from the floor, respectively. The distance between the transmitting node and receiving node is 20 m, and 1000 sets of data are recorded for each test. Random noise is then filtered out by a Kalman filter. The experimental deployment of the first test is shown in Fig. 12.

The distances between nodes in the second test include 5, 10, and 20 m. The data in every scenario are measured in the environment without any changes and with personnel movement between nodes. The node heights from the floor are 1 and 3.2 m, and 1000 sets of data are recorded for each test. Random noise is then filtered out by a Kalman filter. The experimental deployment of the second test is shown in Fig. 13.

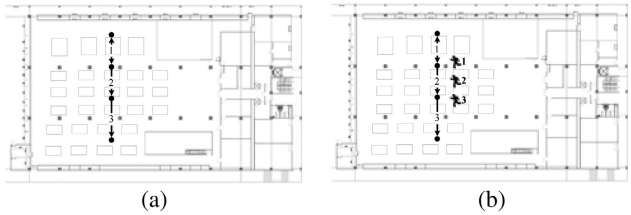


Fig. 13. Experimental deployment of the second test in the chip manufacturing environment (a) without any changes and (b) with personnel movement between nodes.

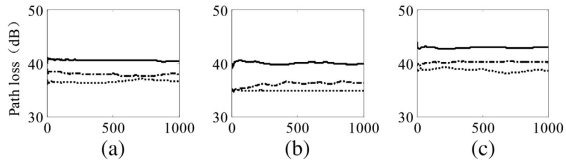


Fig. 14. Test results of the first test. Data of (a) test 1, (b) test 2, and (c) test 3.

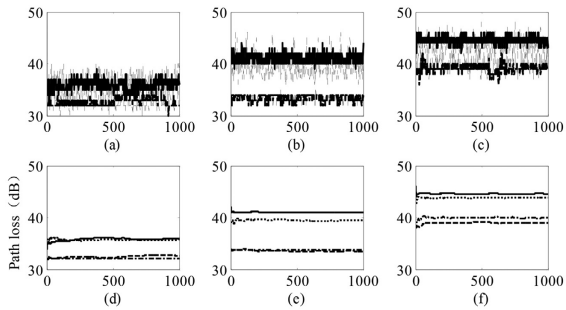


Fig. 15. Experimental results in the second test. Original data of (a) test 1, (b) test 2, and (c) test 3. Filtered data of (d) test 1, (e) test 2, and (f) test 3.

B. Test Result

Fig. 14 shows the experimental results of tests 1–3 in the first test. Fig. 14(a)–(c) shows the test results of tests 1–3, respectively. The solid lines in these figures are the test data from the nodes placed on the floor, the dotted lines are the test data from the nodes that are 1.5-m high from the floor, and the dashed-dotted lines are the test data from the nodes placed on the ceiling.

Fig. 15(a)–(f) is the experimental results of the second test, where Fig. 15(a)–(c) is the original experimental results and Fig. 15(d)–(f) is the experimental results filtered by a Kalman filter. The solid lines in these two figures are the test results in the environment without any changes, the dotted lines are the test results in the environment with personnel movement between nodes when the node height is 1 m from the floor, the dashed lines are the test data in the environment without any changes, and the dashed-dotted lines are the test data in the environment with personnel movement between nodes when the node height is 3.2 m from the floor.

C. Analysis of Test Result

From the experimental results of the first test in Section V-B, the following conclusion can be drawn.

1) The path loss of nodes near a wall is generally greater than that away from the wall.
2) The path loss of nodes near the floors is generally greater than that away from the floors.
3) The path loss of nodes whose heights are 1.5 m is generally smaller than the nodes with other heights because no floors and ceilings exist between the nodes.
4) The path loss of nonline-of-sight nodes is generally greater than line-of-sight nodes.
5) Each pair of sensor nodes with a fixed distance generates the different path losses when they are deployed in different rooms or in different locations of the same room.

From the experimental results of the second test in Section V-B, the following conclusion can be drawn.

1) Path loss increases with the increasing distances between nodes in a constant environment.
2) Path loss experiences significant fluctuation when a person is walking back and forth between the nodes but becomes close to the path loss in the constant environment after being filtered.
3) If a person is moving between nodes, the communication between nodes can be interrupted or the data packets can be lost when the path loss is sufficiently large.
4) When the node height is 3.2 m from the floor, the fluctuations in the path loss between nodes are smaller than when the height is 1 m.

VI. Signal Transmission Characteristics

The signal transmission characteristics of indoor WSN are summarized as follows.

1) The path loss of nodes near a wall is generally greater than that away from the wall.
2) The path loss of nodes near floors and ceilings is generally greater than that away from floors and ceilings.
3) The path loss of nonline-of-sight nodes is generally greater than line-of-sight nodes.
4) Two sensor nodes with a fixed distance apart generate different path losses when they are in different indoor environments such as different indoor materials, layout, and sizes, or in different locations of the same room.
5) Path loss increases with the increasing distances between nodes in a constant environment.
6) Path loss between nodes significantly increases when the obstacles are added between nodes.
7) Path loss experiences significant fluctuation when a person is walking back and forth between the nodes but is close to the path loss in the constant environment after it is filtered.
8) If obstacles exist or a person is moving between nodes, the communication between the nodes can be interrupted or the data packets can be lost when the path loss is sufficiently large.
9) Path loss experiences a decreasing fluctuation with the increasing distances between nodes if a person is walking back and forth between the nodes.

Thus, the WSN signal transmission characteristics are summarized by analyzing extensive test data in different indoor environments. These characteristics can provide a reliable basis for research on key technologies including big data gathering algorithms in indoor WSN.

A big data gathering algorithm can be confronted with the following technical challenges on the basis of the said characteristics.

1) WSN cannot determine whether two sensor nodes can communicate according to the distance information between two nodes.
2) Even if two sensor nodes can communicate, they cannot be guaranteed to communicate again in the future.

Therefore, a big data gathering algorithm should have an adaptive capacity in a complex indoor environment, but the adaptive capacity cannot be too sensitive to the environment. This condition is due to the energy consumption being largely increased if the network structure is frequently adjusted.

VII. RTBDG Algorithm

If sensor nodes are deployed on a rectangular field, the networks are supposed to have the following properties.

1) Sensor nodes are immobile after being deployed.
2) All nodes are homogeneous and energy constrained.
3) BS can send signals of sufficient strength so that all nodes can receive its signals.
4) BS knows the network topology.
5) Sensor nodes send data with a fixed transmission power.
6) The propagation channel is symmetrical.

The RTBDG algorithm proposed in this paper is designed to gather big data used in the risk analysis of industrial operations. Therefore, sensor nodes should screen the gathered data according to the requirements of risk analysis. The screening process can be described as follows.

1) The normal reference ranges of the data collected by sensor nodes are established.
2) Data are collected at a regular time interval t_1.
3) The data collected by sensor nodes with a normal reference value are compared. If these data are within the normal reference ranges, the data are stored at a regular time interval t_2 ($t_2 = n \cdot t_1$, where n is the natural number and can be set according to the requirements of risk analysis). Otherwise, both the abnormal and the stored data are transmitted to BS by the established routes.
4) If the amount of data stored in sensor nodes reaches the preset upper limit value of the storage capacity, all of the stored data are also transmitted to BS by the established routes.

The process of transmitting data to BS in the RTBDG algorithm is also divided into rounds as LEACH [13], but the operating phases of each round are different from LEACH. The RTBDG algorithm consists of six phases in each round.

1) BS broadcasts information to build clusters for a new round in the initial phase, and the sensor nodes calculate their own information.

When a new round begins, BS broadcasts a message (i.e., INIT_MSG) that contains its transmitting signal power value, which can be converted into RSSI and expressed as $\mathrm{RSSI}_{\mathrm{BS}-S}$, weight coefficient $w_{E-\mathrm{CH}}$, w_{E-R}, and the time division multiple address transmission slot allocated to each node. The signals sent by BS cover all sensor nodes. Thus, each node can monitor and receive the signals from BS and obtain the RSSI value defined as $\mathrm{RSSI}_{i-\mathrm{BS}}$ $(i = 1, \ldots, N)$, where i is the identification (ID) number of a sensor node and N is the total number of sensor nodes.

After each node receives a RSSI value from BS, the node uses the RSSI value to estimate the path loss between the BS and the node. Given that the propagation channel is symmetrical and the sensor nodes use fixed transmission power to send data, each node can easily calculate whether it can communicate directly with BS by checking the following formula:

$$\mathrm{RSSI}_{\mathrm{BS}-S} - \mathrm{RSSI}_{i-\mathrm{BS}} < \mathrm{RSSI}_{N-S} - \mathrm{RSSI}_{\mathrm{BS}-M} \tag{1}$$

where RSSI_{N-S} represents the fixed transmission power of a node, and $\mathrm{RSSI}_{\mathrm{BS}-M}$ is the minimum RSSI value identified by a node. If the above formula is satisfied, the node can communicate with BS directly and $\mathrm{Flag}_{i-\mathrm{BS}} = 1$. Otherwise, $\mathrm{Flag}_{i-\mathrm{BS}} = 0$, where $\mathrm{Flag}_{i-\mathrm{BS}}$ is the flag that indicates whether the node s_i can communicate with BS directly.

A cluster-head node must have more energy and be closer from BS than a noncluster-head node to balance the network energy. On the other hand, the RTBDG algorithm is based on indoor WSN for the risk analysis of industrial operations. Thus, the computing speed of this algorithm should be as fast as possible. Therefore, RTBDG uses $s_i \cdot E$ (the residual energy of node s_i) and $\mathrm{RSSI}_{i-\mathrm{BS}}$ as the estimating attributes to achieve a balance between the algorithm performance and the speed. Each node calculates its own $Y_{i-\mathrm{CH}}$ using formula (2) with the above two attributes and weight coefficient $w_{E-\mathrm{CH}}$

$$Y_{i-\mathrm{CH}} = w_{E-\mathrm{CH}} \cdot s_i.E + (1 - w_{E-\mathrm{CH}}) \cdot \mathrm{RSSI}_{i-\mathrm{BS}}. \tag{2}$$

The focus of RTBDG when selecting a cluster-head node is on the node residual energy, and the focus of RTBDG when selecting a forwarding node or relay node is on the distance between the nodes and BS. Therefore, each node should calculate its own Y_{i-R} using formula (3) with the two attributes and weight coefficients w_{E-R}. The value of Y_{i-R} is used to select the forwarding or relay node

$$Y_{i-R} = w_{E-R} \cdot s_i.E + (1 - w_{E-R}) \cdot \mathrm{RSSI}_{i-\mathrm{BS}}. \tag{3}$$

2) Establishment phase of the node neighbor sets.

Each node broadcasts its own control information that consists of its ID $Y_{i-\mathrm{CH}}$, Y_{i-R} and $\mathrm{Flag}_{i-\mathrm{BS}}$ during its allocated time slot. Each node then monitors and receives the control information from its neighbor nodes. Thus, a set of neighbor nodes is formed and can be defined as follows:

$$\begin{cases} S_{i-N} = \{s_j \cdot \mathrm{ID}, s_j \cdot Y_{j-\mathrm{CH}}, s_j \cdot Y_{j-R}, \mathrm{Flag}_{j-\mathrm{BS}}\} \\ j = 1, \ldots, s_i \cdot N\_\mathrm{Nu} \end{cases} \tag{4}$$

where S_{i-N} represents the neighbor set of node s_i, $s_j \cdot \mathrm{ID}$ is the ID of node s_j that is the neighbor node of s_i, $s_j \cdot Y_{j-\mathrm{CH}}$ is the $Y_{j-\mathrm{CH}}$ value of node s_j, $s_j \cdot Y_{j-R}$ is the Y_{j-R} value of node s_j, $\mathrm{Flag}_{j-\mathrm{BS}}$ is the flag of node s_j that represents whether the node can communicate directly with BS, and $s_i \cdot N\_\mathrm{Nu}$ represents the neighbor node number of node s_i.

3) Selection of the cluster-head nodes and their forwarding nodes.

Each node compares its own $Y_{i-\mathrm{CH}}$ value with that of its neighbor nodes. If its own $Y_{i-\mathrm{CH}}$ value is the maximum, the node becomes the cluster-head node. If the cluster-head node satisfies $\mathrm{Flag}_{i-\mathrm{BS}} = 1$, it sends data to the BS directly. Otherwise, the neighbor node with the maximum Y_{j-R} value is selected as the forwarding node and its ID is saved. Each cluster-head node then sends a message (i.e., HEAD_MSG) that contains its ID and its forwarding node's ID to its neighbor nodes. The neighbor nodes of the cluster-head node receive the message. If a noncluster-head node does not receive any HEAD_MSG, it selects the neighbor node s_j with the maximum $Y_{j-\mathrm{CH}}$ value as the cluster-head node. The noncluster-head node then sends a message to node s_j, which receives the message. The node s_j then selects its forwarding node using the same method used by the cluster-head node. The node s_j also sends HEAD_MSG to its neighbor nodes.

4) Noncluster-head nodes select cluster-head nodes to which they will belong.

If a noncluster-head node with $\mathrm{Flag}_{i-\mathrm{BS}} = 0$ is not the forwarding node of a cluster-head node, it should decide which cluster it belongs to after receiving the message HEAD_MSG. The noncluster-head node joins the cluster whose cluster-head node has the maximum $Y_{i-\mathrm{CH}}$ value and then transmits a message (i.e., JOIN_MSG) that contains its ID to the selected cluster-head node. In addition, the nodes with $\mathrm{Flag}_{i-\mathrm{BS}} = 1$ communicate directly with BS and are not members of any cluster.

5) The forwarding nodes of the cluster-head nodes select their forwarding nodes.

If a noncluster-head node is the forwarding node of a cluster-head node, it selects its relay node. The candidate relay nodes include the nodes with $\mathrm{Flag}_{i-\mathrm{BS}} = 1$ and the cluster-head nodes found by this noncluster-head node except those that forward data to this noncluster-head node. If no candidate relay nodes around the forwarding node, all of the forwarding node's neighbors become the candidate relay nodes. After the candidate relay nodes are determined, the node with the maximum Y_{i-R} value

is selected as a relay node. The forwarding node then informs the selected relay node by sending a message.

6) Data are transmitted from the nodes to BS.

All nodes begin to transmit data after the cluster setup process is completed. Initially, each noncluster-head node without a data forwarding task transmits a data packet to its cluster-head node. After a cluster-head node receives all of the data, it begins to aggregate the data and then sends them to its forwarding node. After a forwarding node receives the data sent by its cluster-head node, the forwarding node needs to aggregate the data and send them to its relay node. Finally, all data packets are transmitted to the BS by the nodes with $\mathrm{Flag}_{i-\mathrm{BS}} = 1$. RTBDG then returns to the initial phase for the next round after all nodes have sent their data packets.

The cluster structure in RTBDG is built on the node residual energy and RSSI during the selection of the WSN operation. The network connectivity can then be ensured. However, if a sensor node runs out of energy or two communicating nodes can no longer communicate because of environmental changes, RTBDG should select the different strategies based on its application. If RTBDG is applied to the environment where real-time requirements are low, then RTBDG does not need to take any additional action. This condition is due to the communication interruption nodes being bypassed when building the cluster structure in the next round. The missing data have an insignificant impact on the current round. If the real-time requirements are high, RTBDG should set the time threshold when receiving data in the BS. The time threshold can be set according to the actual applications. For example, the time threshold can be shorter when equipment data are transmitted than when environment data are transmitted. Thus, the procedure of rebuilding the cluster structure comprises the following steps. 1) BS monitors whether the data have been received from all nodes. 2) If the waiting time reaches the time threshold, the information of rebuilding the cluster structure is sent to the entire network, and the network begins to build the cluster structure for a new round.

VIII. Experimental Results of the RTBDG Algorithm

A RTBDG algorithm called RTBDG is proposed in this paper on the basis of indoor WSN for the risk analysis of industrial operations. Experiments are conducted at the chip manufacturing laboratory of the Dalian Institute of Semiconductor Technology to evaluate the efficiency of the RTBDG algorithm. The chip manufacturing equipment and manufacturing environment are monitored in the experiments by the testing system of Section II, and the RSSI values are collected by the sensor nodes. The experimental parameters are shown in Table I and the distribution of sensor nodes is shown in Fig. 16. The RTBDG algorithms and three other classical routing algorithms: 1) LEACH [13]; 2) HEED [14]; and 3) WAIER [21] are used to transmit test data to the control center. The following criteria are then observed.

1) Network lifetime is defined as the time horizon until the first node dies. The network lifetime can reflect the energy efficiency of networks. A longer network lifetime corresponds to higher energy efficiency.
2) The standard deviation of all nodes' residual energy when the first node dies can reflect the balance of network energy. A smaller standard deviation corresponds to more balanced energy.

TABLE I
Experimental Parameters of RTBDG

Parameter	Value
Network size (x, y)	50 m × 40 m
BS location ($x/2$, y)	25 m, 0 m
Number of nodes(2*(x*y)/100)	40
Data packet size	4800 bits
Control packet size	200 bits
Data transmission frequency	240 bits/s
Initial energy	0.5 J
$w_{E-\mathrm{CH}}$	0.8
w_{E-R}	0.2

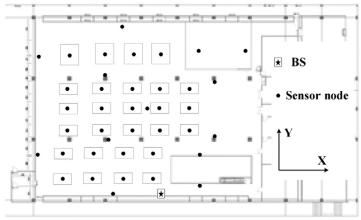


Fig. 16. Experiment environment.

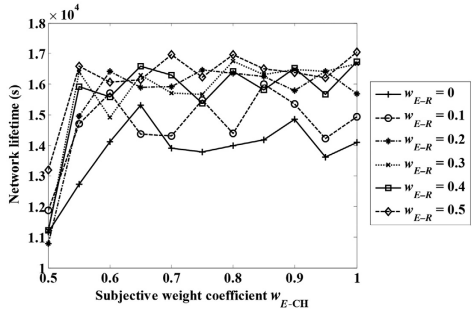


Fig. 17. Changes in the network lifetime with $w_{E-\mathrm{CH}}$ and w_{E-R}.

A. Analysis of the Weight Coefficients

Let the weight coefficient $w_{E-\mathrm{CH}}$ vary from 0.5 to 1 with a step length of 0.05, w_{E-R} vary from 0 to 0.5 with a step length of 0.1, and other parameters remain unchanged. The experimental results are shown in Figs. 17 and 18. Fig. 17 indicates that the network lifetime increases until $w_{E-\mathrm{CH}}$ reaches 0.55 and that the network lifetime slightly changes when $w_{E-\mathrm{CH}} > 0.55$. The network lifetime increases with w_{E-R} and slightly changes when $w_{E-R} > 0.2$. The standard deviations of all nodes' residual energy are shown in Fig. 18. Fig. 18 shows that the standard deviations fluctuate between 0.01 and 0.03 for any value of w_{E-R} when $w_{E-\mathrm{CH}} > 0.55$. The experimental results indicate that the RTBDG algorithm shows an improved performance when $w_{E-\mathrm{CH}} > 0.55$ and $w_{E-R} > 0.2$.

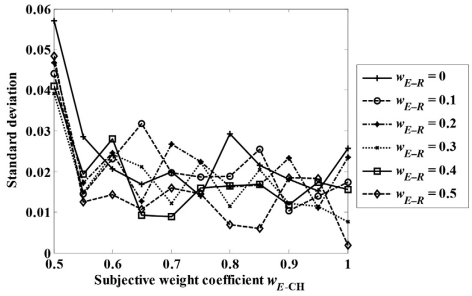


Fig. 18. Changes in the standard deviations of all nodes' residual energy with $w_{E-\mathrm{CH}}$ and w_{E-R}.

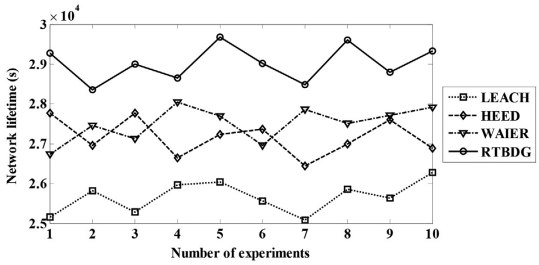


Fig. 19. Comparison of the network lifetime of the four algorithms.

B. Analysis of Network Energy Consumption

The most important problem in the design of the routing algorithm of WSN is the maximization of the network lifetime and balancing the energy consumption of sensor nodes. Fig. 19 illustrates the network lifetime of 10 experiments for the four algorithms, where the network lifetime of RTBDG is clearly the longest among all of the algorithms. Compared with the WAIER algorithm whose network lifetime is the longest among the other three algorithms, the proposed algorithm improves the average network lifetime by approximately 5.5%. The reason for this condition is that the cluster-head nodes consume more energy than the other nodes in the RTBDG algorithm, but they can be rotated in different rounds and be selected by the formulas with weight coefficients so that the energy efficiency is high. The network lifetime in the LEACH algorithm is the shortest among all algorithms. The reason for this condition is that the cluster-head nodes transmit data to the BS using the single-hop method, and the residual energy is ignored in LEACH. Furthermore, the amount of cluster-head nodes is large in LEACH, thus increasing the energy consumption. HEED and WAIER are between RTBDG and LEACH. The cluster-head nodes in the HEED algorithm are selected according to the residual energy and a secondary parameter of the sensor nodes. Thus, the energy expended by cluster-head nodes is lower than in LEACH. The WAIER algorithm uses the multiple attribute decision method to select the next hop for data transmission. Thus, the energy consumption is reduced. However, all data must be forwarded by the nodes that can communicate directly with BS so that they expend the energy rapidly, and these nodes will die earlier than the other nodes.

Fig. 20 illustrates the standard deviations of all nodes' residual energy for the ten experiments of the four algorithms. The standard deviations of RTBDG are clearly far lower than those of the other algorithms. This condition indicates that the energy

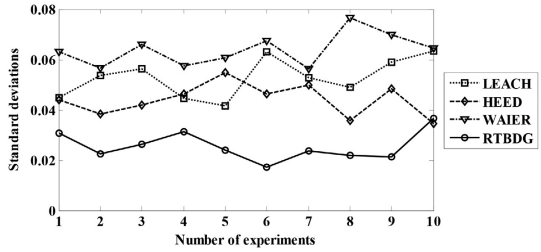


Fig. 20. Comparison of the standard deviations of all nodes' residual energy of the four algorithms.

expended by the sensor nodes in RTBDG is the most balanced among the four algorithms. Compared with the HEED algorithm whose standard deviations are the lowest among the other three algorithms, the proposed algorithm decreases the average standard deviations by approximately 42%. The reason for this condition is that a cluster-head node and its relay node can be selected by a comprehensive consideration of the path loss and residual energy of sensor nodes in RTBDG. Thus, the node energy can be dissipated uniformly. The energy consumption of nodes in WAIER is the least balanced. A relay node in WAIER is only selected from the set of forwarding neighbor nodes, so the nodes near BS are used excessively. Thus, the energy distribution is uneven.

IX. Conclusion

A large amount number of data is tested in this study to identify the WSN signal transmission characteristics in typical residence, office, and factory manufacturing environments. Two types of tests are conducted in every experimental scenario. The signal transmission characteristics of indoor WSN are summarized from the analysis of massive test data, and an energy-balanced big data gathering algorithm called RTBDG is proposed to gather real-time big data for risk analysis of industrial operations. In the algorithm, the RSSI values are used to determine whether two sensor nodes can communicate with each other instead of the distance between two sensor nodes. Thus, the first technical challenge in Section VI is solved. The RTBDG algorithm uses the round strategy and method of rebuilding the cluster structure according to the proposed conditions. Thus, the second technical challenge in Section VI is also solved. The performance of the proposed algorithm is compared with three other well-known algorithms. Experimental results show that the RTBDG algorithm can achieve high performance in the aspect of energy consumption and network lifetime for gathering big data in real time. Future works will concentrate on conducting practical applications of the RTBDG algorithm in manufacturing enterprises such as chip fabrication plants, pharmaceutical factories, and food factories.

References

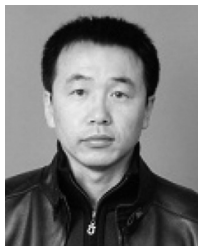
[1] S. Viaene, "Linking business intelligence into your business," *IEEE IT Prof.*, vol. 10, no. 6, pp. 28–34, Nov./Dec. 2008.

[2] J. Braca, J. Bramel, B. Posner, and D. Simchi-Levi, "A computerized approach to the New York City school bus routing problem," *IIE Trans.*, vol. 29, no. 8, pp. 693–702, 1997.

[3] H. K. Chan and F. T. S. Chan, "Early order completion contract approach to minimize the impact of demand uncertainty on supply chains," *IEEE Trans. Ind. Informat.*, vol. 2, no. 1, pp. 48–58, Feb. 2006.
[4] D. Boyd and K. Crawford, "Six provocations for big data," in *Proc. Decade Internet Time Symp. Dyn. Internet Soc.*, 2011, pp. 1–17.
[5] T. M. Choi, "Coordination and risk analysis of VMI supply chains with RFID technology," *IEEE Trans. Ind. Informat.*, vol. 7, no. 3, pp. 497–504, Aug. 2011.
[6] Z. Kremljak and C. Kafol, "Types of risk in a system engineering environment and software tools for risk analysis," *Proc. Eng.*, vol. 69, pp. 177–183, 2014.
[7] P. Giménez, B. Molina, J. Calvo-Gallego, M. Esteve, and C. E. Palau, "I3WSN: Industrial intelligent wireless sensor networks for indoor environments," *Comput. Ind.*, vol. 65, no. 1, pp. 187–199, 2014.
[8] G. M. Gaukler, "Item-level RFID in a retail supply chain with stock-out-based substitution," *IEEE Trans. Ind. Informat.*, vol. 7, no. 2, pp. 362–370, May 2011.
[9] *Big Data: Business Opportunities, Requirements and Oracle's Approach*, Winter Corporation, Cambridge, MA, USA, 2011, pp. 1–8.
[10] A. A. Kumar Somappa, K. Øvsthus, and L. M. Kristensen, "An industrial perspective on wireless sensor networks—A survey of requirements, protocols, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1391–1412, Third Quarter 2014.
[11] Y. Tian and Z. Tang, "Wireless meter reading based energy-balanced steady clustering routing algorithm for sensor networks," *Adv. Elect. Comput. Eng.*, vol. 11, pp. 9–14, 2011.
[12] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 30–40, Feb. 2012.
[13] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application specific protocol architecture for wireless microsensor networks," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 660–670, Oct. 2002.
[14] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad-hoc sensor networks," *IEEE Trans. Mobile Comput.*, vol. 3, no. 4, pp. 660–669, Oct./Dec. 2004.
[15] J. Chen, Z. Li, and Y. H. Kuo, "A centralized balance clustering routing protocol for wireless sensor network," *Wireless Pers. Commun.*, vol. 72, pp. 623–634, 2013.
[16] J. Peng, X. H. Chen, and T. Liu, "A flow-partitioned unequal clustering routing algorithm for wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2014, 12 pp., 2014, Article ID 875268.
[17] H. Lee, M. Jang, and J. W. Chang, "A new energy-efficient cluster-based routing protocol using a representative path in wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2014, 12 pp., 2014, Article ID 527928.
[18] Y. Tao, Y. Zhang, and Y. Ji, "Flow-balanced routing for multi-hop clustered wireless sensor networks," *Ad Hoc Netw.*, vol. 11, pp. 541–554, 2013.
[19] S. S. Wang and Z. P. Chen, "LCM: A link-aware clustering mechanism for energy-efficient routing in wireless sensor networks," *IEEE Sensors J.*, vol. 13, no. 2, pp. 728–736, Feb. 2013.
[20] H. Huang, W. Shi, L. Xu, X. Wang, and W. Zhong, "Weight coefficient adaptive based indoor energy load-balanced wireless sensor networks routing," *Acta Electron. Sin.*, vol. 38, pp. 2493–2498, 2010.
[21] K. Daabaj, M. Dixon, T. Koziniec, and P. Cole, "Reliable routing scheme for indoor sensor networks," in *Proc. IEEE 21st Int. Symp. Pers. Indoor Mobile Radio Commun.*, Instanbul, Turkey, 2010, pp. 1614–1619.
[22] A. E. Tümer and M. Gündüz, "Energy-efficient and fast data gathering protocols for indoor wireless sensor networks," *Sensors*, vol. 10, pp. 8054–8069, 2010.
[23] N. Li, S. McLaughlin, and D. Laurenson, "Traffic-aware routing for wireless sensor networks in built environment," in *Proc. 4th UKSim Eur. Symp. Comput. Modeling Simul.*, Washington, DC, USA, 2010, pp. 397–401.
[24] Z. Khan, N. Aslam, and S. Sivakumar, "Energy-aware peering routing protocol for indoor hospital body area network communication," *Proc. Comput. Sci.*, vol. 10, pp. 188–196, 2012.
[25] P. Padilla, J. Camacho, and F. G. Macia, "On the influence of the propagation channel in the performance of energy-efficient geographic routing algorithms for wireless sensor networks (WSN)," *Wireless Pers. Commun.*, vol. 70, pp. 15–38, 2012.
[26] A. Maskooki, C. Soh, E. Gunawan, and K. Low, "Adaptive routing for dynamic on-body wireless sensor networks," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 2, pp. 549–558, Mar. 2015.
[27] N. Marchenko, T. Andre, G. Brandner, W. Masood, and C. Bettstetter, "An experimental study of selective cooperative relaying in industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 3, pp. 1806–1816, Aug. 2014.
[28] R. Abrishambaf, S. N. Bayindir, and M. Hashemipour, "Energy analysis of routing protocols in wireless sensor networks for industrial applications," in *Proc. Inst. Mech. Eng. I J. Syst. Control Eng.*, vol. 226, no. 5, pp. 678–684, 2012.
[29] W. Shen, T. Zhang, F. Barac, and M. Gidlund, "PriorityMAC: A priority-enhanced MAC protocol for critical traffic in industrial wireless sensor and actuator networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 824–835, Feb. 2014.
[30] B. C. Villaverde, S. Rea, and D. Pesch, "InRout—A QoS aware route selection algorithm for industrial wireless sensor networks," *Ad Hoc Netw.*, vol. 10, no. 3, pp. 458–478, 2012.
[31] K. Yu, Z. Pang, M. Gidlund, J. Akerberg, and M. Björkman, "REALFLOW: Reliable real-time flooding-based routing protocol for industrial wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2014, 17 pp., 2014, Article ID 936379.
[32] J. Zhao, Y. Qin, D. Yang, and J. Duan, "Reliable graph routing in industrial wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 2013, 15 pp., 2013, Article ID 758217.
[33] P. T. A. Quang and D. S. Kim, "Enhancing real-time delivery of gradient routing for industrial wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 8, no. 1, pp. 61–68, Feb. 2012.
[34] A. Tiab and L. Bouallouche-Medjkoune, "Routing in industrial wireless sensor networks: A survey," *Chin. J. Eng.*, vol. 2014, 7 pp., 2014, Article ID 579874.

Xuejun Ding received the B.S. degree in electronic information engineering, and the M.S. degree in signal and information processing from Dalian University of Technology, Dalian, China, in 2001 and 2005, respectively, and the Ph.D. degree in technical economy and management from Dongbei University of Finance and Economics, Dalian, in 2014.

She is an Assistant Professor with the School of Management Science and Engineering, Dongbei University of Finance and Economics. Her research interests include wireless sensor networks, management information systems, and big data.



Yong Tian received the B.S. degree in applied electronic technology from the Wuhan University of Technology, Wuhan, China, in 1998, and the M.S. degree in signal and information processing, and the Ph.D. degree in microelectronics and solid state electronics from the Dalian University of Technology, Dalian, China, in 2005 and 2014, respectively.

He is an Associate Professor with the School of Physics and Electronic Technology, Liaoning Normal University, Dalian. His research interests include wireless sensor networks, Internet of Things, and big data.

Yan Yu received the B.S. degree in industrial automation from the Harbin University of Civil Engineering and Architecture, Harbin, China, in 1999, and the M.S. degree in architecture technique science, and the Ph.D. degree in disaster prevention and reduction engineering from the Harbin Institute of Technology, Harbin, in 2001 and 2006, respectively.

He is an Associate Professor with the Institute of Microelectronics, Dalian University of Technology, Dalian, China. His research interests include wireless sensor networks, data fusion, structural health monitoring, and intelligent buildings.