YONGCHUN HE
FENG ZHANG
XIALONG LI
and
YONG-HENGZHANG

Adaptive clustering routing optimization method for digital mine based on wireless sensor networks

For digital mine applications, it has the characteristics of diverse terrain, complex weather conditions, large monitoring area, large number of sensor nodes, multi-source information for each node, and long monitoring period.With the goal of good environmental adaptability, low power consumption, low cost and standardization, the key technology of wireless sensor network for digital mine is proposed, including network structure, networking mode, node location method, data fusion method, fast self-sufficiency and energy saving strategy. For the application of digital mine, the improved DV-Hop algorithm is proposed to locate the nodes, and then by analyzing the IoT-WSNs network model and the node energy consumption model, a new clustering routing algorithm is proposed. The experimental results show that the wireless sensor network technology designed in this paper satisfies the application requirement of digital mine well. Both hardware and software are convenient for system integration, and it is suitable for standardization and large-scale popularization. The proposed algorithm is effective and reliable.

Keywords: Digital mine, wireless sensor network, clustering routing algorithm, DV-Hop algorithm, IoT-WSNs network

1. Introduction

Digital mine is mainly comes from the social application of information technology such as large network, large concentration and large mobile. It is the result of information technology from single application to multiple fusions. It is the manifestation of the complex analysis of information technology from the front end to the front end. It is the inevitable product of the high information technology of the society[1]. Agriculture is the endless source of big data and has huge data base[2]. In the process of data acquisition, transmission, and analysis of the agricultural IoT still faces many challenges, such as data acquisition and transmission, sensor power consumption, network fault tolerance[3], and self-adaptation. In the analysis and decision-making environment, mass sensor data how to analyze mining and other issues.

Messrs. Yongchun He* and XiaLong Li, Mathematics and Statistics, Yulin University, Yulin 719000, Shaanxi and Feng Zhang and Yong-Heng Zhang, School of Information Engineering, Yulin University, 719000, Yulin, China. *Email of the corresponding author: 244695902@qq.com

In the sensor network, location information is crucial to the monitoring activity of the sensor network[4], and it is related to the information transmission of the entire network. There are many changes in the location of digital mine nodes, and the constraints of node cost, power consumption and hardware conditions of wireless sensor networks[5]. How to achieve low cost, low power consumption and high precision positioning is one of the key points in the research of large-scale wireless sensor network system for agriculture.

In this paper, an improved algorithm is proposed to locate the nodes for the field application. For the multi-source information terminal of digital mine, a fast self-feed method based on solar energy is designed, and the working mode of intermittent sampling and deep dormancy is designed, which effectively increases the power supply reliability of the nodes.

2. A node-oriented low-cost positioning method for precision agriculture

The sensor nodes in the WSN are usually deployed inside or near the data collection object. The sensors form a wireless network in an ad hoc manner. The information in the entire area can be sent to the aggregation node through the multi-hop network and further sent to the remote management and control center by the aggregation node. For further analysis and decision making in field planting, dynamic real-time monitoring of plants, soil, and environment has been achieved. WSN-based automatic control drip irrigation systems have been designed and put into pilot applications. In the process of precision agriculture and modern digital mine, how to use information technology to obtain information on the growth environment of crops, thereby effectively increasing resource utilization and production has become an important issue in digital mine.

2.1. IoT-WSN and Energy Consumption Model

We use G = (V,Y(BS),E) to represent a IoT-WSN, this IoT-WSN contains V heterogeneous sensor nodes and a resource-rich Sink node. It collects corresponding environmental information from sensor nodes, such as temperature, humidity, and soil moisture, which are often collected in digital mine. Each sensor node is equipped with a solar panel, so that the sensor can use solar energy to supplement battery power. In the data collection process, the data collection range of each node is fixed. If the distance between two sensor nodes or between the sensor node and the Sink node is smaller than the respective transmission range, there is a communication link between the two nodes.

Where the whole IoT-WSN is divided into several clusters, each cluster consists of cluster head nodes and cluster nodes. The main task of cluster nodes is to collect the information of the environment and send the collected information to the cluster head nodes[6]. The main task of the cluster head node is to receive the data of the cluster nodes and transmit these data to the Sink nodes through direct communication or multi hop. At the same time, when multi hop transmission is adopted, the cluster head node is also responsible for the relay task of data transmission for the corresponding cluster[7]. The task of Sink node is to collect data from the entire network and provide it to users, and at the same time responsible for complex computation and decision making.

The energy consumed by sensor nodes to transmit k bit data to nodes with a distance of *d* can be expressed as:

$$E_{tx}(k,d)=\begin{cases}E_{tx\_elec}\times k+\varepsilon_{fx}\times d^{n}\times k\\ E_{tx\_elec}\times k+\varepsilon_{mf}\times d^{n}\times k\end{cases} \tag{1}$$

The energy consumed by a sensor node to receive k bits of data is:

$$E_{RX}(K)=E_{RX\_elec\times k} \tag{2}$$

2.2 Node position calculation method

The three side measurement method is a common method for calculating node locations.The idea of the algorithm of the three side measurement method is:

The coordinates of the reference nodes are known to calculate the locations of the unknown nodes. It is known that the coordinates of the three nodes of A, B, and C are (x_a,y_b), (x_b,y_b), (x_c,y_c) respectively. The distance of the unknown node D from the three anchor nodes A, B, and C is d_a, d_b, d_c respectively. With the shop node as the canter and the distance value as the radius, three circles are drawn. The intersection of these three circles is the unknown node D(x,y) position. Fig. 1 shows a schematic of a trilateration method.

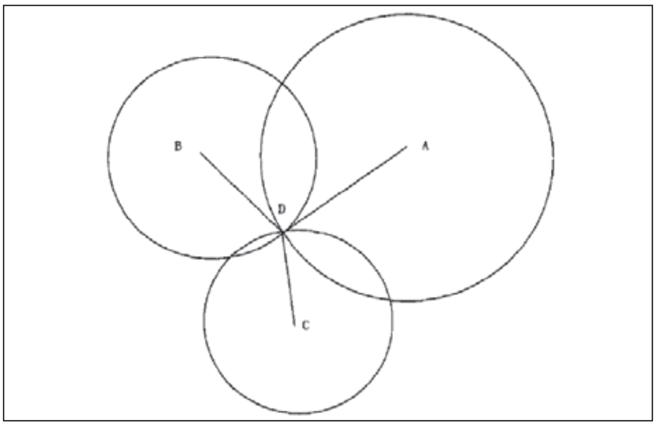


Fig. 1: A schematic diagram of three side measurement

Then, there are the following formulas:

$$\begin{cases}\sqrt{(x-x_a)^2+(x-x_a)^2}=d_a\\ \sqrt{(x-x_b)^2+(x-x_b)^2}=d_b\\ \sqrt{(x-x_c)^2+(x-x_c)^2}=d_c\end{cases} \tag{3}$$

The coordinates of D can be obtained by the formula (3).

$$\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}2(x_a-x_c)2(y_a-y_c)\\ 2(x_b-x_c)2(y_b-y_c)\end{bmatrix}^{-1}\begin{bmatrix}x_a{}^2-x_c{}^2+y_a{}^2-y_c{}^2+d_c{}^2-d_a{}^2\\ x_a{}^2-x_c{}^2+y_b{}^2-y_c{}^2+d_c{}^2-d_b{}^2\end{bmatrix} \tag{4}$$

2.3 Location Algorithm

By using the empirical model at different transmission distance, antenna height, different crop height, crop density, and different terrain, the radio frequency signal attenuation and packet loss rate under 433MHz wireless channel are analyzed, and the connectivity of information transmission is calculated.

$$CM_{k,a}=\frac{n_{recv}(a_i,t)}{n_{send}(a_i,t)}100\% \tag{5}$$

$n_{recv}(a_i,t)$ is in the time t period, the number of signals received by the node k is transmitted by the beacon node a_i.$n_{send}(a_i,t)$ is the number of signals sent by the anchor node a_i during the time period t. Take the threshold = 90%. The node with connectivity $CM_{k,a}$ > 90% serves as a neighbour node.

In the distance vector hopping location algorithm, the unknown node first calculates the minimum number of hops with the anchor node, then estimates the average distance per hop, and uses the minimum hop count times the average hop distance to obtain an estimate between the unknown node and the pin node[8]. Distance, then use the trilateration method or maximum likelihood estimation method to calculate the coordinates of unknown nodes. The beacon node broadcasts a packet of its own location information to the neighbouring node. The packet includes a hop segment number field, and the initial value is 0. The receiving node records the minimum number of hops for each beacon node, while ignoring the larger number of hops from the same pin node, and then transfers the value to 1 and forwards it to the neighbor node[9]. By this method, the minimum hops of all beacon nodes in the network are recorded. Each pin node is based on the location information and distance hopping of other wrong nodes recorded in the first stage. Then we can write a formula to estimate the average distance per hop.

$$HopSize_i = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{H_j} \tag{6}$$

Where $(x_i - y_i)$, $(x_j - y_j)$ are the coordinates of the beacon node i, j, H_j, is the number of hops between the beacon node i and the j. The average distance per hop calculated by the beacon node is broadcast to the network by the packet with the lifetime field. The unknown node only records the first received average distance per hop and forwards it to the neighbour node. This strategy ensures that most nodes receive the average distance per hop from the nearest wrong node. After receiving the average distance per hop, the unknown node calculates the hop distance to each beacon node according to the recorded hop count.

2.4 Improved distance vector hopping algorithm

The distance vector hopping location algorithm uses the communication radius of the node as the average distance per hop, and the positioning error is large. When the nodes are not dense enough, the number of hops from the unknown node to the beacon node increases, and the error between the obtained hop distance and the actual distance increases. In order to solve the problems in practical applications and improve the accuracy, according to formula (6), each beacon node can get an average distance of $HopSize_j$ per hop. Through flood routing, the unknown node obtains the average hop distance of each wrong node and weights the distance value to obtain the final average hop distance.

(1) Assume that the number of hops from the unknown node to each beacon node is h_1, h_2, h_3, ..., h_n, then the weighted value of the average hop distance of each wrong node is recorded as:

$$w_i = \frac{h_i}{\sum_{i=1}^{n} h_i} \tag{7}$$

Calculate the final average hop distance of unknown nodes, denoted as Q,

$$Q = \sum_{i-1}^{n} w_i HopSize_i \tag{8}$$

If we add a wrong node to participate in the positioning based on the three-side ranging, as shown in Fig. 2, then according to the three-side ranging the principle is to take three wrong nodes to estimate the coordinates of the unknown node each time, and the four wrong nodes obtain a total of four groups of position estimates. The coordinates of B, C, and D points for E points are $E_1(x_{E_1}, y_{E_1})$, A, B, C get the E point coordinate estimation is $E_2(x_{E_2}, y_{E_2})$.

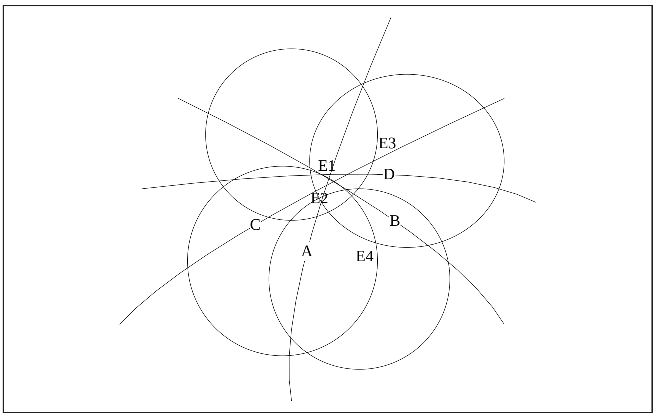


Fig. 2: Quadrangular location and distance measurement

3. Adaptive clustering routing optimization method for digital mine

Because the energy obtained by the nodes from the outside is unstable and unbalanced, the sensor nodes that are cluster nodes adopt different sampling frequencies so that the purpose of sustainable work can be achieved, under this premise, the sampling frequency of the cluster head node can be determined by the correlation algorithm. However, sensor nodes that are cluster head nodes need to periodically transmit the collected data to Sink nodes. Similar to the clustering algorithm of traditional WSNs, it also works in the form of rounds, but each round is divided into cluster establishment phase and data transmission phase. Because the cluster head node is not only responsible for the collection, processing, and transmission of cluster node data within the cluster, it also needs to be responsible for relaying tasks when other clusters need to transmit data, especially cluster head nodes close to Sink nodes, so the cluster head node plays an important role in the clustering routing of IoT-WSN. It is specially important to establish a proper cluster to balance the energy consumption among cluster heads. Based on the characteristics of IoT-WSN, we design a new algorithm for cluster establishment process is as follows:

Step 1: For all nodes we can set Net(i) = 0, Kind(i) = 0;

Step 2: BS sends out partition_cluster() signals, node i calculates its distance based on partition_cluster() and updates H_j, Q and E_i;

Step 3: Node i calculates the average residual energy of its neighbor nodes Q;

$$f(E_1(x_{E_1}, y_{E_1})) > E_2(x_{E_2}, y_{E_2})$$

$$kind(i) = \sum_{i-1}^{n} w_i HopSize_i$$

endif

Step 4: BS sends out CH_select() signal, after receiving CH_select(), the candidate cluster head starts the timer T_i and monitors other packets.

$if\,(T_i < w_i)$

$if\,(Received\ cluster\ head\ contention\ information$

$CH\_select(j, net(i))\ from\ node\ j)$

$if\,(Net(j) = Net(i))$

$Storage\ node\ j's\ ID\ and\ close\ the\ timer\ T_1;$

$kind(i) = HopSize_i;$

$else$

$Discarding\ the\ packet\ and\ continuing\ to\ monitor\ it;$

$endif$

$endif$

$endif$

$if\,(T_i = w_i)$

$Send\ out\ cluster\ head\ competition\ information\ Q$

$kind(i) = Q;$

$endif$

At the initial stage of cluster establishment, the network gradient values and node types of each node are set to 0 respectively. Then the Sink node sends out the clustering signal partition_cluster(). As mentioned earlier, the transmission radius of the Sink node can cover the entire IoT-WSN area, so every node can receive partition_cluster().

4. Experimental analysis

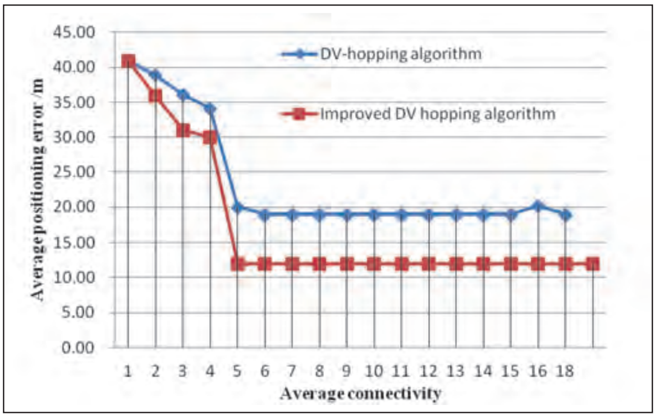
During the experiment, a mine base in Yulin was selected as the test area.The communication radius of the node is 120m, and 50 nodes are randomly deployed in the test area. The position information of the sensor nodes is tested 150 times, and the experimental data are the average of the 150 experiments. Using different algorithms, the location error results are compared under different connectivity and anchor nodes ratio. The simulation results are shown in Fig. 3. As is observed from Fig. 3, we can see that the improved DV-Hop achieves fairly good results in the average positioning error of unknown nodes, and improves the accuracy of the algorithm.

In order to verify the performance of the improved DV-Hop algorithm, simulation is carried out in MATLAB R2010a.The machine used in the simulation is: Windows 10 operating system, Intel (R) Core (TM) i7 processor, 8GB memory. The simulation environment is as follows: 200 rechargeable sensor nodes are deployed in the two-dimensional space of 200m * 200m, and the Sink nodes are located at (50m, 50m). Other related parameters are set as shown in Table 1.

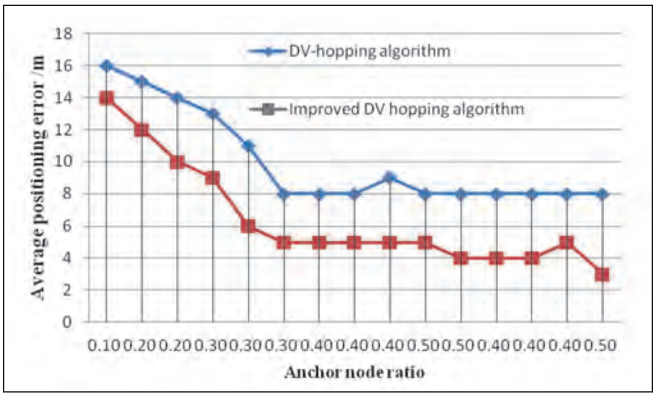
Due to the important role of cluster head nodes in clustering routing, the quality of cluster head nodes is first compared. The ratio of the residual energy of cluster head nodes to the average residual energy of all nodes of the cluster is used to judge the quality of cluster head nodes. In order to reflect the effectiveness of comparison, 15 clusters are randomly selected from IDV-Hop and DV-Hop for comparison, and the comparison results are shown in Fig. 4.

TABLE 1. PARTIAL SIMULATION PARAMETER SETTING

Parameter	Set value
$E_1(x_{E_1}, y_{E_1})$	10.40
$E_2(x_{E_2}, y_{E_2})$	12.31
$HopSize_1$	45.98
CH_select	130.343
Net(1)	12.21
kind(1)	56.29



(a) Relationship between mean location error and average connectivity



(b) Relationship between average positioning error ratios.

Fig. 3: Simulation result

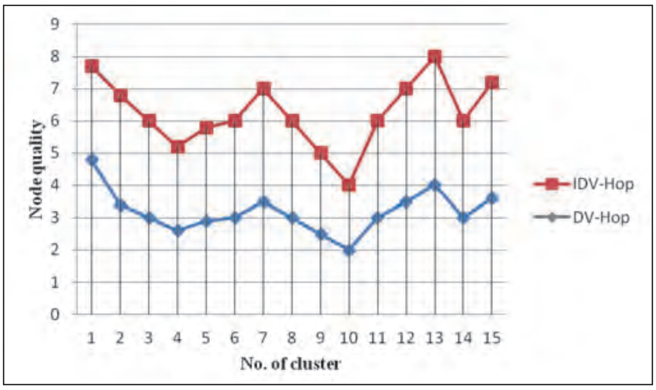


Fig. 4: Quality comparison of cluster head nodes

5. Conclusions

In this paper, with the goal of good environmental adaptability, low power consumption, low cost and standardization, the key technology of wireless sensor network for digital mine is studied in digital mine, including network structure, networking mode, node location method, data fusion method, energy fast self-sufficiency and energy saving strategy. The performance evaluation method of wireless sensor network is also introduced. Based on the different energy absorption efficiency and limited battery capacity of the rechargeable sensor nodes for the application of the digital mine network, a new clustering routing protocol is proposed for IoT-WSNs.

Acknowledgements

This work is partially supported by the Natural Science Basic Research Plan in Shaanxi Province of China (2017NY-089, 2017NY-134, 2016NY141, 2016KJXX-62,2017NY132),Funding Project for Department of Yulin University (16GK24), and thanks for the help.

References

1. Sharma T, Tomar G S, Gandhi R,et al. (2015): Optimized Genetic Algorithm (OGA) for Homogeneous WSNs, *International Journal of Future Generation Communication and Networking*, 8(4): 131-140.
2. F Zhang, H-F Xue, D-S Xu. (2013): Big data cleaning algorithms in cloud computing, *International Journal of Online Engineering*, 9(3): 77-81.
3. Hussain S, Matin A W, Islam O. (2007): Genetic algorithmforhierarchical wireless sensor networks, *Journal of Networks*, 2(5): 87-97.
4. Gupta S K, Jana P K. (2015): Energy Efficient Clustering and Routing Algorithms for Wireless Sensor Networks: GA Based Approach, *Wireless Personal Communications*, 83(3): 2403-2423.
5. Long C, Zhou X, Liao S, et al. (2014): An improved LEACH multi-hop routing protocol based on genetic algorithms for heterogeneous wireless sensor networks, *Journal of Information & Computational Scienc*e, 11(2): 415-424.
6. Zhang Yong-Heng, Zhang Feng. (2013): A New Time Synchronization Algorithm for Wireless Sensor Networks Based on Internet of Things, *Sensors and Transducers*, 151(4):95-100.
7. JI Yan, Zhang Feng. (2013): Wireless Sensor Traceability Algorithm Based on Internet of Things in the Area of Agriculture, *Sensors and Transducers*, 151(4):101-106.
8. Cao X, Chen, Sun Y. (2009): An Interface Designed for Networked Monitoring and Control in Wireless Sensor Networks, *Computer Standards and Interfaces*, 31(3):579-585.
9. Lazzari. (2011): Caetano Decian. Wireless crankarm dynamometer for cycling, *Sensors and Transducers*, 128(5):39-54.

WATER RESOURCES BIG DATA CLASSIFICATION BASED ON MULTI-OBJECTIVE OPTIMIZATION FOR MINING AREA

Continued from page 723

2. F Zhang, H-F Xue, D-S Xu. (2013): Big data cleaning algorithms in cloud computing, *International Journal of Online Engineering*, 9(3), 77-81.
3. X Li, L Zhang and X Cao et al. (2016): Retrieval of precipitable water vapor using MFRSR and comparison with other multisensory over the semi-arid area of northwest China, *Atmospheric Research*, 172(1),83-94.
4. Shao Xin-guang. (2006): Parameters selection and application of support vector machines based on particle swarm optimization algorithm, *Control Theory & Applications*, 23(5),740-743.
5. Yuan Xiao-yan. (2007): Kernel parameter selection of support vector machine based on particle swarm optimization, *Automatic technique & Applications*, 26(5), 5-8.
6. Li Zhao-Xing. (2014): A multi objective optimization algorithm for recommender system based on PSO, *Computer Modelling and New Technologies*, 18(7),231-235.
7. Zhou SH L, Liao J, Shi J X. (2014): Kernel parameter selection of RBM-SVM and its application in fault diagnosis, *Journal of Electronic Measurement and Instrument*, 28(3): 240-246.
8. Rahi, O.P. (2011): Optimization of hydro power plant design by Particle Swarm Optimization (PSO), *Procedia Engineering*, 30(1), 418-425.
9. Lian K, Chen Sh J, Zhou J M, et al. (2009): Study on a GA-based SVM decision-tree multi-classification strategy, *Control and Decision*, 2009(1), 7-12.
10. Wang R, Wong S, Chen D, et al. (2013): A vector-valued support vector machine model for multiclass problem, *Information Sciences*, 2013(235): 174-194.

Printed by Pradip Kumar Chanda at The Indian Press Pvt. Ltd. 93A Lenin Sarani, Kolkata 700 013 and published by him for Books & Journals Pvt. Ltd. from 6/2 Madan Street, Kolkata 700 072.