A Cross Unequal Clustering Routing Algorithm for Sensor Network

Wang Tong[1], Wu Jiyi[2], Xu He[3], Zhu Jinghua[4], Charles Munyabugingo[1]

[1]College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
[2]Key Laboratory of E-Business and Information Security, Hangzhou Normal University, Hangzhou 311121, China
[3]College of Mechanical and Electrical, Harbin Engineering University, Harbin 150001, China
[4]School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China

In the routing protocol for wireless sensor network, the cluster size is generally fixed in clustering routing algorithm for wireless sensor network, which can easily lead to the "hot spot" problem. Furthermore, the majority of routing algorithms barely consider the problem of long distance communication between adjacent cluster heads that brings high energy consumption. Therefore, this paper proposes a new cross unequal clustering routing algorithm based on the EEUC algorithm. In order to solve the defects of EEUC algorithm, this algorithm calculating of competition radius takes the node's position and node's remaining energy into account to make the load of cluster heads more balanced. At the same time, cluster adjacent node is applied to transport data and reduce the energy-loss of cluster heads. Simulation experiments show that, compared with LEACH and EEUC, the proposed algorithm can effectively reduce the energy-loss of cluster heads and balance the energy consumption among all nodes in the network and improve the network lifetime.

Keywords: Sensor networks, routing algorithm, energy balance.

1. Introduction

THE RAPID DEVELOPMENT of the Wireless communications and low power consumption embedded system gives birth to Wireless Sensor Network (WSN). And its characteristics of low-cost, low-power consumption, and self-organization have brought a revolution of information perception. WSN is composed of a large number of tiny sensor nodes deployed in monitoring areas, and forms a multi-hop self-organizing network through wireless communication. Its potential applications have a variety of forms such as environmental monitoring, target tracking and battlefield surveillance, etc. [1] Sensor nodes have three functional modules named perception, data processing, and transmission. Compared with other networks, sensor network's energy, computing, and storage ability is limited. Therefore, the design of effective routing algorithms, and efficient use of node energy to extend the network lifetime is becoming a focus of the WSN research.

In Wireless Sensor Network routing protocol, hierarchical routing algorithm based on clustering was considered one of the most favorable energy-saving methods. Clustering routing algorithm uses nodes within the network to organize the structure of cluster, the nodes in cluster are responsible for collecting data and the cluster head is responsible for data forwarding of the final data through a single hop or multi-hop way to reach the base station. In addition, recent studies show that the multi-hop communication between cluster head and aggregation node is more conducive to save energy [2].

In this paper we present a cross non-uniform clustering routing algorithm, which is based on the distance between the base station and how much the node remaining energy have to establish clusters of varying sizes. The closer distance to the base station, the lower the energy and the smaller the radius of the cluster. Cluster head can reserve more energy for the data forwarding. In the data transfer process, the cluster adjacent node is used to forward the inter-cluster data to avoid long-distance communication between clusters. Simulation results show that in the way of multi-hop communication the routing algorithm can effectively solve the problem of too fast consumption of cluster head energy, balanced network node energy loss and significantly extend the network lifetime.

2. Related work

In recent years, many clustering routing algorithms for WSN have been proposed. In order to avoid the cluster head consume, too much energy was spent and communication traffic was reduced The low energy adaptive clustering hierarchy algorithm (Low Energy Adaptive Clustering Hierarchy, LEACH) is presented in [3]. It uses circular random clustering methods and each node in the network can be a cluster head in rotation, this makes the energy-carrying of the network balancing to each node, extending the lifetime of the network, but LEACH uses a single hop communication mode, the cluster head and the base station can communicate directly, this causes a large cluster head communication overhead. Studies have shown that in the data transmission process, using multi-hop communication mode between the cluster head and base station is more conducive to save energy [4]. However, it can easily cause the cluster head, which is near the base station, undertake too much task of forwarding and always lead to a premature death. This phenomenon that makes the network energy consumption extremely uneven is called the "hot spot" [5] problem. In order to solve this, an uneven clustering algorithm, called EEUC (Energy-Efficient Uneven Clustering) [6] is proposed, this algorithm is calculated according to the distance of the candidate cluster head to the base station, the network is divided into different size classes of non-uniform structure of clusters by geographical location in order to balance the load of cluster head. But in

the cluster head election, the residual energy of the nodes is not considered, and the iterative phenomenon can be raised easily in the clustering process, and the overhead of clustering is relatively large. Thereafter people have been oriented on these algorithms. Taking the deployment model into consideration, the ant colony algorithm is applied [7] to optimize the energy consumption of each node in the network. Although the simulation shows that it can prolong the survival time of the network, manual deployment limits its application in a real environment. An adaptive clustering algorithm based on load balance is proposed in [8], the cluster radius, residual energy, and the spacing between cluster heads have been considered in the cluster head selection. The residual energy of neighboring nodes has been taken into account in [9]. The cluster head is selected by the node which has most residual energy, and the size of the cluster is also limited [10].Then, according to the principle of energy balance network, EBCA (Energy Balancing Clustering Algorithm) [11] sets the node energy threshold to select a cluster head, and to certain extent, it can balance the overall energy consumption of the network; but it is not the optimal path selection and it is relatively complex. In order to avoid the energy hole, the transmission way inspired by residual energy cooperation has been put forward [12], it just points at the "hot spot" problem. However, the long distance communication problems when transmitting data between clusters have not been considered in these routing algorithms.

In order to solve the above problems, a Cross Unequal Clustering Routing Algorithm for Sensor Network (called CUCRA) is proposed in this paper. Compared with the EEUC algorithms, CUCRA has improved EEUC and has the following advantages: (1) the residual energy of nodes is introduced when EEUC competition radius is computed in order to effectively balance the energy loss of the nodes within the network; (2) it enhances the EEUC cluster construction process so as to improve the converge speed; (3) cluster crossing nodes between the clusters are used to forward data, which can successfully reduce the long-distance communication spending, thus prolonging the lifetime of the whole network.

3. SYSTEM MODEL AND ANALYSIS OF EXISTING PROBLEMS

A. c System model

Above all, the Network Model is introduced. Assume that the wireless sensor network consists of N nodes. Nodes are distributed in the area of a fixed size. Assume the i-th node is s_i, and the node set $S = \{s_1, s_2, s_3, ..., s_{N-2}, s_{N-1}, s_N\}$, $|S| = N$ .In addition, the wireless sensor network has the following properties:

1) The network has only one base station or sink node located outside the sensing area. After the deployment of the base station and nodes, they cannot be moved.
2) All nodes have the same properties and functions. Each node has a unique identifier.
3) Node does not need to have the GPS function to obtain precise location information.

4) Wireless transmit power can be controlled, a node can adjust the size of the transmit power based on distance.
5) The link is symmetrical. If transmit power is given, the sensor node can calculate the distance from the transmission source based on signal strength.

In this paper, the wireless channel model in [13] - Radio Hardware Energy Dissipation - is applied. The model assumes a threshold d_0 . And if the distance d between nodes is shorter than d_0, then the energy consumed by the node for sending data is proportional to the square of the distance between two nodes. If $d > d_0$, the energy consumed by the node to send data is proportional to the fourth power of distance between two nodes. These two different energy loss calculation models are known as the free space model and multi-path fading model.

Node A to node B sends l bytes of data, the distance between two nodes is d , the energy loss is:

$$
\begin{aligned}
E_{Tx}(l,d) &= E_{Tx-elec}(l) + E_{Tx-amp}(l,d) \\
&= lE_{elec} + l\varepsilon_{amp}f(d) \\
&= \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, d < d_0 \\ lE_{elec} + l\varepsilon_{mp}d^4, d \geq d_0 \end{cases}
\end{aligned} \tag{1}
$$

Node B receives the energy consumed by the l bytes of data:

$$
E_{Rx}(l,d) = E_{Rx-elec}(l) = l * E_{elec} \tag{2}
$$

Threshold d_0 determined by (3) is as follows:

$$
d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}} \tag{3}
$$

In the above equation, $E_{Tx}(l,d)$ means the energy consumed by sending l bits of data to the distance d away from its position; $E_{Tx-elec}(l)$ means the energy consumed by starting the wireless transmitter circuit; $E_{Rx-elec}(l)$ means energy consumed by the Receiver circuit; $E_{Tx-amp}(l,d)$ means the energy consumed by enlarging the transmission signal; E_{elec} means energy consumed by transmission circuit and receiving circuit to send or receive per-bit data; ε_{amp} means multiple of the signal amplifier; $\varepsilon_{fs}d^2$, $\varepsilon_{mp}d^4$ means radio power amplifier loss.

B. Analysis of existing problems

Because wireless sensor network node energy is limited, how to make efficient use of node energy so as to extend the network lifetime has become an important indicator of WSN routing designs [14]. However, current routing algorithms have more or less the following three issues.

First, the general clustering routing algorithm is divided into two processes: Cluster building process and data transfer process [15]. Even if there is unequal clustering, only the distance from the base station or node residual energy is considered for one situation in the formation of cluster size. Thus, the cluster size does not possess the flexibility and dynamic nature resulting from uneven network energy loss.

Second, according to the wireless transmission model, if the transmission distance is greater than d_0, the energy loss is the fourth power of distance. During data forwarding between clusters, if distance between cluster head is greater than d_0, it will result in long-distance communication between the clusters, cluster head energy is lost quickly, and this loss affects the cluster and the whole network and hence, shortening the network lifetime.

Third, in the WSN, node's ability of computing, storage, and its energy are limited. Existing algorithms are too complex in routing address, while nodes with high computational and storage capacity are required.

4. CROSS UNEQUAL CLUSTERING ROUTING MECHANISM

For the above problems, CUCRA has taken some action, which will be discussed here. The CUCRA is composed of three parts: hierarchy network structure establishment, Clustering mechanism, and Multi-hop transmission.

A. Hierarchy network structure establishment

After sensor nodes are deployed in the sensing area, base stations broadcast a "hello" message across the sensing area and the entire network is divided into the domain of m equivalent width with different level. The closer to the base station the higher level of the domain, as shown in Fig.1.
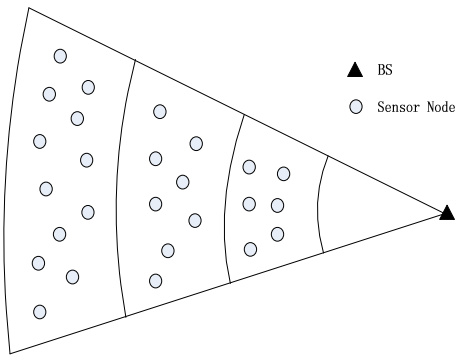

Fig.1. CUCRA domain structure

The Node calculates the distance to the base station according to the received signal strength, and it determines their domain. The Base station according to (4), (5) in order to calculate the upper and lower boundaries of each domain:

$$UB_i = d_{min} + i \times \frac{d_{max} - d_{min}}{m}, i = 1, 2 \ldots m \qquad (4)$$

$$LB_i = d_{min} + (i-1) \times \frac{d_{max} - d_{min}}{m}, i = 1, 2 \ldots m \qquad (5)$$

UB_i is the upper boundary of the domain i, LB_i is the lower boundary of the domain i. d_{min}, d_{max} are the shortest and the longest distance between sensing area and the base station, respectively. Assume that the distance between node j and the base station is d_j, if $LB_i < d_j < UB_i$, then the node belongs to the domain i.

B. Clustering mechanism

In the clustering process, the algorithm uses the mechanism of competition based on the distance to the base station and remaining energy, the nodes in the network calculate the competition radii which are also known as cluster radii. The competition radius of node i is given by the following equation.

$$R_{cmp} = \left[\varepsilon_1 \left(1 - c \times \frac{d_{max} - d(s_i, BS)}{d_{max} - d_{min}} \right) + \varepsilon_2 \times \frac{E_{residual}}{E_{init}} \right] \times R_0 \qquad (6)$$

Among them, d_{min} and d_{max} are the shortest and longest distance between sensing area and the base station, respectively. $d(s_i, BS)$ is the distance from node i to the base station. c, ε_1, ε_2 is a constant between 0 and 1. R_0 is the system setting of the cluster radius. From (6) can be seen that the cluster radius R_{cmp} of node varies with residual energy of node and the distance between nodes to the base station. That is to say, the closer it is to base station, the lower is the residual energy and the smaller is the radius of the clusters.

The clustering algorithm process for CUCRA will be specifically explained according to the following pseudo-code and Fig.2. In the first stage, the node generates a random number between 0 and 1. If the random number is less than T, then the node is chosen to become a candidate. Then, in the next stage, the candidate sends a message of (COMPETE_HEAD_MSG) in its competitive range, which parameters include location, competition radius and residual energy of node cluster head according to the (6). In stage three, if node i can receive a message of (COMPETE_HEAD_MSG) from node j, and the distance between both nodes is less than any competition radius of each other, then node i stores node j in S_{ct}. In stage four, the node i compares its own energy with the energy of each node in S_{ct}. If the energy of node i is larger than any node stored in S_{ct}, then the node i becomes a cluster head and releases a message of FINISH_ELECT_MSG within its cluster radius; otherwise the node i is electoral defeated. In stage five, after cluster head election, the Non-cluster head which received a message of (COMPETE_HEAD_MSG) chooses to join the nearest cluster head according to the information within (COMPETE_HEAD_MSG); if there is a Non-cluster head which received more than one message of (COMPETE_HEAD_MSG), then it selects one with the largest signal strength as its own primary cluster head and selects the cluster head with lower than it's own signal strength as its own sub-cluster head which will be a crossing node for two adjacent clusters, so it forms the CUCRA

clustering structure in Fig.2. where the solid node is the cluster head, the hollow one is the normal node and the star is the crossing node.
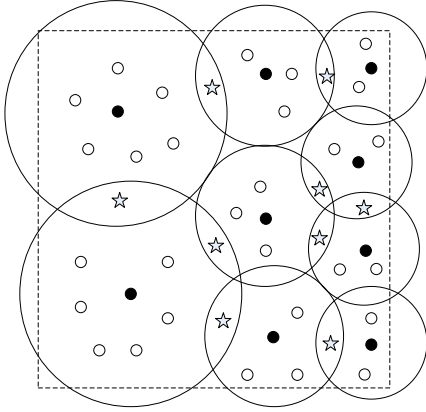

Fig.2. CUCRA clusters structure

The pseudo - code of CUCRA clusters is as follows:

```
1. α ← rand(0,1)
2. if α < T then
3.   beCandidateHead ← True
4. endif
5. if beCandidateHead = True then
6.   Send(COMPETE_HEAD_MSG)
7. endif
8. S_i receiving a COMPETE_HEAD_MSG from node S_j
9. if d(s_i,s_j) < s_i.R_cmp and d(s_i,s_j) < s_j.R_cmp then
10.   Add s_j → s_i.S_ct
11. endif
12. while beCandidateHead = True then
13.   if s_i.ResidueEnergy > s_j.ResidueEnergy,
        ∀s_j ∈ S_ct then
14.     s_i.Send(FINISH_ELECT_MSG) and Exit
15.   endif
16. end while
```

C. Multi-hop transmission

After Network cluster class structure is established, each cluster head in the local environment establishes a set of adjacent cluster heads G_{ch} and a collection of crossing nodes in the cluster C_{nodes}. Multi-hop communication between layers for data transmission is proposed in CUCRA. The cluster inner node will send the collected data directly to data cluster head, and the cluster head on this layer then sends the data to the cluster head of the upper layer, until the data reaches the base station. In the construction of multi-hop path between cluster heads, when the cluster head s_i chooses the next hop path, the forwarding strategy is as follows:

1) If s_i is located on the first floor, then it sends data directly to the base station.

2) If s_i is located in the other layers, and $s_i.G_{ch}$ in the presence of a certain layer so that $d^2(s_i,s_j)+d^2(s_j,BS)$ is the minimum cluster head, and s_i will select s_j as the next hop.

3) When s_i selects s_j as the next hop, if $d(s_i,s_j) > DIS\_MAX$, then s_i searches whether $s_i.C_{nodes}$ is the element of both cross-node clusters or not. If there exist the crossing nodes, they are selected as a bridge for both data forwarding. If $d(s_i,s_j) < DIS\_MAX$, then s_i sends data directly to the s_i.

4) If $s_i.G_{ch}$ there is no next level cluster head, then s_i forwards the data to the same layer from its nearest cluster head, otherwise, the data is forwarded directly to the base station.

5. ALGORITHM ANALYSIS AND SIMULATION ANALYSIS

A. Algorithm analysis

In this section, we will analyze the complexity of CUCRA compared with the classic EEUC algorithm.

On the one hand, we focus on the message complexity of CUCRA. At the beginning of clustering, there are $N \times T$ nodes as candidate cluster head and it need $N \times T$ competition news broadcast COMPETE_HEAD_MSG. However, K nodes are elected as the actual cluster head and need to broadcast K competition to end message FINISH_ELECT_MSG. As a result, the clustering process message overhead that need to be broadcast is $N \times T + K$, while the message overhead of the EEUC algorithm is $(2T+1)K$. Therefore, compared with EEUC algorithm, CUCRA clustering overhead is relatively small. In addition, the EEUC need several iterative messages to complete the selection of cluster head while the CUCRA does not need as much, so that its convergence speed is faster.

On the other hand, unlike the EEUC algorithm, CUCRA considers the residual energy of the nodes in the competition of the radius R_{cmp}. The competition radius of the candidate cluster head is determined by both the candidate cluster head and the distance between the base station and its remaining energy. With the operation of the network, the farther away the node is from the base station, the smaller is the cluster which is built by the cluster heads with less residual energy. The parameter ε_1 and ε_2 determine the influence on competition radius. When the value of ε_1 is 0, the radius of the competition is only determined by the residual energy of candidate cluster head, and when the value of ε_2 is 0, the radius of the competition is determined only by the candidate cluster head and the base station distance. For different network sizes, optimal values of parameters are also different. According to the simulation scale of this paper, a large number of experiments shows that the network has a good performance when the value of ε_1 and ε_2 are 0.6 and 0.4, respectively.

B. Simulation parameter set

This paper simulated and compared CUCRA with LEACH and EEUC algorithms by MATLAB. For simplicity, assuming that we use the ideal MAC protocol and ignore the packet loss, the collision errors happening in wireless links, etc., we use the simulation parameters in [16], which is shown in Table 1.

Table 1. The network simulation parameters

parameter	value
Area acreage	(500, 500)
Node size	100
Base position	(250, 550)
Node initial energy	5 J
E_{elec}	50 nJ/bit
ε_{fs}	10 pJ/bit/ m^2
ε_{mp}	0.0013 pJ/bit/ m^4
$\varepsilon_1, \varepsilon_2, c$	0.5, 0.5, 0.5
Packet size	4000 bits

C. Analysis of simulation results

Based on the above parameters, we make a simulation of LEACH, EEUC and CUCRA. In the simulation, it is assumed that when the residual energy of nodes is less than 0.002J, then the node has failed. When 50% of the nodes in the network are not functioning, it is identified that the network lifetime ends. 15 rounds were randomly selected from the experimental data, and statistics of the total energy consumption of cluster head of each round was produced. Fig.3. shows three kinds of clustering routing algorithm corresponding to the total energy consumption of cluster head in each round.
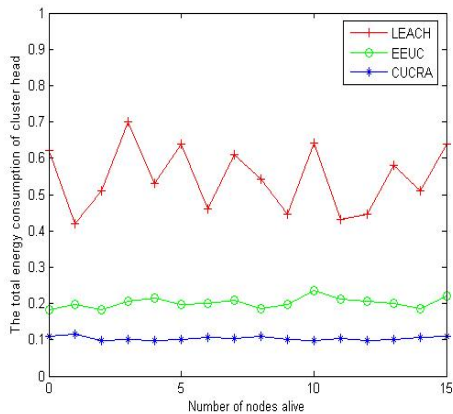

Fig.3. The total energy consumption of cluster head in each round

As it can be seen from Fig.3., in the energy loss of the cluster head, LEACH consumes the most, followed by EEUC, and CUCRA the least because CUCRA and EEUC algorithms are using multi-hop communication, they successfully avoid long-distance communication between cluster head and the base station. At the same time, CUCRA improved the communication between clusters and avoided long-distance communication between the clusters.
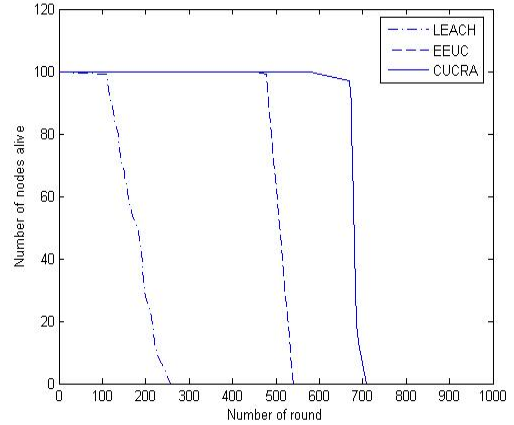

Fig.4. Network lifetime

Fig.4. shows the number changes of surviving nodes in the network as time goes by. It can be seen from Fig.4., whether it is the death time of the first node and last node, CUCRA has an obvious improvement compared with EEUC and LEACH. Because CUCRA takes into account both the remaining energy and the distance between the node and the BS in the clustering radius, cluster size can dynamically change with time and also the distance between the node and the BS. This contributes to the effective balance of energy loss and the extended network lifetime.

6. CONCLUSION

This paper presents an efficient energy-saving non-uniform cross-clustering routing algorithm. The CUCRA combines the advantages of various clustering algorithms and improves them. In the calculation of the cluster radius, two of the dynamic variables of the cluster size were introduced in the parameters. The node residual energy and the distance between node and the BS implement the dynamic variable of the cluster size. At the same time, introducing cluster adjacent node into the communication between clusters for data forwarding, it effectively reduces the energy loss of the cluster head. In addition, the routing address uses hierarchical addressing from lower grades to the high level, it is simple and efficient to implement. The experimental results show that compared with LEACH, EEUC algorithm, CUCRA significantly improve the network lifetime, and balance energy loss of the network nodes.

REFERENCES

[1] Takayama, S., Hiraoka, M., Mori, K., Kariya, K. (2008).Variable data flow management in wireless sensing network for landslide disaster. *Measurement Science Review*, 8 (1), 14-17.

[2] Mitsugu, T. (2009). Application of ZigBee sensor network to data acquisition and monitoring. *Measurement Science Review*, 9 (6), 183-186.

[3] Heinzelman, W., Chandrakasan, A., Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1 (4), 660-670.

[4] Marcelloni, F., Vecchio, M. (2010). Enabling energy-efficient and lossy-aware data compression in wireless sensor networks by multi-objective evolutionary optimization. *Information Sciences*, 180 (10), 1924-1941.

[5] Chamam, A., Pierre, S. (2009). On the planning of wireless sensor networks: Energy-efficient clustering under the joint routing and coverage constraint. *IEEE Transactions on Mobile Computing*, 8 (8), 1077-1086.

[6] Li, C., Ye, M., Chen, G., Wu, J. (2005). An energy-efficient unequal clustering mechanism for wireless sensor networks. In *IEEE International Conference on Mobile Adhoc and Sensor Systems*, 7 November 2005. IEEE, 604-611.

[7] Wen-Hwa Liao, Yucheng Kao, Ru-Ting Wu. (2011). Ant colony optimization based sensor deployment protocol for wireless sensor networks. *Expert Systems with Applications*, 38 (6), 6599-6605.

[8] Li, Z.Y., Shi, H.S. (2009). Exploring load-balancing adaptive clustering algorithm in wireless sensor network. *Journal of Northwestern Polytechnical University*, 27 (6), 822-826.

[9] Yonggang He, Tingrong Xu. (2009). An improved uneven clustering routing algorithm for sensor networks. In *International Symposium on Computer Network and Multimedia Technology (CNMT 2009)*, 18-20 January 2009. IEEE, 1-5.

[10] Peng, D., Zhang, Q.Y., Jia, K.J. (2009). Energy-efficient clustered Routing Protocol for wireless sensor network. *Computer Engineering*, 35 (17), 123-125.

[11] Li, L.Y., Jiang, X.L., Zhong, S., Hu, L. (2009). Energy balancing clustering algorithm for wireless sensor network. In *International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC '09)*, 25-26 April 2009. IEEE, Vol. 1, 61-64.

[12] Xue, X., Qi, R., Qian, F. (2011). Energy-balanced multi-hop uneven clustering routing algorithm for wireless sensor networks. *Journal of East China University of Science and Technology*, 37 (3), 355-358.

[13] Shi, W., Bai, D., Gao, P. (2012). Cluster-head range adaptive adjustment clustering routing protocol for wireless sensor networks. *Chinese Journal of Scientific Instrument,* 33 (8), 1779-1783.

[14] Tong Wang, Chunhui Zhao. (2011). An intelligent processing model on heterogeneous information from the internet of things. *Journal of Computational Information Systems*, 7 (2), 578-584.

[15] Sadeghi, M., Khosravi, F., Atefi, K., Barati, M. (2012). Security analysis of routing protocols in wireless sensor networks. *International Journal of Computer Science Issues*, 9 (1), 465-472.

[16] Heizelman, W., Chandrakasan, A., Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1 (4), 660-670.