Study of Wireless Sensor Network Location Algorithm and Data Collection Protocol

Weimin Han[1], Shijun Li[2,4], Weidong Li[3]

[1] Hunan Vocational Institute of Technology, Xiangtan, China;
[2] Hunan Institute of Engineering; Hunan Provincial Key Laboratory of Wind Generator and Its Control Xiangtan, China
[3] Powerchina Zhongnan Engineering Co.Ltd, Changsha, China
[4] Hunan province Cooperative Innovation Center for Wind Power Equipment and Energy Conversion, Xiangtan, China

Abstract—**It is deduced that the network life time is inversely proportional to the zone area of the collection data that each sink is responsible for. That is to say, when the zone area in the charge of a sink is smaller, the average load of nodes distributed in this area and the energy consumption are smaller, while the network life time is longer. MSGP protocol is proposed, with several round cluster head candidate zones, fixed member zones and public zones being delimited to make approximately uniform clustering. The cluster head is responsible for collecting a meaningful event in the cluster, making data fusion and then forwarding it to sink. Multi-sinks move according to the zone division rules and the rule of free movement. Specific analysis is made from three aspects of data direction, single round data collection of sink, and network life time.**

Index Terms—**wireless sensor network, data collection protocol, location algorithm, multi-sinks.**

I. INTRODUCTION

The application and study of static wireless sensor network are based on the assumption that the nodes do not move. With the expansion of the application, many scholars propose the concept of mobile nodes. In such applications, some nodes can be moved. For example, in the application of ocean monitoring, the sensor node is installed on a buoy and it is inevitable for the buoy to move. Additionally, the sensor node installed on the automobile can also move. It is by relying on the mobile nodes that the traffic center obtains, analyzes and processes the traffic flow in this region so as to make new and favorable traffic scheduling [1].

The application of mobile nodes brings new opportunities to the wireless sensor network, which can improve the locating accuracy of the nodes, and prolong the lifetime of the network. However, it also leads to some challenge. The frequent node movement results in failure of the established communication link and solution of the time slot distribution. Some classical routing methods can also fail completely, so this raises higher requirements for the design of communication protocol.

II. ANALYSIS OF THE PROBLEM

In the wireless sensor network, the load of nodes surrounding the sink is the largest in the whole network whose life cycle depends on the life of these nodes. The average node load formula is deduced targeting the round network. The lifetime of a multi-sinks network is further analyzed with this formula. When there is only one sink in the network [2], the average node load Li is as follows, with the sink being at the center of the circle:

$$L_i = \begin{cases} \dfrac{\beta/2\left(R^2-d^2\right)\lambda e}{\pi/2\cdot r^2} + \lambda e & d \geq r \\ R^2\lambda e/r^2 & d < r \end{cases} \quad (1)$$

Among which representing the radius of the round network, r is the communication radius of the node, λ is the probability in which the node transmits the data within the unit time, e is the energy that the node consumes to receive and transmit the unit data, J is the distance from the node i to the circle center of the network, and β is constant, $\beta = 2\arcsin(r/d)$. When the node is close to sink, the load is large, as well is the energy consumption. As shown in formula (1), the maximum load of the node can be represented as follows:

$$\left(L_{\max}^i\right)_{1-link} = \frac{R^2\lambda e}{r^2} \quad (2)$$

By multiplying the numerator and the denominator with π the formula above becomes:

$$\left(L_{\max}^i\right)_{1-link} = \frac{\pi R^2\lambda e}{\pi r^2} \quad (3)$$

πR^2 and π are the area of the whole network and the area of the node communication scope respectively. When they are represented by S1-link and S_r, we obtain:

$$\left(L_{\max}^i\right)_{1-link} = \frac{S_{1-link}\lambda e}{S_r} \quad (4)$$

Assuming the initial energy of all nodes in the network is the same, which is represented by E_{ini}, then the network lifetime can be represented as:

$$T_{multi-\sin k} = \alpha \frac{E_{ini}}{\left(L_{\max}^i\right)_{multi-\sin k}} \quad (5)$$

$$T_{1-\sin k} = \alpha \frac{E_{ini}}{L_{\max}^i} \quad (6)$$

Among which α represents the coefficient, T multi-sink and T1-sink are the lifetime of the network owning multiple sinks and single sink respectively. The ratio of this tow is:

$$\frac{T_{multi-\sin k}}{T_{1-\sin k}} = \frac{\alpha \dfrac{E_{ini}}{\left(L_{\max}^i\right)_{multi-\sin k}}}{\alpha \dfrac{E_{ini}}{L_{\max}^i}} = \frac{\dfrac{S_{1-\sin k}\lambda e}{S_r}}{\dfrac{S_{multi-\sin k}\lambda e}{S_r}} = \frac{S_{1-\sin k}}{S_{multi-\sin k}} \quad (7)$$

We take the lifetime of a single sink network as the reference, set T1-sink=1, then the formula above can be simplified as follows:

$$T_{multi-\sin k} = \frac{S_{1-link}}{S_{multi-\sin k}} \qquad (8)$$

As known from formula (9), when there are multi-sinks, the network lifetime is inversely proportionate to the zone area that each sink is responsible for. That is to say, when the zone area is smaller of the data collection that sink is responsible for, the average load of the nodes distributed in this zone is smaller and the node energy consumption is smaller but the network lifetime is longer. Aiming at a multi-sinks scenario, we propose two approaches for comparison [3]. One is to divide the data collection zone that each sink is responsible for, another is that sink moves freely. It is considered in this section that multiple mobile sinks and n sensor nodes collect the data in the network (L×L). The sensor nodes are supplied by the battery, their communication radius is the same and their own coordinates are known. Multi-sinks do not have a limitation of energy.

III. MULTI-SINKS DATA COLLECTION PROTOCOL

Clustering treatment for MSGP agreement is completed due to the advantages of clustering. Several rounds of cluster head candidate zones, fixed member zones and public zones are delimited to make approximately uniform clustering[4]. The cluster head is responsible for collecting the relevant event in the cluster, making data fusion and then forwarding it to sink [5]. Multi-sinks move according to the zone division rules and free movement rule. Additionally, these two approaches are compared.

A. Round clustering law

$$\begin{cases} N_x^i = \left\lceil \dfrac{x_i}{l} \right\rceil \\[2mm] N_y^i = \left\lceil \dfrac{y_i}{l} \right\rceil \\[2mm] \left(N_x^i - \dfrac{1}{2}\right)l - r_0 \le x_i \le \left(N_x^i - \dfrac{1}{2}\right)l + r_o \\[2mm] \left(N_y^i - \dfrac{1}{2}\right)l - r_0 \le y_i \le \left(N_y^i - \dfrac{1}{2}\right)l + r_0 \\[2mm] \left(x_i - \left(N_x^i - \dfrac{1}{2}\right)l\right)^2 + \left(y_i - \left(N_y^i - \dfrac{1}{2}\right)l\right)^2 \le r_0^2 \end{cases} \qquad (9)$$

(1) Round cluster head candidate zone

Several round cluster head candidate zones are set and the nodes deployed in the zone can compete for cluster heads. The advantages of uniform head distribution have already been introduced above. As shown in Figure 1, the inside of the small cycle is the round cluster head candidate zone. The node decides whether it belongs to this zone depending on formula (10).

In the formula, x_i and y_i are the coordinates of node *i*. [xi/l] is rounded up to the integer; i.e., the minimum integer larger than xi/l, r_0 is the radius of cluster head candidate zone (r_0=l/2), $\left(\left(N_x^i - 1/2\right)l, \left(N_y^i - 1/2\right)l\right)$ is the coordinate of central point of the candidate zone [6].

(2) Fixed member zone

One circular fixed member area is set outside each round cluster head candidate zone. The nodes deployed at this zone can only enter the cluster formed by the cluster
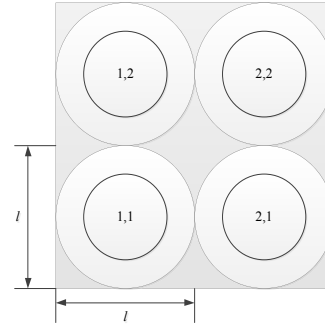

Figure 1.   Partition model

head candidate zone in the middle. r_i is the circular ring width (r_i=l/2). The circular ring area between the large and small circles in Figure1 is the fixed member zone [7].

(3) Public zone

Outside the fixed member zone is the public area. The nodes deployed in this zone join the cluster formed by the cluster heads closest to them, as shown in the shaded area in Figure 1. $f_d = (r_0 - r_i)/r_0$ is the distance from the node to the center of round cluster head candidate zone. Figure 2 shows the cluster heads produced with the traditional clustering approach. Their position distribution is uniform and cluster heads in this part of the zone is too low. Figure 3 shows the cluster heads produced with the round clustering approach whose position distribution is relatively uniform.
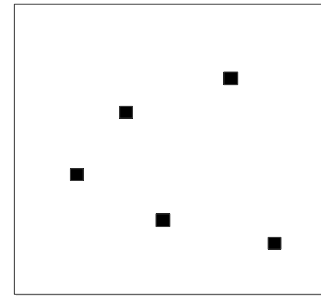

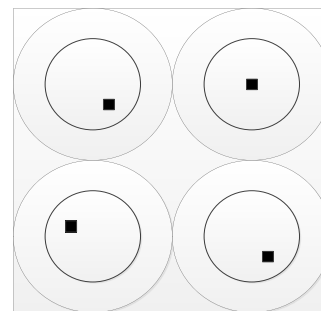Figure 2.   Uneven distribution of head

Figure 3.   Nearly uniform distribution of head

(a) Non-uniform distribution of cluster heads (b) Approximately uniform distribution of cluster heads [8-9].

B. MuIti-sinks movement rules

With the goal of multi-sinks, two kinds of movement rules are proposed: regional divisions rule and free movement rule. S is used to represent the multi-sinks set $S = \{s_1, s_2, L, s_k\}$, among which k is constant $(k \ge 2)$. Assuming a sink moves once each other T rounds [10].

Figure 4 shows the data flow chart of sink stop points. The circle represents the round cluster head candidate zone which is abstracted to one cluster for convenience of analysis. The black triangular is sink and the dotted arrow is the data flow direction. In Figure 4, the sink stops at the non-boundary area and there are eight clusters in total in the surrounding which are A, B, C, D, E, F, G and H. In Figure 5, there are five clusters in total which are A, B, C, D and E. It is assumed that each cluster sends n data packets to sink each round. The total data transmitted by peripheral cluster heads which will be received by the eight adjacent clusters of the sink in Figure 4 is 16n while that received by the five adjacent clusters of the sink in Figure 5 is 19n. Therefore, the average data received by the adjacent clusters of sink in Figure 4 is

$$\overline{data_a} = \frac{16n}{8} = 2n \qquad (10)$$

The average data received by the adjacent clusters of sink in Figure 5 is

$$\overline{data_a} = \frac{19n}{5} \qquad (11)$$

Obviously, $\overline{data_b} > \overline{data_a}$; i.e., the load of adjacent clusters of sink is relatively large, and so is the energy consumption in Figure 4. In order to reduce the load of neighboring clusters and prolong the network lifetime so that the sink can stop in the middle of responsible data collection zone as much as possible, the target selection function to select the next stop point of sink is proposed.
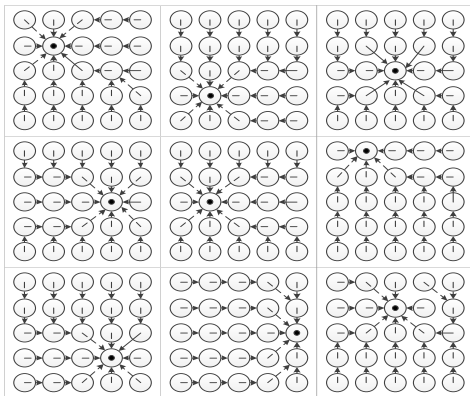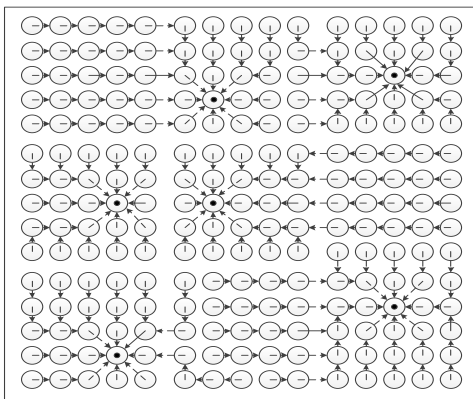

Figure 4.   Data Direction of MSGPZDK

Figure 5.   Data Direction of MSGPFMR

C. Target selection function

$$V = \theta_1 V_1 + \theta_2 V_2 \qquad (12)$$

Among which $V_1 = (d_{center} - d_i)/d_{center}$, $V_2 = 1/(p_i + 1)$ represents the distance from apex A to the center 0 of the data collection responsible zone, di represents the distance from the center of each round cluster head candidate area in the responsible zone to 0; pi represents the number that sink stops in the round cluster head candidate area, $\theta 1$ and $\theta 2$ are adjustment parameters. The target selection function considers the next stoping point of sink from two aspects which are the distance from the central point of responsible zone to the number of sink stops. The cluster closer to the central point and with a smaller number of stops will be taken as the stop point.

IV. PERFORMANCE ANALYSIS AND SIMULATION TEST

Performance analysis is conducted for two MSGP protocols in this section. The total amount of data produced by the network is D_{total}. In $MSGP_{ZDR}$, the amount of data each sink collects is D_{total}/k, while in $MSGP_{FMR}$, the maximum amount of data that sink can collect is D_{total} and the minimum is 0. We choose the mean value of $D_{total}/2$. As can be seen in Figure 5, the adjacent clusters of sink can be eight at most and three at least, and we choose the mean value of 5.5. The data that the adjacent cluster has to transfer in $MSGP_{ZDR}$ and $MSGP_{FMR}$ are $D_{total}/5.5k$ respectively. While $k \geq 2$, so $D_{total}/11 \geq D_{total}/5.5k$; i.e., the more the data transferred by the neighboring clusters of sink in $MSGP_{FMR}$, the more the energy consumption (as other parameters shown in Table 1).

A. Data flow

Figures 6 and 7 show the data flow of MSGPZDR and MSGPFMR protocol at a certain time. Since the responsibility zone is divided, the sink distribution in MSGPZDR is relatively uniform and the load amount in each responsibility zone is approximately the same. In this way, the energy consumption of the neighboring clusters surrounding sink tends to be the same. Additionally, the position information of sink is only updated in the responsibility zone which avoids energy consumption for forwarding due to the overlapping of multiple sink position information. In MSGPFMR, all the sinks move freely, which will lead to the non-uniform distribution of sinks and a portion of sink neighboring load is too large, increasing the energy consumption [11-13].
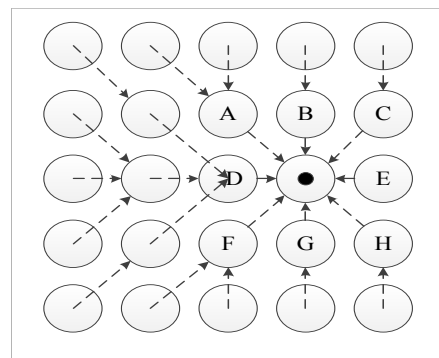

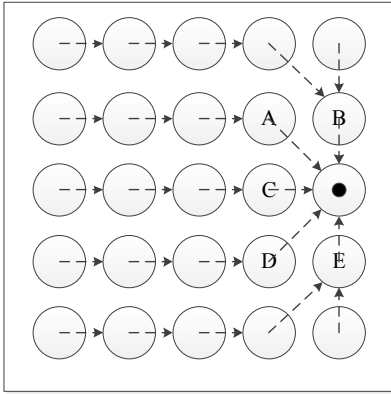Figure 6.   Non-edge region for stopping position

Figure 7. Edge region for stopping position

The application and study of static wireless sensor network are based on the assumption that the nodes do not move. With the expansion of the application, many scholars propose the concept of mobile nodes. In such an application, some nodes can be moved. For example, in the application of ocean monitoring, the sensor node is installed on the buoy and it is inevitable that the buoy will move. Additionally, the sensor node installed on the automobile can also move. It is by relying on the mobile nodes that the traffic center obtains, analyzes and processes the traffic flow in this region so as to make new and favorable traffic scheduling.

In the process of building the routing, a child radio base station began to constantly transfer an information package to the surrounding node, establishing a communication radius of the expansion of the tree structure[14]. The routing tree structure, base station is considered as the root, the data is converting from the source node to the leaves of the tree, and the data flow from the leaves to the root node. After completing the establishment of the tree structure, in the process of the data transfer from the source node to others, the source node can adjust the routing link on a small scale, such when the original link in source node becomes dormant[15-16]. Because the wireless sensor network (WSN) is in broadcast mode, all the work nodes can receive information within the scope of the radio, so the new link can be established. Where the data flow is concentrated, in the original tree structure, the node can rely on the tree structure to establish multiple pathways and finally get to the base station, without the need to refactor all routing processes [17].

The application of mobile nodes brings new opportunities to the wireless sensor network, which can improve the locating accuracy of the nodes, and prolong the lifetime of the network [18-20]. However, it also leads to some challenges. The frequent node movement results in failure of the established communication link and time slot distribution solution. Some classical routing methods can also fail completely, so this raises higher requirements for the design of communication protocol.

B. Data collection by sink each round

Tables 2 and 3 show the data collection by sink each round of these two protocols. In Table 2, the data collection by sink each round is basically the same, which suggests that the load of sink neighboring clusters is basically the same and the energy consumption is uniform, and this is attributed to the division of sink data collection respon-

sibility zones in $MSGP_{ZDR}$. In Table 3, the data collection amount is shown of sink 1 and sink 2 as 432.61kb and 63.24kb. Obviously, the single round data collection of each sink in $MSGP_{FMR}$ is largely different. The load of neighboring clusters of the sink with a large data collection amount in single round is too large, so the energy consumption is too high, while the load of the adjacent clusters of the sink with a small data collection amount in a single round is small, so the energy consumption is also small. The unbalance in energy consumption will lead to the "energy hole" and shortens the network lifetime. This is because each sink of $MSGP_{FMR}$ can move freely, which will inevitably make the sink density in a certain part higher and lower in another part. The load is large where the sink density is low and the load is small when the sink density is high.

The application and study of static wireless sensor networks are based on the assumption that the nodes do not move. With the expansion of the application, many scholars propose the concept of mobile nodes. In such applications, some nodes can be moved. For example, in the application of ocean monitoring, the sensor node is installed on a buoy and it is inevitable for the buoy to move. Additionally, the sensor node installed on the automobile can also move. It is by relying on the mobile nodes that the traffic center obtains, analyzes and processes the traffic flow in this region so as to make new and favorable traffic scheduling.

C. Network life time

The network lifetime of these two protocols is shown in Table 4. In $MSGP_{ZDR}$, the best situation and the worst situation of the network lifetime is 5631.52 rounds and 3451.26 rounds respectively while the network lifetime corresponding to $MSGP_{FMR}$ is 1549.74 rounds and 679.32 rounds. The sink load of $MSGP_{ZDR}$ is relatively uniform, so that the node energy consumption is relatively uniform and the network lifetime is longer; while the sink load in some portions of $MSGP_{FMR}$ is too large, so that the energy consumption of their adjacent clusters is too large and they fail in advance, which shortens the network lifetime.

TABLE I.
PARAMETER SET

Parameter	Value
Side length of the network	2700mm
l_{sink}	900mm
L	180mm
R_0	90mm
R_1	90mm
Communication radius	30mm
The node density	0.003/m2
Elect	50nj/bit
ε_{fs}	10pj/bit/m2
ε_{mp}	0.0012pj/bit/m2
Eda	5nj/bit/m2
θ_1：θ_2	1:1
K	8
Data size	2000bit
Initial energy	2J

TABLE II.
SINK DATA COLLECTION FOR ONE TERM(MSGPZDR)

	Sink1	Sink2	Sink3	Sink4	Sink5	Sink6	Sink7	Sink8	Sink9
data size	298.36	280.36	299.14	301.25	292.34	298.23	290.45	301.14	281.11

TABLE III.
SINK DATA COLLECTION FOR ONE TERM(MSGPFMR)

	Sink1	Sink2	Sink3	Sink4	Sink5	Sink6	Sink7	Sink8	Sink9
data size	432.21	63.21	596.32	104.25	190.36	54.12	341.25	496.58	369.48

TABLE IV.
THE COMPARISON FOR NETWORK SURVIVAL TIME

	MSGPZDR		MSGPFMR	
	The best condition	*The worst condition*	*The best condition*	*The worst condition*
Network survival time	5621.36	3425.36	1548.45	678.12

V. CONCLUSIONS

The largest communication model of a wireless sensor network lifetime is based on multiple source links and multiple base stations. The data that was produced from source node can be transmitted to different links of base station nodes. Source node product data, and intermediate node provides data forwarding business. Data translates from the edge of network nodes to base station node collection, and data source node perception though a jump or jump forward to base station node. Homologous node is possible to be routed to different base station nodes through different data links. Homologous node data has a strong correlation, and different source nodes can cognate data in the process of data fusion. Considering the source data fusion processing that the base station node needs from multiple base station nodes to a tree structure, the root node is connected to the base station through a virtual link. One of the base station nodes is as a root node, and the root node deletes redundant compressed data before sending data to the terminal receiver.

One wireless sensor network data collection protocol based on multiple mobile sinks is presented in this paper. With the expansion of the application zone, the energy consumed by the node farther away from sink to transmit the relevant event will increase and the real time performance is reduced. This study is focused on the data collection protocol with multiple mobile sinks. In MSGP protocol, several round cluster head candidate zones, fixed member zones and public zones are delimited to make approximately uniform clustering. The sink moves according to the rule of zone division and the rule of free movement ($MSGP_{ZDR}$ and $MSGP_{FMR}$). It is analyzed theoretically that multiple sins can balance the network load and save the energy consumption of node. The simulation results show $MSGP_{ZDR}$ is better than $MSGP_{FMR}$ in terms of data collection amount by sink in each round and network lifetime.

REFERENCES

[1] M.A. Adnan, M.A. Razzaque, "A novel cuckoo search based clustering algorithm for wireless sensor networks," *Lecture Notes in Electrical Engineering*, vol. 3, pp. 621-634.August 2016. http://dx.doi.org/10.1007/978-3-319-24584-3_53

[2] R. Anane, R. Bouallegue, K. Raoof, "Medium access control (MAC) protocols for wireless sensor network: An energy aware survey, "*Lecture Notes in Electrical Engineering*, vol. 3, pp. 561-569, June 2016. http://dx.doi.org/10.1007/978-3-319-30301-7_60

[3] C. Emmanouilidis, P. Pistofidis, "Design requirements for wireless sensor-based novelty detection in machinery condition monitoring. Engineering Asset Lifecycle," *Management-Proceedings of the 4th World Congress on Engineering Asset Management*, WCEAM 2009, pp. 420-429.

[4] S.D. Jang, B.W. Kang, "Frequency selective surface based passive wireless sensor for structural health monitoring," *Smart Materials and Structures*, vol. 22, pp. 221-231, June 2013. http://dx.doi.org/10.1088/0964-1726/22/2/025002

[5] S. Meena, N. K. Prakash, "Runtime reconfiguration of wireless sensor node using FPGA," *5th International Conference on Computing Communication and Networking Technologies*, ICCCNT 2014, pp. 269-274. http://dx.doi.org/10.1109/iccnt.2014.6963048

[6] N. Wu, J. Zhang, "Mobile nodes deployment scheme design based on perceived probability model in heterogeneous wireless sensor network," *Journal of Robotics and Mechatronics*, vol. 26, pp. 616-621, August 2014. http://dx.doi.org/10.20965/jrm.2014.p0616

[7] X.H. Yu, "Content of smart wireless sensor network security and its network security policy," *Computer Modelling and New Technologies*, vol. 18, pp. 569-573, June 2014.

[8] Y. Zhou, D. Xiao, "Policy-based distributed management for wireless sensor networks," *IET Conference Publications*, vol. 52, pp. 49-50, May 2006.

[9] J. Zhipeng, Z. Sanguo, "Reliability evaluation of wireless sensor networks using logistic regression," *2010 WRI International Conference on Communications and Mobile Computing*, CMC 2010, Vol. 3, pp. 334-338, April 2010.

[10] B. Steffen, L. Cyril, "Rollout Algorithms for Wireless Sensor Network-Assisted Target Search," *IEEE Sensors Journal*, Vol. 15, pp. 3835-3845, July 2015. http://dx.doi.org/10.1109/JSEN.2015.2393893

[11] K. Sanjeev, V. Poor, "Regression in sensor networks: Training distributively with alternating projections," *Proceedings of SPIE - The International Society for Optical Engineering*, Vol. 5910, pp. 1-15, April 2005.

[12] C. Aditi, V. Palaniandavar, "An efficient statistical approach for time synchronization in wireless sensor networks," *International Journal of Communication Systems*, Vol. 29, pp. 722-733, March 2016. http://dx.doi.org/10.1002/dac.2944

[13] J.B.Predd, S.R. Kulkarni, H.V. Poor, "Distributed kernel regression: An algorithm for training collaboratively," *2006 IEEE Information Theory Workshop*, ITW 2006, pp. 332-336. http://dx.doi.org/10.1109/ITW.2006.1633840

[14] C. Jongyoo, "Development and evaluation of an ambulatory stress monitor based on wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 16, pp. 279-286, March 2012. http://dx.doi.org/10.1109/TITB.2011.2169804

[15] O.E.K. Aktouf, I. Parissis, "SMART service for fault diagnosis in wireless sensor networks," *Proceedings - 6th International Con-*

ference on Next Generation Mobile Applications. NGMAST 2012, pp. 211-216.

[16] I. Aydin, M. Karaköse, "Wireless sensor network based fault diagnosis approaches," *2013 21st Signal Processing and Communications Applications Conference*, SIU 2013, pp. 201-211.

[17] S.H. Chang, "A causal model method for fault diagnosis in wireless sensor networks," *Proceedings-10th IEEE International Conference on Computer and Information Technology*, CIT-2010, pp. 155-162. http://dx.doi.org/10.1109/cit.2010.65

[18] C.L. He, "Technology research on the fault diagnosis of wireless sensor network system," *Advanced Materials Research*, vol. 84, pp. 442-445, April 2014.

[19] X.H. Jin, T.W.S. Chow, "Kuiper test and autoregressive model-based approach for wireless sensor network fault diagnosis," *Wireless Networks*, vol. 21, pp. 829-839, May 2015. http://dx.doi.org/10.1007/s11276-014-0820-0

[20] P.M. Khilar, "Intermittent fault diagnosis in wireless sensor networks," *Proceedings-10th International Conference on Information Technology*, ICIT 2007, pp. 145-147. http://dx.doi.org/10.1109/icit.2007.15

[21] Q. Li, "Wireless sensor network fault diagnosis method of optimization research and simulation," *Applied Mechanics and Materials*, vol. 347, pp. 955-959, May 2009.

AUTHORS

Weimin Han is an associate professor with Department of Wind engineering, Hunan Vocational institute of Technology, Xiangtan 411000, China. Her research interests include wind power engineering, information engineering and automation. (e-mail: B12090018@hnu.edu.cn).

Shijun Li is a doctoral student of College of Electrical and Information Engineering, Hunan University, Xiangtan 411101, China. He is mainly engaged in the study of computer control technology. He is also with Hunan province Cooperative Innovation Center for Wind Power Equipment and Energy Conversion, Xiangtan 411101, China (email: lishijun820404@qq.com).

Weidong Li is a master of Powerchina Zhongnan Engineering Co. Ltd., Changsha 410014, China. He is mainly engaged in communication technology research. (e-mail: Lsj181627@qq.com).