

DPSO Based Clustering Algorithm for Location in Wireless Sensor Networks

Baojian Liu, Qing Li*, Xiaoyi Zhang

National Digital Switching System Engineering & Technology R&D Center, Zhengzhou, Henan, China.

* Corresponding author. Tel: 13838086977; email: liqing0206@163.com

Manuscript submitted July 10, 2015; accepted October 16, 2015.

doi: 10.17706/ijcce.2016.5.4.260-268

Abstract: In this paper, a new clustering algorithm for location is presented to improve the location accuracy in the energy constrained wireless sensor networks (WSNs). TDOA (time difference of arrival) location accuracy is affected by the position of the sensor nodes. As WSNs are composed of a huge number of sensor nodes deployed randomly, the optimal nodes for TDOA location are selected by trading off the location accuracy and the networks energy consumption using DPSO (discrete particle swarm optimization). Then a detailed process of the proposed clustering algorithm is given to group the selected node-sets into clusters. Finally, the proposed algorithm is tested under the RSSI (received signal strength indication)/TDOA two rounds cooperative location scheme. Simulation results show that this algorithm improves the location accuracy with low energy consumption.

Key words: Clustering algorithm, DPSO, TDOA, wireless sensor networks.

1. Introduction

With the development of wireless sensor networks (WSNs), the applications have tremendous potential in both commercial and military environments. One typical application scenario is object detection and tracking where sensors are deployed within urban areas to monitor the intruders. There have been many papers on object detection and tracking with sensor networks. But we will present a new clustering algorithm to help with the location accuracy of the sensors.

The TDOA (time difference of arrival) location has higher location accuracy, but the nodes have to transmit amounts of interactive information with each other for TDOA location. It will be a huge challenge for the energy limited WSNs. To reduce energy consumption, a RSSI (received signal strength indication)/TDOA two rounds cooperative location scheme is proposed in [1]. However, TDOA location accuracy is affected not only by the number of nodes but also the position of the nodes. In order to measuring the effect of the position of the nodes on the location accuracy, GDOP (geometry dilution of precision) is introduced. The location accuracy will get higher with the lower GDOP. The author analysis and gives the optimal node deployment in [2]. The lowest value will occur when the range are with respect to N points located at the vertices of a regular N-sided polygon, and will be found at the most favorable location, which is the center of the polygon. Therefore, the nodes with the best position for location will be chosen and grouped into one cluster for further TDOA location.

To reduce the network energy consumption and prolong the network lifetime, grouping sensor nodes into clusters has been widely used such as LEACH, HEED, [3], [4]. The cluster structure is gradually adopted for

target tracking problem. In [5], [6], dynamic clustering approaches are used for target tracking. The proposed clustering algorithms widely vary depending on the node deployment, the pursued network architecture, and the function of the application, etc. Those clustering algorithms do not take consideration of the effect of nodes' position on the accuracy of TDOA location.

With the low network energy consumption, how to choose the nodes with the best position and group them into clusters for further TDOA location is still a question.

In this paper, we propose a novel clustering algorithm, called *Clustering Algorithm for Location (CAL)*. Under RSSI/TDOA two rounds cooperative location scheme, the nodes within the radiation area make full use of the result of RSSI location to re-clustering. Firstly, we define the energy consumption model and location accuracy model. Then the optimal nodes for TDOA location are selected by trading off the location accuracy and the network energy consumption using DPSO (Discrete Particle Swarm Optimization). Finally, based on the CAL, the selected node-sets are grouped into cluster for further TDOA location.

The remainder of the paper is organized as follows. Section 2 describes the problem model and gives an algorithm of choosing the optimal node-sets for TDOA location. In Section 3, we give the detail description of the CAL. Section 4 compared the performance with the HEED. Finally, the paper is concluded in Section 5.

2. Problem Model

For our proposed model, we adopt a few reasonable assumptions as follows:

- 1) The nodes are deployed randomly with uniform distribution in the monitoring area. The nodes are stationary and homogeneous with same initial energy. Boolean model is used for node perception model. The links between nodes are peer-to-peer links.
- 2) There are some anchor nodes which are equipped with GPS to get its geographical location. Others can use the location information of anchor nodes to obtain their absolute position.
- 3) The target node, with constant transmission power, is single, stationary and deployed randomly.
- 4) The Sink node is deployed in the monitoring area without limited energy. All cluster-heads (CHs) communicate with Sink node using multi-hop. In cluster, member nodes (MNs) communicate with cluster-head (CH) using one-hop.

2.1. Energy Consumption Model for the CAL

As in [7] gives an energy consumption model for CAL. We assume that there are n nodes deployed in the area $M \times M$ with sink node at the center of area. N is the optimal cluster sizes. The length of data received from target node is considered for a constant value K_{bit} bit. And the result of location is K_{result_bit} bit. So, the energy expend for a cluster with $N-1$ MNs is

$$E_c = \left[2N \times K_{bit} - (2N - 1) \times K_{result_bit} \right] \times E_{elec} + (N - 1) \times E_{TDE} + K_{bit} \times \varepsilon_{fs} \times d_{toCH}^2 + K_{result_bit} \times \varepsilon_{amp} \times d_{toSink}^4 + (N - 1) \times K_{result_bit} \times \varepsilon_{fs} \times d_{toCH}^2 + E_{TDOA} \quad (1)$$

where E_{TDE} , E_{TDOA} respectively represents: the energy expended on estimating the difference of the arrival time, performing TDOA location. d_{toCH} , d_{toSink} represent the distance between the cluster head and the member nodes, the distance between the cluster head and the sink.

In order to calculate the distance between MNs and CH node within a cluster, we assume that CHs locate at the center of a cluster. Since the nodes are deployed randomly with uniform distribution in the monitoring area. The density of CH is $\rho(x, y) = n / (M^2 N)$, and the average distance between MNs and CH node is

$$E[d_{\text{toCH}}^2] = \iint (x^2 + y^2) \rho(x, y) dx dy = \iint r^2 \rho(r, \theta) r dr d\theta = NM^2 / 2\pi n \quad (2)$$

The same method is used for calculating the average distance of CH node to the sink node. The density of sink node is $\rho_1(x, y) = 1/M^2$, then

$$E[d_{\text{toBS}}^4] = \iint (x^2 + y^2)^2 \rho_1(x, y) dx dy = \int_{\theta=0}^{2\pi} \int_{r=0}^{\frac{M}{\sqrt{\pi}}} r^4 \times (1/M^2) r dr d\theta = M^4 / 2\pi^2 \quad (3)$$

By substituting for (2) and (3) into (1), we get an expression for the energy consumption of a single cluster.

$$E_c = [2N \times K_{\text{bit}} - (2N - 1) \times K_{\text{result_bit}}] \times E_{\text{elec}} + E_{\text{TDOA}} + [(N - 1) \times K_{\text{result_bit}} + K_{\text{bit}}] \times \mathcal{E}_{fs} \times (NM^2 / 2\pi n) + (N - 1) E_{\text{TDE}} + K_{\text{result_bit}} \times \mathcal{E}_{\text{amp}} \times (M^4 / 3\pi^2) \quad (4)$$

From (4), we can find that the energy consumption of a cluster is proportional to the cluster sizes. To prolong the network lifetime, the cluster sizes should be small. But the reducing of nodes will lead to low location accuracy. To balance the energy consumption and location accuracy, the nodes should be grouped into cluster with appropriate number.

2.2. Location Accuracy Model

Reference [8] gives an expression of GDOP in which the position of sensor nodes is converted to the angles between the target and sensor nodes. The formula is further deduced and the main factors affecting GDOP value are given.

Assumed that there is a set of N ($N > 2$) sensor nodes placed at locations $p_i = [x_i, y_i]^T$, $i = 1, 2, 3 \dots N$, and a single stationary target located at $p_T = [x_T, y_T]^T$.

$$GDOP = \sqrt{N / \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \sin^2(\alpha_i - \alpha_j) \right)}, \text{ s.t., } N \geq 3. \quad (5)$$

where $[\alpha_1, \alpha_2 \dots \alpha_N]$ are the angles of the ray from the target to the sensor nodes relative to the positive x-axis.

Therefore, we present an idea based on RSSI/TDOA two rounds cooperative location scheme that the nodes within radiation area to re-clustering as the error of initial RSSI location under the allowable range. To reduce the energy consumption and improve the target location accuracy, the nodes which are best for location are selected and grouped into cluster for further TDOA location.

2.3. Optimal Node Selection

To ensure the low energy consumption, how to choose the optimal node-sets is actually a combinatorial optimization problem. There are many optimization algorithms in solving combinatorial optimization problem such as genetic algorithm, artificial neural network, PSO [9]. For comprehensive consideration the complexity and effectiveness of the algorithm, the DPSO is selected as the optimization tool.

The DPSO operates on discrete binary variables, which is evolved from previous versions of the PSO

operated in continuous space. In DPSO, the position of each particle is converted to a binary sequence. For each row of the matrix X represents a particle, the dimension of the particle is equal to the number of nodes within the radiation area. In the matrix X , 1 represents the node is selected, 0 is not selected. To meet the requirement of TDOA location, when the number of nodes selected in the initialization or updating process is less than 3, the corresponding particle should be reset.

In order to balance energy consumption and improve location accuracy, the fitness function is defined as follow.

$$f_1(N) = \sqrt{N / \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \sin^2(\alpha_i - \alpha_j) \right)} \quad (6)$$

$$f_2(N) = \left[2N \cdot K_{bit} - (2N-1) \cdot K_{result_bit} \right] \cdot E_{elec} + \left[(N-1) \cdot K_{result_bit} + K_{bit} \right] \cdot \varepsilon_{fs} \cdot (NM^2 / 2\pi n) + (N-1) E_{TDE} \quad (7)$$

$$F_{fitness}(N) = \varepsilon \cdot (f_1(N) / \|f_1(N)\|) + (1-\varepsilon) \cdot (f_2(N) / \|f_2(N)\|), \text{ s.t. } N > 3, 0 < \varepsilon < 1 \quad (8)$$

where (6) reflects the effect of the number and the position of sensor nodes on TDOA location accuracy. Equation (7) reflects the effect of the cluster sizes on energy consumption. The fitness function is the result of normalized energy consumption and location accuracy (as shown in (8)). In this paper, the weighting factor ε is equal to 0.5.

3. The Clustering Algorithm for Location

Based on RSSI/TDOA two rounds cooperative location scheme, a twice clustering idea is proposed for completing once high-precision location. The progress of cooperative location is divided into rounds. In order to ensure there is sufficient time for location, a round is defined as completion once location (as shown in Fig. 1). According to the TDOA location accuracy is affected by the distribution of sensor nodes relative to the target, the re-clustering is performed by the nodes within the radiation area before the TDOA location.

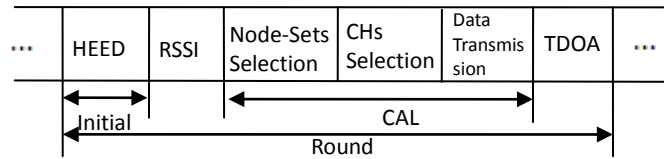


Fig. 1. The progress of RSSI/TDOA location.

The HEED algorithm is performed to initialize the network. The uniform distribution of CHs ensures the monitoring of the target. The process of CAL is performed after getting the coarse estimation of target location by using RSSI location scheme. The detail progress of CAL is divided into three phases, the optimal node-sets selection, cluster head selection, and data transmission.

3.1. The Optimal Node-Sets Selection

After completing the RSSI coarse location, the Sink node transmits the result of the RSSI location to each CH. The CH nodes determine whether it is in the radiation area as they received the message transmitted by Sink node. If the CH node is in the radiation area, CH nodes forward the result of the RSSI location to MNs and

dismiss the cluster structure. Otherwise, only the result information is transmitted to MNs. If there are parts of MNs within a cluster located in the radiation area, the covered MNs will send a request message for disassociation after receiving the result of the RSSI location. Therefore, if there are parts of MNs within a cluster located out of the radiation area, the uncovered MNs perform the HEED algorithm to re-clustering after receiving the dismiss message sent by CH node.

To reduce energy consumption, we introduce a node called *temporary central node*. The roles of this node are to collect the information of nodes which is located at the radiation area and to choose the optimal node-sets by implementing the optimal node selection algorithm as above. The temporary central node is selected based on the principle of maximum residual energy.

The nodes covered by the target transmit a message containing the nodes' ID, position, and residual energy to the temporary central node. Then the temporary central node repeatedly performs the optimal node-sets selection algorithm as described in Section 2. In each iteration, the selected node-set will be deleted from the node-set which is covered by the target. It isn't terminated until the number of remaining nodes is less than 3. The selected node-sets will be marked. Finally, the results of selected node-sets are transmitted to the covered nodes by the temporary central node.

3.2. The Cluster Head Selection

In this phase, CHs are selected from the optimal node-sets to provide a foundation for TDOA location and prolong network lifetime. CH selection is primarily based on the residual energy of each node and the distance to the Sink node.

When the covered nodes received the message transmitted by the temporary central node, the nodes within the same node-set set the back-off timer based on the residual energy and the distance to the Sink. Once the back-off timer expires, the node broadcasts an advertisement message to select itself as the CH in this round. Other nodes within the same node-set stop the timer and send a message to join the cluster. This makes each node-set to be a cluster which is best for TDOA location with less energy consumption.

3.3. Data Transmission

After the phase of cluster head selection, the nodes start to collect the signal of the target and forward the data to the sink. In this phase, it is divided into two parts: intra-cluster and inter-cluster routing. As is mentioned in Section 2, we assumed single-hop communication among cluster head and their registered cluster members. The member nodes collect the data of the target and transmit it to the cluster head node. Then the cluster head aggregate the received data and forward the data to the sink node for TDOA location. In this process, cluster head overlay routes are used to communicate among clusters. In this case, a routing protocol is employed for data forwarding among cluster heads, such as AODV and Dijkstra.

4. Simulation and Analysis

The simulation environment is composed of 100 sensor nodes which are distributed in a region of 100m×100m randomly. The target node is single, stationary and deployed randomly. The sink node located at the center of the area. It is assumed that the error ratio of the initial RSSI location is limited in [0, 0.3]. The algorithm is performed in Matlab7.0. Table 1 outlines the key simulation parameters.

Table 1. Simulation Parameters

Parameter	Units	Values	Parameter	Units	Values
Initial energy	J	0.1	E_{RSSI}	nJ	5
Perception radius	m	20	E_{TDE}	nJ	50

Radius of target	m	20	E_{GDOP}	nJ	5
E_{elec}	nJ/bit	50	E_{TDOA}	nJ	5
\mathcal{E}_{mp}	pJ/bit/m ⁴	0.0013	\mathcal{W}		0.5
\mathcal{E}_{fs}	pJ/bit/m ²	10			

In order to verify the effectiveness of the proposed algorithm, we compared the algorithm with HEED algorithm. According to the above simulation conditions, the HEED algorithm needn't re-clustering after getting the result of initial RSSI location. Instead it uses the coarse estimation of the target position to choose a part of clusters which have lower GDOP for further TDOA location.

Due to the random of the algorithm, the result of clustering is different in the size and structure of the cluster each time. The lower GDOP will be got with the larger of cluster size. Therefore, it can't simply compare the value of GDOP to decide which algorithm is better without the same size and structure of the cluster. So the error ratio of GDOP is defined as follow to measure the location accuracy.

$$GDOP_{error} = (GDOP_{real} - GDOP_{ideal}) / GDOP_{ideal}, GDOP_{ideal} = 2/\sqrt{N} \quad (9)$$

where $GDOP_{real}$ is the actual value, $GDOP_{ideal}$ is the Cramer-Rao Bound with N sensor nodes.

4.1. Location Accuracy

We performed the algorithm for 50 rounds. In each round, the nodes and the target are deployed randomly with uniform distribution in the monitoring area. The CAL and HEED are performed 500 times in each round. Then the corresponding cluster sizes and error ratio of GDOP are calculated each time. Finally, the average cluster sizes and error ratio of GDOP are calculated as the final value of this round. The simulation results are shown as follows.

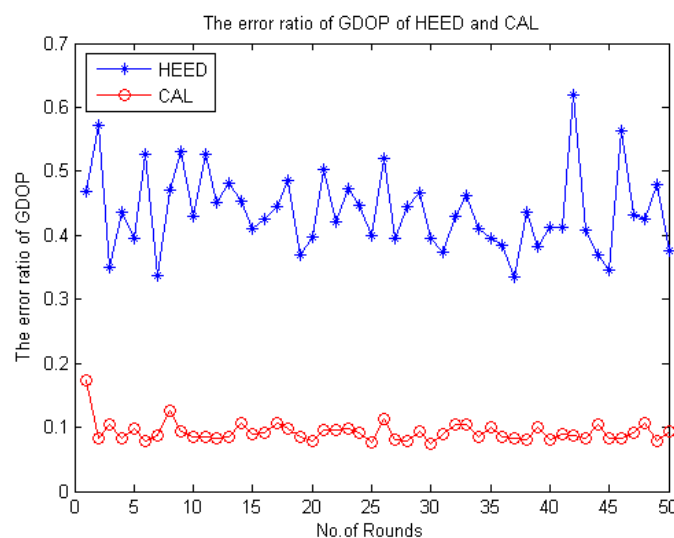


Fig. 2. The error ratio of GDOP.

Fig. 2 shows that the CAL is better than HEED on the location accuracy of TDOA. This is because the CAL chooses the optimal node-sets and grouping them into cluster before performing the TDOA location. Since the effect of the size and structure of a cluster on the accuracy of TDOA location isn't considered by HEED, this leads to a large range of the error ratio of GDOP.

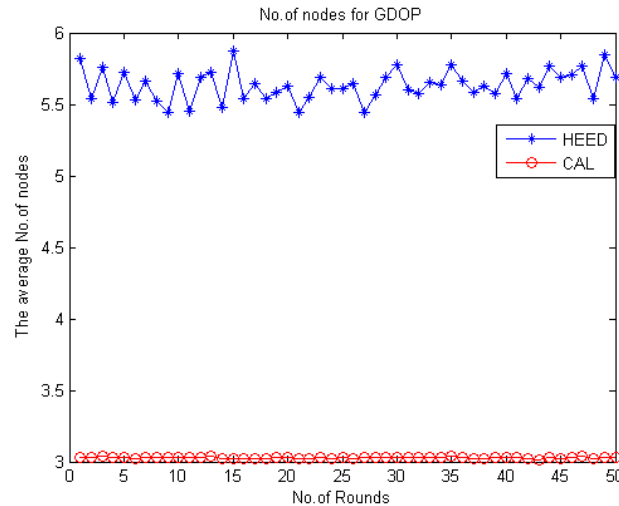


Fig. 3. The average number of nodes.

Fig. 3 shows that the average cluster sizes of HEED changes in a wide range. In some cases, there may be no cluster covered by the target. So the TDOA location cannot be guaranteed in HEED. But, in CAL, due to the optimal node-sets selection, the cluster sizes are substantially the same in each round. Hence, the TDOA location can be guaranteed unless the nodes covered by the target are less than 3.

4.2. Energy Consumption

The simulation parameters are set as shown in Table 1. Due to the nodes outside the radiation area do not take part in the TDOA location, the energy consumption is alike in CAL and HEED. The two algorithms are both performed for 1000 rounds. We observe the number of alive nodes in the radiation area.

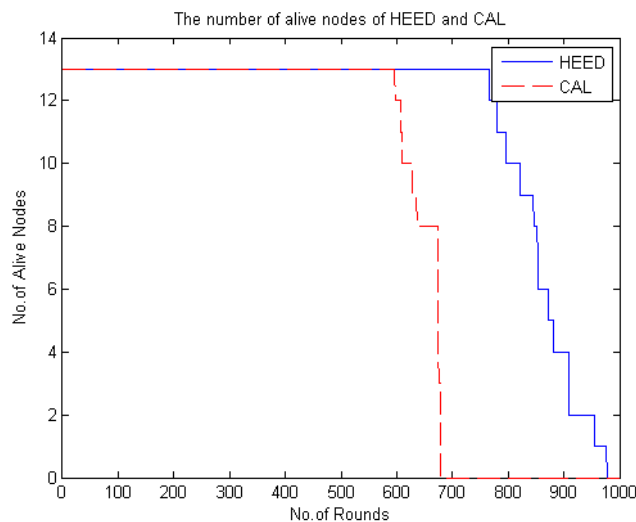


Fig. 4. The number of alive nodes.

As shown in Fig. 4 the number of alive nodes within the radiation area is compared between the CAL and HEED. The death of the first node is earlier in CAL than in HEED. This is because the process of the CAL needs to consume a lot of energy while the HEED need not. But it can't prove that the HEED performs better than the CAL on the energy consumption under the same accuracy of TDOA location. The changing cluster size in each round makes the accuracy of TDOA location can not be guaranteed in HEED. In CAL, the energy consumption keeps roughly the same due to the cluster size remains unchanged. Hence, the gap between the

death of the first node and the last node is less in CAL than in HEED.

4.3. The Effect of Initial RSSI Location

According to RSSI/TDOA two rounds cooperative location scheme, we need to know the coarse position of the target to choose the optimal node-sets for TDOA location. Thus the accuracy of initial RSSI location affects the optimal node-sets selection.

The error ratio of RSSI location is given from 0 to 1 where the maximum error is 10 m. In order to simulation the influence of the error of initial RSSI location on the TDOA location, we perform 500 times Monte Carlo simulation in each given error, then calculate the average error ratio of GDOP. The result is shown as follow.

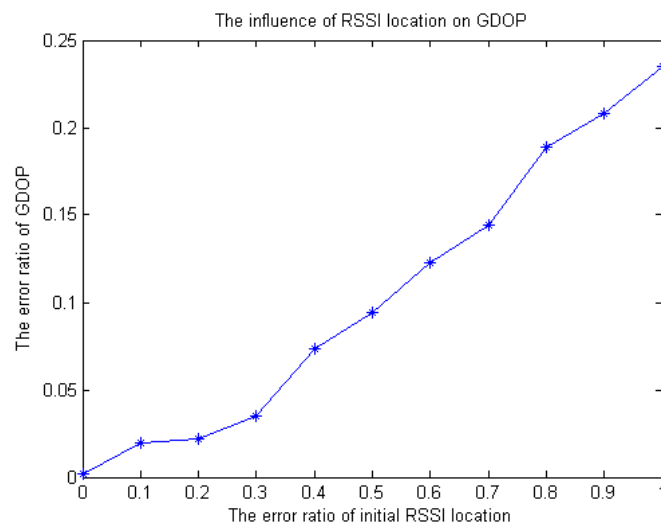


Fig. 5. The influence of RSSI location on GDOP.

Fig. 5 shows that the accuracy of TDOA location decreasing with the increase of the error ratio of initial RSSI location. To ensure the effectiveness of CAL, the error ratio of initial RSSI location should be limited to the range of [0, 0.3].

5. Conclusions

In this paper, we have presented a clustering algorithm for location based on RSSI/TDOA two rounds cooperative location scheme. The clustering algorithm gives a new way to improve the location accuracy by grouping the optimal node-sets into clusters for further TDOA location. It has been shown that the CAL has a better performance on location with a little energy consumption.

This work focus on the single and stationary target. In near future multi-targets or moving target will also be considered. The appropriate clustering algorithm will be presented in different situations as future work.

Acknowledgment

The authors wish to thank Dr. Q. Li and Dr. X. Y. Zhang for valuable comments and a fruitful discussion.

References

- [1] Li, X., Zhang, X. Y., & Zeng, Y. (2012). An efficient passive location method based on clustering structure wireless sensor network. *Signal Processing*, 28(4), 587-594.
- [2] Levanon, N. (2000). Lowest GDOP in 2-D scenarios. *IEE Proceedings-Radar, Sonar and Navigation*, 147(3), 149-155.

- [3] Bara'a, A. A., & Khalil, E. A. (2012). A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Applied Soft Computing*, 12(7), 1950-1957.
- [4] Selvi, G. V., & Manoharan, R. (January 2013). Unequal clustering algorithm for WSN to Prolong the network lifetime (UCAPN). *Proceedings of 4th International Conference on Intelligent Systems Modelling & Simulation* (pp. 456-461).
- [5] Kong, J. I., Kim, J. W., & Eom, D. S. (2014). Energy-aware distributed clustering algorithm for improving network performance in WSNs. *International Journal of Distributed Sensor Networks*.
- [6] Wang, Z., Lou, W., Wang, Z., Ma, J., & Chen, H. (2013). A hybrid cluster-based target tracking protocol for wireless sensor networks. *International Journal of Distributed Sensor Networks*.
- [7] Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (January 2000). Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd annual Hawaii international conference on System Sciences* (p. 10).
- [8] Quan, Q. (June 2012). Low bounds of the GDOP in absolute-range based 2-D wireless location systems. *Proceedings of 8th International Conference on Information Science and Digital Content Technology: Vol. 1* (pp. 135-138). IEEE.
- [9] Kulkarni, R. V., & Venayagamoorthy, G. K. (2011). Particle swarm optimization in wireless-sensor networks: A brief survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 41(2), 262-267.



Baojian Liu was born in Henan Province, China, in 1989. He received the B.S. degree from Henan University of Urban Construction (HUUC), Pingdingshan, in 2013. He is currently pursuing the M.S. degree in communication engineering from National Digital Switching System Engineering & Technology R&D Center, Zhengzhou. His research interests include wireless sensor networks and clustering algorithm.



Qing Li was born in Hebei Province, China, in 1975. She received the M.S. degree and the Ph.D degree in time measure from National Digital Switching System Engineering & Technology R&D Center, Zhengzhou, in 2003 and 2008, respectively, where she is an associate professor now. Her research interests mainly focus on ad hoc and wireless sensor networks, protocol analysis, visible light communication etc.



Xiaoyi Zhang was born in Henan Province, China, in 1966. He received the M.S. degree in electrical and telecommunications engineering from Xidian University in 1997. Now he is a professor in National Digital Switching System Engineering & Technology R&D Center, Zhengzhou. His research interests mainly focus on wireless communication system, software radio, ad hoc and wireless sensor networks, visible light communication etc.