A Novel Nodes Deployment Assignment Scheme with Data Association Attributed in Wireless Sensor Networks

Zeyu Sun[1,2], Longxing Li[1], Xiaofei Xing[3], Zhiguo Lv[4], Neal N. Xiong[5]

[1] School of Computer Science and Engineering, Luoyang Institute of Science and Technology, China
[2] Department of Computer Science and Technology, Xi'an Jiaotong University, China
[3] School of Computer Science and Educational Software, Guangzhou University, China
[4] School of Telecommunications Engineering, Xi'dian University, China
[5] Department of Mathematics and Computer Science, Northeastern State University, USA
lylgszy@163.com, jsjsdoc@163.com, xxfcsu@163.com, 709009460@qq.com, xiongnaixue@gmail.com

Abstract

Data association plays a very important role in wireless sensor networks. This paper proposes a node deployment assignment method with data association attributes (NDADA). In this method, the relationship between the sensor nodes and the target nodes is used to give the three nodes correlation model, so as to reduce the energy consumption of nodes. In addition, the proposed algorithm uses the pre-distribution of random numbers in the node to distribute the original data of the whole network. The simulation results show that the proposed algorithm is effective and stable.

Keywords: Wireless sensor networks, Node deployment, Data association attributes, Data packets

1 Introduction

Wireless Sensor Networks (WSNs) are composed of a large number of low-cost sensor nodes with sensing, computing and communication capabilities, and widely used in various fields of science, such as national defense, medical treatment, exploration, environmental monitoring and so on [1-2]. Research of coverage quality is a fundamental problem in wireless sensor networks [3-4]. The deployments of sensor nodes have two kinds: deterministic deployment and random deployment. In the process of deterministic deployment, it mainly studies the optimal node deployment strategy in 2D and 3D space, and realizes the requirements of different wireless sensor networks coverage in an artificial way [5-6]. Most applications don't require the wireless sensor network has been completely covered, as long as the nodes in sensor networks can maintain a reasonable ratio of coverage of the monitoring area [7-8]. The random deployment of sensor nodes is a very important problem. It is a challenging problem that how to use the minimal sensor nodes to complete the coverage the specified monitoring area under the condition of satisfying the coverage rate [9-10]. In order to solve these problems, we need to take the following three measures:

(1) The monitored multi-target nodes are programmed into a square area and the correlation model between multi-sensor nodes and multi-target nodes are established.

(2) For sensor nodes sensing radius and required coverage rate, we use the expected value to calculate the maximum networks lifetime under the covering condition.

(3) Under the premise of ensuring the state balance of scheduling nodes and satisfying the quality of network coverage, a probability model is designed to make the results more accurate.

2 Related Works

Many researchers have come up with different solutions to solve the exposed terminals problem [11-14]. AMACA (CSMA/CA without carrier sensing) protocol abandons the channel listening and the ACK message while the data transmission is only controlled by the RTS/CTS message. The required transmission time T_d of the DATA message is included in the CTS message. The neighbors of the destination terminal will remain silent within the time interval T_d when they hear the CTS message, while the neighbors of the transmitting terminal may be permitted to perform transmissions. In this way, the exposed terminals problem can be solved. Since the neighbors of the transmitting terminal can also perform transmission during the dialogue, the CTS message of the destination terminal may not be correctly received by the transmitting terminals [15-17]. As a result, the network performance may deteriorate sharply when the system load is heavy.

A DBTMA (dual busy tone media access) protocol

---

[*] Corresponding Author: Xiaofei Xing; E-mail: xxfcsu@163.com

DOI: 10.3966/160792642019032002018

was proposed in paper based on the dual channel hardware structure [18]. The data is transmitted through a dialogue composed of the RTS+DATA message. When terminal A transmits the RTS message to terminal *B*, the busy tone BTt is transmitted through the control channel. After terminal *B* receives the RTS message, the response to BTt will be transmitted from terminal *B* through the control channel if the receiver of data from terminal A is permitted. The data transmission at terminal *A* is started upon receiving the BTt response from terminal B. During the data transmission, terminal *B* keeps transmitting BTt through the control channel while all the neighbors perform the back-off scheme once they detect the busy tone BTt. The data transmissions of the terminals are initiated according to the existence of the busy tone on the control channel, instead of the carrier sensing method. During the data transmission, the neighbors of terminal *A* are permitted to transmit data as long as they do not affect the receiving of terminal *B*, as a result of which the exposed terminals problem can be solved. Simulation results prove that the DBTMA protocol out-performs the MACA protocol by 140% in terms of the performance. However, without the ACK scheme, erroneous data of the DBTMA protocol requires further processing on higher layer protocols, which affects the transmission efficiency [19]. Furthermore, the dual channel structure of the DBTMA protocol increases the hardware complexity at the terminals.

The MACA-P protocol solves the exposed terminals problem by synchronizing multiple neighboring terminals to perform transmission in parallel [20]. The transmission of exposed terminals is aligned by employing the control interval. In this protocol, the dialogue process of normal terminals is identical to that of the 802.11DCF protocol while the dialogue of the exposed terminals is only composed of the DATA message and the ACK message. The RTS/CTS message for the dialogue of normal terminals includes the starting and ending time instances of the DATA message, between which the transmission for the DATA message of exposed terminals must be completed [21]. Therefore, the parallel transmission for exposed terminals is realized. Parallel transmission can be achieved for the MACA-P protocol when there is no conflict between the ACK messages. Simulation results show that, under certain network topologies, the MACA-P protocol achieves twice the throughput of the 802.11DCF protocol [22]. However, the major disadvantages of the MACA-P protocol are as follows. (1) The transmission of the DATA message is performed without the handshake process through the RTS/CTS message. As a result, correct receiver of the DATA message cannot be guaranteed due to the interference caused by the transmitting messages from the neighbors of the destination terminal. Hence, the hidden terminals problem can be caused. (2) Extra delay is needed before the transmission at the transmitting terminals due to the data alignment of the exposed nodes, which undermines the original dialogue efficiency. (3) The conflict between multiple ACK messages cannot be solved by the MACA-P protocol, i.e., this protocol is only applicable to the special scenarios where the transmitting area of the destination terminals does not overlap and no ACK message conflicts exist. Please, leave two blank lines between successive sections as here.

3 Problem Solutions

In general, the coverage is more directly reflect the target, the multi-target node caused by regional concern with high coverage, taking into account the sensor nodes in the monitoring area at the function relation between value and coverage area as shown in Figure 1.

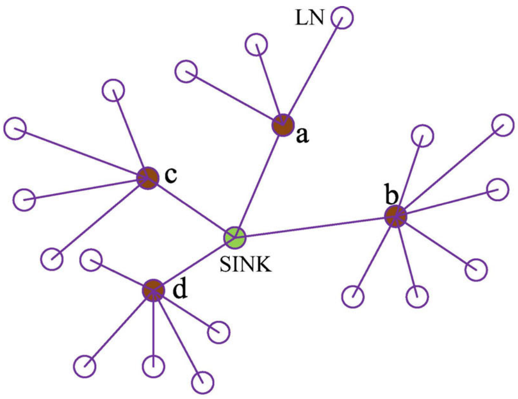

Figure 1. Data association and nodes coverage

In this wireless sensor network, there are three types of nodes: the leaf node LN (Leaf Node), the intermediate node IN (Intermediate Node) and the base station (Sink Node). In WSN, the data fusion operation must be completed in the process of data transmission, in order to reduce the amount of data communication, reduce the possibility of collision and improve the communication efficiency [23-24].

If the node receives the information d_i at time t, the data fusion model can be given as:

$$f(t) = f(d_1, d_2, \cdots d_n) \tag{1}$$

According to the definition of fusion, other typical data fusion functions are as follows:

$$f(t) = \sum_{i=1}^{n} d_i \tag{2}$$

$$f(t') = \sum_{i=1}^{n} \frac{d_i}{n} \tag{3}$$

$$f(t) = \min(d_i \mid i = 1, 2, \cdots \mathrm{n}) \tag{4}$$

$$f\left(t^{'}\right) = \max\left(d_i \mid i = 1,2,\cdots \mathrm{n}\right) \quad \textbf{(5)}$$

$$H\left(K \mid \left\{g\left(P_i\right) \mid P_i \in X\right\}\right) = 0 \quad \textbf{(6)}$$

$$\forall X \in \beta, H\left(K \mid \left\{g\left(P_i\right) \mid P_i \in X\right\}\right) = H\left(K^{'} \mid \left\{g\left(P_i\right) \mid P_i \in X\right\}\right) \textbf{(7)}$$

In which, *K'* denotes any element in a secret space.

Each pair (x_i,y_i) is a point on the curve *f*(*x*), because the *t* points uniquely determine a polynomial, so *k* can be reconstructed from the *t*. Given a *t* shared y_i, the *h*(*x*) reconstructed from the Lagrange polynomial is:

$$h\left(x\right) = \sum_{s=1}^{t} y_i \prod_{j=1}^{t} \frac{x - x_i}{x_{i_s} - x_{i_j}} \quad \textbf{(9)}$$

The coefficient matrix of cluster head *CHk* can be given as:

$$A^k = \begin{pmatrix} a_{00}^k & a_{01}^k & \cdots & a_{0,t-1}^k \\ a_{10}^k & a_{11}^k & \cdots & a_{1,t-1}^k \\ \cdots & \cdots & \cdots & \cdots \\ a_{t-1,0}^k & a_{t-1,1}^k & \cdots & a_{t-1,t-1}^k \end{pmatrix} \quad \textbf{(10)}$$

In which, $a_{ij}^k = a_{ji}^k \ \left(0 \le i, j \le t-1\right)$.

Then calculate the coefficients of the other *n*-1 cluster head node *ID* coefficient matrix, and sent to the corresponding cluster head CH_L.

$$P^{kl} = \left(p_{ij}^{kl}\right)_{t \times t} = \begin{pmatrix} p_{00}^{kl} & p_{01}^{kl} & \cdots & p_{0,t-1}^{kl} \\ p_{10}^{kl} & p_{11}^{kl} & \cdots & p_{1,t-1}^{kl} \\ \cdots & \cdots & \cdots & \cdots \\ p_{t-1,0}^{kl} & p_{t-1,1}^{kl} & \cdots & p_{t-1,t-1}^{kl} \end{pmatrix} \quad \textbf{(11)}$$

In which, $p_{ij}^{kl} = a_{ij}^k + \sum_{m-1}^{t-1} b_{ijm}^k \left(\beta\right)^m$ and b_{ijm}^k , b_{ijm}^k are a random selection from a finite field by CH_k. CH_K will be the key information sent to CH_L. After receiving the information of other nodes, CH_L decrypts the information and calculates,

$$p_{ij}^l = \sum_{k=1}^{n} p_{ij}^{kl} = \sum_{k=1}^{n} a_{ij}^k + \sum_{m=1}^{t-1} \sum_{k=1}^{n} b_{ijm}^k \beta^m \quad \textbf{(12)}$$

$$p_{ij}^l = a_{ij} + \sum_{m=1}^{t-1} b_{ijm} \beta^m \quad \textbf{(13)}$$

Therefore, the coefficient matrix is the sum of the coefficient matrix of all cluster head nodes [25-26]. A certain number of head nodes can be restored by the Lagrange polynomial of the main two element symmetric polynomial. It is assumed that each node has a pair of temporary keys when it is initialized [27-28].

Signal to interference and noise ratio is the effective signal strength and noise intensity received by the node is indicated by *γ*, and the node can receive the message correctly only when the signal-to-ratio received by the node is greater than a certain value [29]. When the node receives the node message, the signal to noise ratio can be expressed as:

$$\gamma_{(b \to a)} = 10lg \frac{\mathrm{Pr}_{(b \to a)}}{\phi + N_0} \quad \textbf{(14)}$$

Where N_0 is the ambient noise, and ϕ is the accumulation interference of node *N* for all nodes except the node *b* in the network.

Theorem 1. During the dialogue between terminal *a* and terminal *b*, when no node synchronization is performed, the newly-built dialogue of the exposed terminal *c* will cause no interference on the existing dialogue if *f*(*Cov*(*a*,*b*), *Cov*(*c*,*d*))≤1/(*N*+1), where terminal *d* is the destination of the exposed terminal *c* and $N = 10^{\gamma/10}$.

Proof. If the exposed terminal *c* is a neighbor of terminal *b* and terminal *a* is silent, then according to equation (14), the critical condition for terminal *b* to correctly receive the data from terminal *c* is:

$$\mathrm{Pr}_{(c \to b)} = \left(\phi + N_0\right) \times N \quad \textbf{(15)}$$

When terminal *a* and terminal *c* transmit data simultaneously, the condition for terminal *b* to correctly receive the data from terminal *a* is:

$$\left(\phi + \mathrm{Pr}_{(c \to b)} + N_0\right) \times N \le \mathrm{Pr}_{(a \to b)} \quad \textbf{(16)}$$

Substituting equation (15) into equation (16), the condition for terminal *b* to correctly receive the data from terminal *a* is:

$$\frac{\mathrm{Pr}_{(a \to b)}}{\mathrm{Pr}_{(c \to b)}} \ge N + 1 \quad \textbf{(17)}$$

Similarly, when terminal *b* and terminal *c* transmit data simultaneously, the condition for terminal *a* to correctly receive the data from terminal *b* is:

$$\frac{\mathrm{Pr}_{(b \to a)}}{\mathrm{Pr}_{(c \to a)}} \ge N + 1 \quad \textbf{(18)}$$

Generalizing equation (17) and equation (18) and, the condition for transmission of terminal *c* to pose no influence on *Cov*(*a*,*b*) is:

$$\frac{\mathrm{Pr}_{(b \to a)}}{\max\left(\mathrm{Pr}_{(c \to a)}, \mathrm{Pr}_{(c \to b)}\right)} \ge N + 1 \quad \textbf{(19)}$$

Similarly, we can derive the condition for transmission of terminal *d* to pose no influence on *Cov*(*a*,*b*) Therefore, the condition for *Cov*(*c*,*d*) to pose no influence on *Cov*(*a*,*b*) is:

$$\frac{\Pr_{(b\leftrightarrow a)}}{\max\left(\Pr_{(d\to a)},\Pr_{(d\to b)},\Pr_{(c\to a)},\Pr_{(c\to b)}\right)}\geq N+1 \quad \textbf{(20)}$$

Finally, we can derive the condition for the case that *Cov*(*a*,*b*) and *Cov*(*c*,*d*) do not influence each other, i.e. The proof is completed.

Corollary 1. When no node synchronization is performed, the condition for the exposed terminal *c* to build a new dialogue during the dialogue between terminal *a* and terminal *b* is:

$$f^{'}=\frac{DX}{DM}\geq N^{'} \quad \textbf{(21)}$$

where *d* is the destination terminal of terminal *c*, *DX*=min{*d*(*a*,*c*),*d*(*b*,*c*),*d*(*a*,*d*),*d*(*b*,*d*)}, *DM*=max{*d*(*a*,*b*), *d* (*c*,*d*)} and $N'=(N+1)^{1/\lambda}$.

Proof. According to equation (18), equation (19) and (20)

$$\frac{\Pr_{(a\to b)}}{\Pr_{(c\to b)}}=\frac{d(a,b)^{\lambda}}{d(c,b)^{\lambda}}\geq(N+1)^{\lambda} \quad \textbf{(22)}$$

According to the deriving process which is similar to that of Theorem 1, we have min{*d*(*a*,*c*),*d*(*b*,*c*),*d*(*a*,*d*), *d*(*b*,*d*)}/ max{*d*(*a*,*b*),*d* (*c*,*d*)}=*DX*/*DM*≥*N'*, The proof is completed.

4 Method of Realization

4.1 Neighbors Quantification Based Relay Strategy

When node *i* receives a data packet *p* whose destination address is *D*, a node from *N*(*i*) is chosen as the relay node for the next hop according to the relay strategy of the NDADA algorithm, which is elaborated as follows.

(1) If node *i* is exactly the destination of data packet *p*, the transmission of the data packet is successful.

(2) If node *j* from *N*(*i*) is the destination of data packet *p*, then the data packet is relayed to node *j* from node *i*.

(3) If the destination *D* lies within the wireless transmission area of node *j* from *N*(*i*), the data packet p is relayed to node *j* from node *i*. If multiple nodes from *N*(*i*) satisfies the above condition, *F*(*i*,*j*,*D*) is calculated according to equation (10) and equation (11) for those nodes. Then the node *j* with the maximal *F*(*i*,*j*,*D*) is chosen as the relay node.

(4) If none of the conditions are satisfied, then for an arbitrary node from *N*(*i*), the potential propulsion degree from this node to *D* is calculated according to the corresponding neighbors quantification value.

$$F_{tn}(i,j,D)=d(i,D)-\min \mathrm{G}_{t}(j,k,D)\ \ j\in N(i) \quad \textbf{(23)}$$

Where $G_t(j,k,D)$ is the approximate distance from the nodes in the *k*-th sector of node *j* to *D*, which can be calculated as follows.

$$G_{t}(i,j,D)=\left(G_{a}(j,k)^{n}+\left(d(j,D)-G(j,k)^{2}\right)\right)^{\frac{1}{2}} \quad \textbf{(24)}$$

$$\xi_{cs}(i)=\left\{j\,|\,\mathrm{F}_{tn}(i,j,D)>F(i,\mathrm{j},\mathrm{D})\right\} \quad \textbf{(25)}$$

Equation (25) Construct the set of candidate relay nodes $\xi_{cs}(i)$. If $\xi_{cs}(i)$ is not empty, then chose the node *j* with the maximal $F_{tn}(i,j,D)$ in $\xi_{cs}(i)$. As the relay node. Otherwise, data packet *p* is likely to suffer from routing void at any neighbor of node *i*. To reduce unnecessary energy consumption, the edge recovery scheme can be employed in advance at node *i*.

In comparison with the conventional greedy relay strategy, routing void can be avoided by the nodes utilizing the NDADA algorithm data packet. Therefore, the average energy consumption for data packet transmission can be reduced.

4.2 Data Flow Congestion Avoidance

The data packet is propelled towards the destination node along the shortest path utilizing the greedy relay strategy [30-32]. Normally, one data flow would consistently transmit data packets within a certain period. Traffic congestion can be caused if all the data packets are transmitted along the shortest path, as a result of which the transmission delay is increased. If the congestion degree further increases, the probability of data packet collision grows higher. Meanwhile, the packet loss rate increases due to the buffering overflow. Furthermore, employing the retransmission recovery scheme would increase the average energy consumption for transmission.

In order to solve this problem, the greedy relay strategy can be modified by introducing a greediness factor *g*, which indicates the expected minimal propulsion degree. The factor *g* is real-valued and ranges from 0 to 1. A small *g* would increase the path length and the energy consumption for transmission. On the contrary, if *g* is too large, the number of candidate nodes would decrease and congestion avoidance cannot be achieved [33-34]. Since the propulsion degree for the candidate node is expected to be at least 80% of the maximal propulsion degree here, we take the value *g*=0.8. To avoid data flow congestion, we construct the set of candidate relay nodes which satisfies the greediness factor based on the set $\xi_{cs}(i)$.

$$\xi_{GCS}(i)=\left\{j\,|\,F_{tn}(i,j,D)>g\ max\,F_{tn}\right\} \quad \textbf{(26)}$$

Where max F_{tn} is the maximal $F_{tn}(i,j,D)$ in the set $\xi_{cs}(i)$ for node *i*. A node with the smallest data flow congestion degree can be chosen as the relay node for the next hop as long as the number of nodes in $\xi_{cs}(i)$ is larger than one.

Considering the sharing property of the wireless links between the nodes in the sensor network, the nodes will have more chances to relay data packets when the surrounding links are idle. Otherwise, data packets will be buffered in the node array. The congestion degree of a node within a period of T_{cw}. is defined as:

$$\Psi_{cw}(i) = \frac{P_{in}(i)}{P_{tr}(i)} \tag{27}$$

Where $P_{tr}(i)$ is the number of successfully relayed data packets of node i and $P_{tr}(i)$ is the average array length for node i within T_{cw}.

$\Psi_{cw}(i)$ reflects the node congestion degree within T_{cw}. However, during the transmission of data packets, it is the future congestion degree after the arrival of data packets that we are always concerned about. In order to predict the future congestion degree for the nodes, we employ the exponentially weighted moving average method and define the node congestion degree as:

$$\Psi_c(i) = (1-\lambda)\Psi_c(i) + \lambda\Psi_{cw}(i) \tag{28}$$

Where λ is the weight factor which reflects the real-time varying extent for the node congestion degree. Experiments have shown that when the weight factor takes the value of 2, the mean squared error between the predicted value and the realistic value is the smallest. Therefore, we assume λ=2 here.

Considering the fact that the NDADA algorithm employs the greedy relay strategy based on two-hop neighbors, the congestion degree of neighbors should also be taken into consideration for the calculation of node congestion degree. The neighbors' congestion degree is define as the data flow congestion degree $\Psi_{dc}(i)$.

$$\Psi_{dc}(i) = (1-\lambda)\Psi_c(i) + \lambda A(j) \tag{29}$$

Where $A(j)$ is the average congestion degree for the neighbor j of node i and λ is the weight factor for the neighbors.

In order to obtain the data flow congestion degree of the neighbors, $\Psi_{dc}(i)$ of node i can be included in the Hello message. Therefore, we assume T_{cw} to be equal to the exchange period of Hello message here.

4.3 Node Energy Consumption Balance

As is mentioned above, the choice of the relay node for the next hop would influence the network performance to some extent when there are multiple data flows. The reason is that when one node lies on the greedy relay path of multiple data flows, the node energy would be exhausted quickly. If one node is the only greedy relay path for one of the data flows, the failure of this node will cause routing void to this data flow, which will result in sharply increased average energy consumption for this data flow.

In order to solve this problem, the remained energy of the node and the number of data flows being relayed have to be taken into consideration to choose the relay node for the next hop. The data flows corresponding to each data packets have to be identified for the counting of data flows. Here, the data flows are distinguished by the source and destination of the data packets. The information on the source node and destination node is extracted from the received data packets by the nodes. If the data flow passes the node for the first time, it will be recorded in the data flow table at the nodes. Otherwise, corresponding records will be updated in the data flow table. Considering the fact that the memory of sensor nodes is limited, only the data flow within the latest period T_{cw} is remained. The number of data flows is:

$$I_n(i) = (1-\lambda)I_n(i) + \lambda I_{win}(i) \tag{30}$$

Where $I_{win}(i)$ is the number of data flows within the latest period T_{dw}. Just like the data flow congestion degree $\Psi_{dc}(i)$ of nodes, $I_n(i)$ is also included in the Hello message. Furthermore, T_{dw}=T_{cw}.

$I_n(i)$ reflects the importance of node i to the data packet transmission in the network. To avoid the increase of energy consumption for data packet transmission caused by the failure of important nodes, more energy has to be reserved for the nodes with larger $I_n(i)$, which is referred to as the node energy consumption balance scheme.

The energy balance degree of node i is defined as:

$$B_e(i) = \frac{E(i)}{I_n(i)} \tag{31}$$

Where $E(i)$ is the remained energy of node i. A larger $E(i)$ means that node i hopes to relay more data flows or data packets. After the introduction of the node energy balance degree, the relay node weight of each node j in $\xi_{GCS}(i)$ is calculated as follows for node i when the relay node for the next hop is chosen based on the greedy relay strategy.

$$W(j) = \frac{B_e(j)}{\Psi_{de}(j)} \quad j \in \xi_{GCS}(i) \tag{32}$$

Then the node with the largest $W(j)$ is chosen as the relay node for the next hop. It is shown from the definition of $W(j)$ that node i usually choose the neighbor with smaller congestion degree and higher energy balance degree as the relay for the next hop.

4.4 Data Fusion

In the phase of data fusion, sensor nodes use sensitive data and private random data as their own data [35-36]. The intermediate node p collects its child node information, according to the information as well as the p sub node list children set, can't get the p to send data to the child node, and in the p record the identity of the node failure-node.

At the same time, the HASH integrity of the received data is verified, and the failure of integrity verification is regarded as the node's failure. After the parent node receives the data, it is decrypted using the shared key. The hash value is re calculated to verify the data integrity. At the end of the fusion, each intermediate node detects whether it contains the failure node, and the node containing failure node will report the failure information along the tree structure to the sink node [37-38]. In the ideal environment, the number of failed nodes caused by packet loss or collision is very small, so the data traffic generated by reporting the failure information does not increase the network communication burden.

The sink node will be received from the first layer of nodes and the fusion results in the fusion failure node list failures when data fusion stage is completed, because the child nodes of the parent node information must be fused to upload. The node information of the immediate descendant of all failed nodes is not converged to the final fusion results, and should also be regarded as the failure nodes:

$$Fail_{agg} = \bigcup_{p\in failures} \left(fnode_p \bigcup subtree\left(fnode_p \right) \right) \quad \textbf{(33)}$$

All failed node sets in WSNs are:

$$Fail_{total} = Fail_{tree} \bigcup Fail_{agg} = \left\{ node_1, node_2, \cdots node_n \right\} \quad \textbf{(34)}$$

In WSNs, the data fusion results of the effective nodes are:

$$K_{agg} = L_0 + L_1 + \cdots L_n \quad \textbf{(35)}$$

Finally, the fusion result is the fusion value of sink minus the random number of all the valid nodes:

$$K = K_{agg} - \left(\sum_{i=0}^{n} \sec\ data_i - \sum_{j=1}^{n} \sec\ data_j \right) \quad \textbf{(36)}$$

Therefore, it can be obtained:

$$\frac{K_r}{K_t} = \frac{K_{agg} - \left(\sum_{i=0}^{n-1} \sec\ data_i - \sum_{j\in Failure} \sec\ data_j \right)}{\sum_{i=0}^{n-1} raw\ data_i - \sum_{j\in Failure} raw\ data_j} \quad \textbf{(37)}$$

$$\frac{K_r}{K} = \frac{K_{agg} - \left(\sum_{i=0}^{n-1} \sec\ data_i - \sum_{j\in Failure} \sec\ data_j \right)}{\sum_{i=0}^{n-1} raw\ data_i} \quad \textbf{(38)}$$

5 Simulations and Result Analysis

In order to verify the effectiveness and stability of the Energy-efficient Coverage Control with Multi-nodes Redundancy Verification, NDADA Algorithm, this paper selects the MATLAB7 as the simulation platform. And the NDADA Algorithm is compared with the four algorithms of the paper [35] (Energy-efficient Target Coverage Algorithm, ETCA) and the paper [37] (Multi-targets Cycle Control Coverage Preservation Protocol, MTCPP). And the simulation parameters are shown in the Table 1.

Table 1. The tabulation of the simulation parameters

parameter	value	parameter	value
area I	100*100m^2	R_c	10m
area II	200*200m^2	$E_{\text{R-elec}}$	50J/b
areaIII	300*300m^3	$E_{\text{T-elec}}$	50J/b
R_s	5m	ε_{fs}	10(J/b)/m^2
energy	5J	ε_{amp}	100(J/b)/m^2
time	800s	e_{min}	0.005J

The Figure 2 provides the different network and sensor nodes energy consumption in different monitoring area, and the contracted algorithms are the algorithms in the paper [35] and paper [37]. At the initial time of the three method, the network lifetime is increasing with the increase number of the sensor nodes. However, the behavioral parameter range limitation in this paper and closing statement of the redundant sensor nodes make the network lifetime of the algorithm in this paper, which is smaller than the other two method, when the network and sensor nodes energy reaches balancing statement. Figure 3 provides the network lifetime and the target node. When the target nodes are being covered, the network energy of this paper is also smaller than that of the other two algorithms, and the reason is the same as the above situation. In the Figure 4, it makes some of redundant sensor nodes in the working statement, which has prolong the network lifetime. When the parameter λ=3, the network lifetime of NDADA algorithm in this paper is larger than that of the ETCA Algorithm; when the parameter λ=4, the network lifetime is larger than that of the other algorithms; and the Figure 5 provides the change of the network lifetime comparison of the coverage of the multi-target nodes. With the increasing of the multi-target nodes, the network lifetimes of the three method have been the declining trance, and finally provide the energy balance statement. But during the declining process, the average speed of the algorithm in this paper is smaller than that of the other algorithms, the main reason of the sensor nodes are deployed in the monitoring coverage area, where are much too intensive, that is, the expected value of the coverage area is a little larger, and some of the redundant sensor nodes are waken up to be working sensor nodes, which increases the coverage probability and prolongs the network lifetime. And so as for the monitoring area of 300*300m^2, and the same reason. According to the analysis from the Figure 4 to the

Figure 5, the algorithm in this paper has more energy during the running time; the reason is that in the coverage process, the algorithm in this paper closes some of the redundant sensor nodes, which greatly prolongs the network lifetime. Thus, numbers of same sensor nodes, running time of the method in this paper is a little higher than that of the other algorithms, as shown in the Figure 6 and Figure 7.

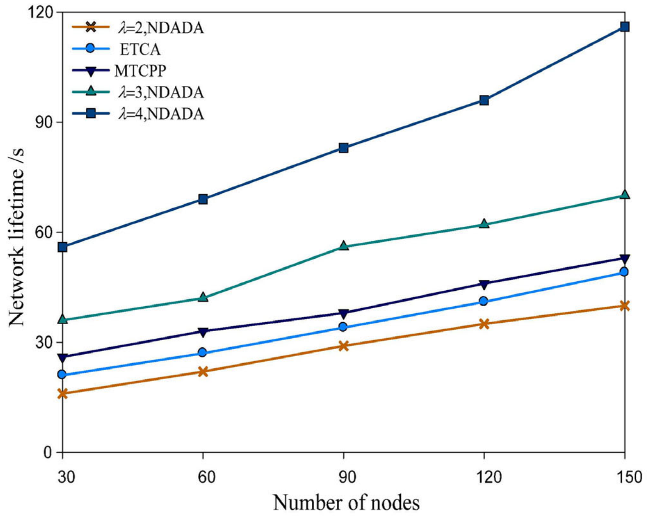

Figure 2. 100*100m^2, network lifetime and sensor nodes

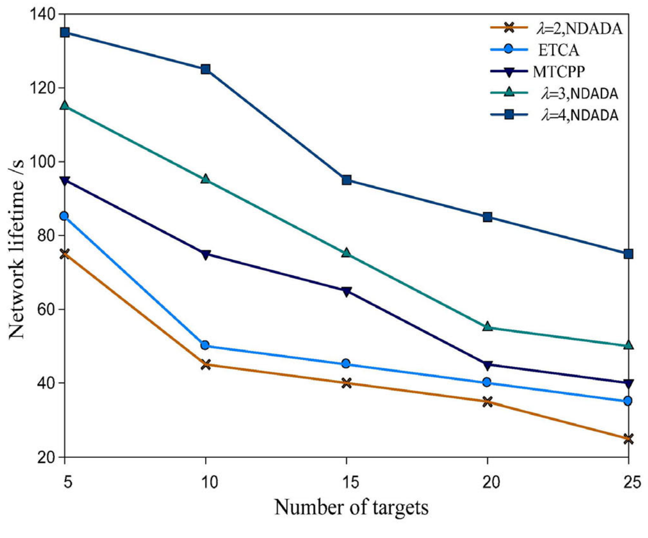

Figure 3. 100*100m^2, network lifetime and multi-target nodes

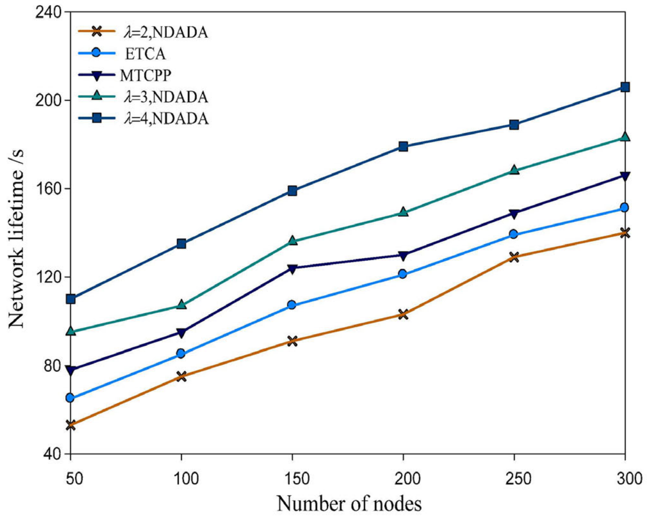

Figure 4. 200*200m^2, network lifetime and sensor nodes

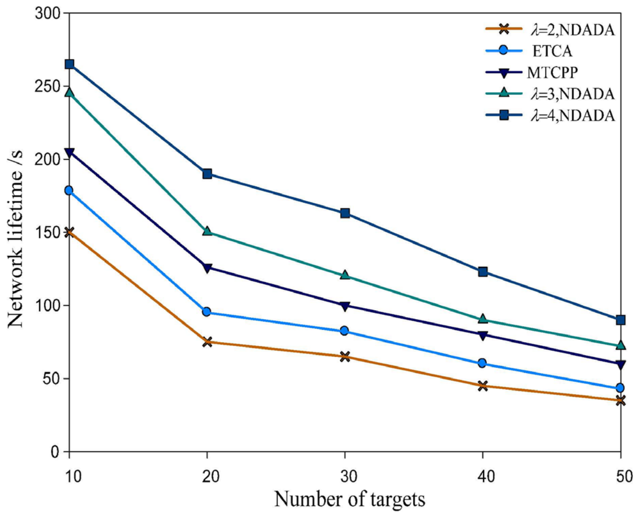

Figure 5. 200*200m^2, network lifetime and multi-target nodes

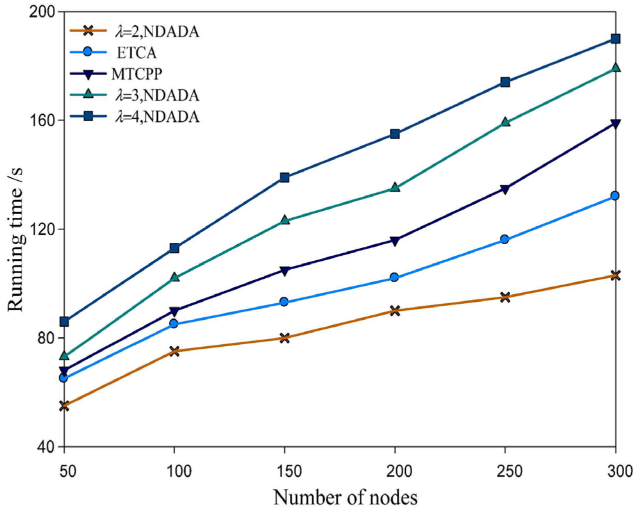

Figure 6. 200*200m^2, running time and sensor nodes

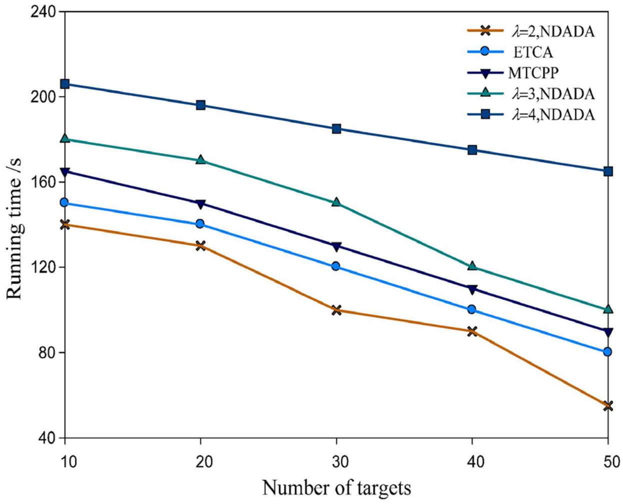

Figure 7. 200*200m^2, running time and multi-target nodes

In the amount of data communication, the NDADA algorithm is compared with TAG algorithm and SMART (λ=2, 3, 4) that are shown in Figure 8 to Figure 10.

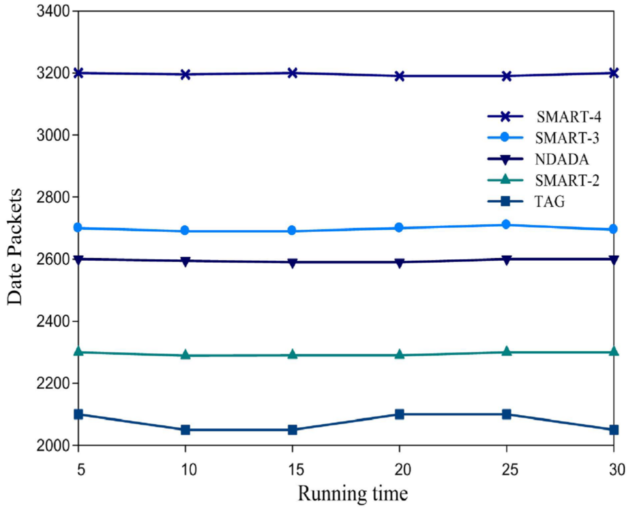

Figure 8. 100*100m^2, comparison of data packets

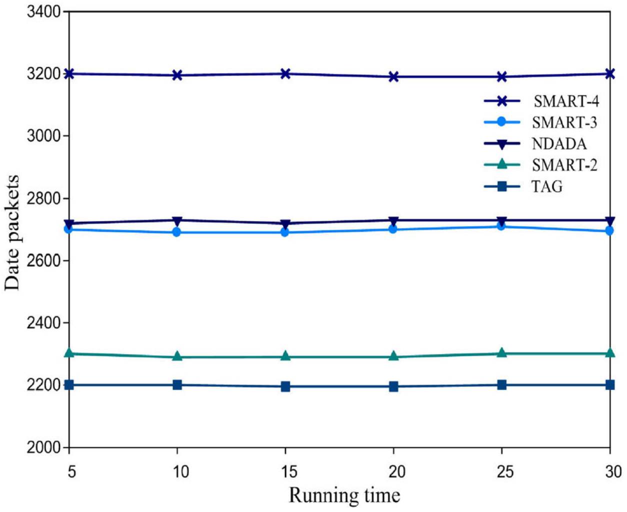

Figure 9. 200*200m^2, comparison of data packets

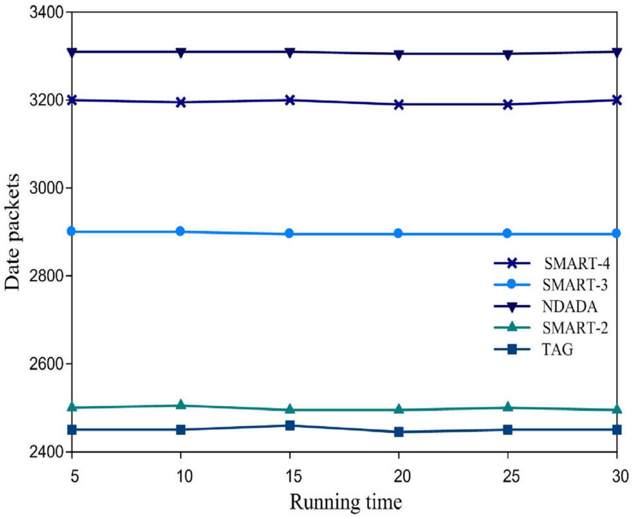

Figure 10. 400*400m^2, comparison of data packets

Energy consumption is an important index to measure the effectiveness of the algorithm. In WSNs, the TAG algorithm will have the sensor nodes of sensitive data fusion, but does not provide the function of privacy protection. SMART and NDADA algorithm introduces privacy protection mechanism, which is bound to increase the algorithm's energy consumption. From the analysis we can know that the energy consumption is mainly concentrated in the calculation and data transceiver two parts. At the same time, several algorithms are similar in calculation, and computational overhead does not affect the energy consumption. By comparing the amount of data communication, we can indirectly measure the characteristics of the three algorithms in terms of energy consumption. It can be seen that the results are consistent with the theoretical analysis. The simulation value is slightly smaller than the theoretical analysis value, because of the impact of the network environment, the data in the process of the data packet collision, collision, resulting in the phenomenon of packet loss. In NDADA, with the increase of the fusion time slice, the number of nodes in the data packet lost is found to be reduced, so there is a trend of decrease in the amount of data communication.

For each type of network topology, twelve node deployment scenarios are randomly generated according to the uniform and random principle. The running time for each experiment is 800s while 10 data flows are randomly chosen from the nodes. The direct discard strategy is adopted when the data packets suffer from routing void and the final data is obtained by averaging on the results of the twelve experiments.

Figure 11 illustrates the results on the average number of hops for the paths of data packets. Observe that when the number of neighbors is smaller than 15, the number of hops increases with the number of neighbors, especially when the number of neighbors is 5. There are fewer hops for the TNEB algorithm [8] than for the NDADA and the MTTA algorithm [39].

This is due to the fact that, when there are fewer neighbors, more routing voids exist in the network. Since the NDADA algorithm and the MTTA algorithm employ the two hop based relay strategy, routing voids can be predicted and avoided in advance. Therefore, data transmission can still be completed for the NDADA algorithm and the MTTA algorithm even when the source node and the destination node of the data flow are far from each other. However, data transmission is only possible for the TNEB algorithm when the source is close to the destination. On the other hand, when the number of neighbors exceeds 15, the number of hops tends to decrease with the number of neighbors. This is due to the fact that routing voids tend to vanish with the increasing number of neighbors. Consequently, the average distance between the source and the destination of the data flow becomes smaller.

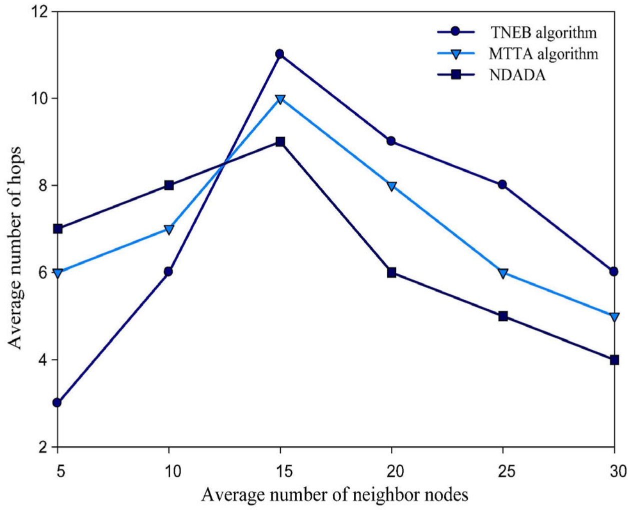

Figure 11. Average number of hops vs average number of neighbors

The average energy consumption for transmitting one data packet at different time instances is presented in Figure 12. Note that the average number of nodes is smaller than 15 for the network. It can be observed that the average energy consumption tends to increase with time. For the first 350s of the experiment, the energy consumption of the NDADA and the MTTA algorithm is higher than that of the TNEB algorithm. This is because the information of the neighbors within the range of two hops has to be exchanged through the Hello message for the NDADA and the MTTA algorithm. Therefore, extra overhead is needed for these two algorithms when compared with the TNEB algorithm. Since the node energy balance is not considered by the greedy relay strategy of the TNEB and the MTTA algorithm, when the energy of nodes on the shortest path runs out, the length of paths for subsequent data packets would be increased obviously. As a result, the average energy consumption is also increased. However, the remained energy and the data flow of the nodes are taken into consideration for the greedy relay of the NDADA algorithm. Therefore, the relay for the data packets in the NDADA algorithm is chosen in a more reasonable way, which explains the slow increase for the energy consumption curve of the NDADA algorithm in Figure 12. At the final stage of the experiment, the NDADA algorithm reduces it energy consumption by 23.6% and 37.1% in comparison with the MTTA and the TNEB algorithm.

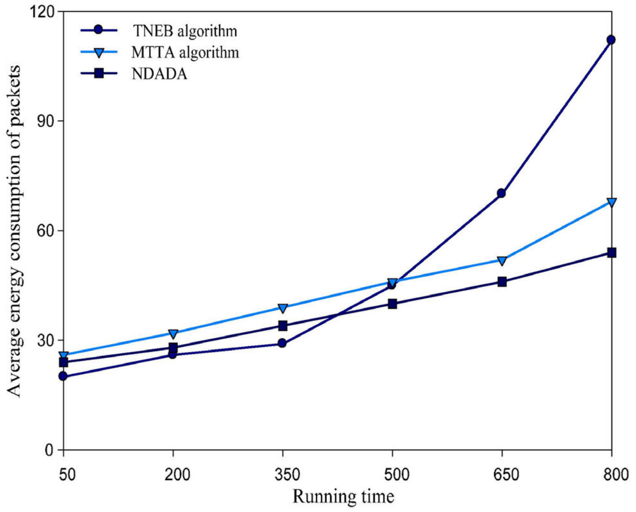

Figure 12. Average energy consumption for data packet transmission at different running time

The average end-to-end delay of data packets is illustrated in Figure 13. The end-to-end delay is determined, to some extent, by the number of hops of data packets. Therefore, the delay curves in Figure 13 shows much in common with the curves in Figure 11. However, the NDADA algorithm exhibits an obviously lower delay than the TNEB algorithm and the MTTA algorithm when the number of neighbors is large. This is because, on the wireless shared link, the available bandwidth for each node becomes smaller with the increase of neighbors. As a result, much longer time is required to successfully relay one data packet. However, the NDADA algorithm would give priority to the nodes with smaller link congestion degree for the choice of relay nodes. Hence, it requires shorter delay than the TNEB and the MTTA algorithm. When the number of neighbors is 15, the NDADA algorithm reduces its end-to-end delay by 17.7% and 26.8% in comparison with the MTTA and the TNEB algorithm, respectively.

6 Conclusions

This paper proposes a node deployment assignment method with data association attributes (NDADA). Aided by the information of the neighbors within the range of two hops, according to the dialogue correlation model, dialogues with small correlation are allowed to be performed in parallel. Therefore, the exposed terminal problem can be solved and network

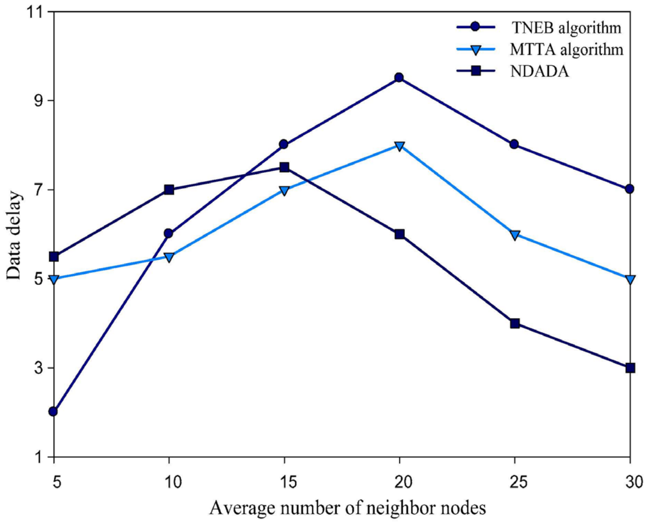

Figure 13. Average end-to-end data delay for packets

spatial efficiency can be increased. The NDADA algorithm employed the greedy relay strategy based on the information of neighbors within two hops. In order to prevent data packets from being relayed to congested links, the data flow congestion degree of neighbors was taken into consideration. This algorithm further exploited the node energy consumption balance degree to prevent the influence on data packet transmission caused by the failure of important nodes. The proposed algorithm required no node synchronization or update for the hardware. In comparison with existing algorithms [8], algorithm [35], algorithm [37] and algorithm [39], NDADA algorithm shows better scalability and efficiently deals with the impacts of exposed terminals on the network throughput.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant No. 61628210; Henan Province Education Department Cultivation Young Key Teachers in University of Under Grant No. 2016GGJS-158; Henan Province Education Department Natural Science Foundation under Grant No. 17A520044; Luoyang Institute of Science and Technology High-level Research Start Foundation under Grant No. 2017BZ07; Natural Science and Technology Research of Henan Province Department of Science Foundation under Grant No. 182102210116 and No. 162102210113; Guangdong Natural Science Foundation of China under Grant No. 2016A030313540; Guangzhou Education Bureau Science Foundation under Grant No. 1201430560; Science and Technology Planning Project of Guangzhou under Grant No. 201707010284; Shaanxi Education Bureau Science Foundation under Grant No. 2016SF-428.

References

[1] Z. Y. Sun, W. G. Wu, H. Z. Wang, H. Chen, W. Wei, An Optimization Strategy Coverage Control Algorithm for WSN, *International Journal of Distributed Sensor Networks*, Vol. 2014, No. 1, pp. 1-12, July, 2014.

[2] L. Shu, Y. Zhang, L. T. Yang, Y. Wang, M. Hauswirth, N. Xiong, TPGF: Geographic Routing in Wireless Multimedia Sensor Networks, *Telecommunication Systems,* Vol. 44, No. (1-2), pp. 79-95, June, 2010.

[3] N. Xiong, A. V. Vasilakos, L. T. Yang, L. Song, Y. Pan, R. Kannan, Y. Li, Comparative Analysis of Quality of Service and Memory Usage for Adaptive Failure Detectors in Healthcare Systems, *IEEE Journal on Selected Areas in Communications*, Vol. 27, No. 4, pp. 495-509, April, 2009.

[4] A. A. Mogren, M. Iftikahar, M. Imran, N. X. Xiong, S. Guizani, Performance Analysis of Hybrid Polling Schemes with Multiple Classes of Self-similar and Long-range Dependent Traffic Input, *Journal of Internet Technology*, Vol. 16, No. 4, pp. 615-627, August, 2015.

[5] Y. Yoon, Y. H. Kim, An Efficient Genetic Algorithm for Maximum Coverage Deployment in Wireless Sensor Network, *IEEE Transactions on Cybernetics*, Vol. 43, No. 5, pp. 1473-1483, May, 2013.

[6] K. X. Huang, Q. Zhang, C. J. Zhou, N. X. Xiong, Y. X. Qin, An Efficient Intrusion Detection Approach for Visual Sensor Networks Based on Traffic Pattern Learning, *IEEE Transactions on Systems, Man, and cybernetics: Systems*, Vol. 1, No. 99, pp. 1-10, May, 2017.

[7] G. Wei, A. V. Vasilakos, Y. Zheng, N. Xiong, A Game-theoretic Method of Fair Resource Allocation for Cloud Computing Services, *The Journal of Supercomputing*, Vol. 54, No. 2, pp. 252-269, February, 2010.

[8] H. Chen, D. P. Qian, W. G. Wu, H. Z. Wang, Neghbors Quantification Based Energy Balance Routing in Wireless Sensor Networks, *Journal of Xi'an Jiaotong University*, Vol. 46, No. 4, pp. 1-6, April, 2012.

[9] S. Mini, S. K. Udgata, S. L. Sabat, Sensor Deployment and Scheduling for Target Coverage Problem in Wireless Sensor Networks, *IEEE Sensors Journal*, Vol. 14, No. 3, pp. 636-644, October, 2014.

[10] H. Cheng, N. Xiong, A. V. Vasilakos, L. T. Yang, G. Chen, X. Zhuang, Nodes Organization for Channel Assignment with Topology Preservation in Multi-radio Wireless Mesh Networks, *Ad Hoc Networks*, Vol. 10, No. 5, pp. 760-773, May, 2012.

[11] N. Tamboli, Y. Mohamed, Coverage-Aware Connectivity Restoration in Mobile Sensor Networks, *Journal of Network and Computer Applications*, Vol. 33, No. 4, pp. 363-374, May, 2010.

[12] M. Cardei, J. Wu, Energy-efficient Coverage Problems in Wireless Ad-hoc Sensor Networks, *Advanced Computing An Internationl Journal*, Vol. 29, No. 4, pp. 413-420, February, 2006.

[13] Y. C. Tseng, P. Y. Chen, W. T. Chen, *k*-angle Object Coverage Problem in a Wireless Sensor Network, *IEEE*

Sensors Journal, Vol. 12, No. 12, pp. 3408-3416, December, 2012.

[14] H. Li, H. Y. Yu, B. W. Yang, A. Liu, Timing Control for Delay-constrained Data Aggregation in Wireless Sensor Networks, *International Journal of Communication Systems*, Vol. 20, No. 7, pp. 875-887, July, 2007.

[15] J. Zhang, H. Zhang, X. Jia, A Data-aggregation-centric Wireless Sensor Network Routing Architecture, *Journal of Information and Computational Science*, Vol. 3, No. 3, pp. 635-642, May, 2012.

[16] G. Huang, A. Y. Zomaya, F. C. Delicato, P. F. Pires, An Accurate On-demand Time Synchronization Protocol for Wireless Sensor Networks, *Journal of Parallel & Distributed Computing*, Vol. 72, No. 10, pp. 1332-1346, October, 2012.

[17] H. Mahboubi, K. Moezzi, A. G. Aghdam, K. Sayrafian-Pour, V. Marbukh, Distributed Deployment Algorithms for Improved Coverage in a Network of Wireless Mobile Sensors, *IEEE Transactions on Industrial Informatics*, Vol. 10, No. 1, pp. 163-174, February, 2014.

[18] Q. F. Fan, K. Zeitouni, N. X. Xiong, Q. L. Wu, A General Nash Equilibrium Semantic Cache Algorithm in a Sensor Grid Database System, *IEEE Transactions on Systems, Man and Cybernetics: Systems*, Vol. 4, No. 99, pp. 1-12, April, 2016.

[19] S. Sengupta, M. Chatterjee, S. Ganguly, Improving Quality of Voio Streams Over WiMax, *IEEE Transaction on Computers*, Vol. 57, No. 2, pp. 145-156, February, 2008.

[20] H. J. Cheng, Z. H. Su, N. X. Xiong, Y. Xiao, Energy-Efficient Nodes Scheduling Algorithms for Wireless Sensor Networks Using Markov Random Field Model, *Information Science*, Vol. 329, No. 1, pp. 461-477, February, 2016.

[21] Z. Y. Sun, R. Tao, L. X. Li, X. F. Xing, A New Energy-Efficient Multi-target Coverage Control Protocol Using Event-driven-mechanism in Wireless Sensor Networks, *Internet Journal of Online Engineering*, Vol. 13, No. 2, pp. 53-67, February, 2017.

[22] Q. F. Fan, N. X. Xiong, K. Zeitouni, Q. L. Wu, A. Vasilakos, Y. C. Tian, Game Balanced Multi-factor Multicast Routing in Sensor Grid Networks, *Information Sciences*, Vol. 367-368, No. 1, pp. 550-572, January, 2016.

[23] E. Amiri, H. Keshavarz, M. Alizadeh, M. Zamani, T. Khodadadi, Energy Efficient Routing in Wireless Sensor Networks Based on Fuzzy Ant Colony Optimization, *International Journal of Distributed Sensor Networks*, Vol. 16, No. 1, pp. 579-597, July, 2014.

[24] H. Wang, H. Yi, L. N. Wu, F. C. Zhou, N. X. Xiong, Mining Probabilistic Representative Gathering Patterns for Mobile Sensor Data, *Journal of Internet Technology*, Vol. 18, No. 2, pp. 321-332, May, 2017.

[25] M. Wu, L. S. Tan, N. Xiong, Data Prediction, Compression, and Recovery in Clustered Wireless Sensor Networks for Environmental Monitoring Applications, *Information Science*, Vol. 329, No. 1, pp. 800-818, February, 2016.

[26] G. X. Tong, G, N. Wu, J. Tian, N. X. Xiong, A. V. Vasilakos, A Digital Noise Reduction Scheme in Communication Systems for Internet of Things, *Journal of Internet Technology*, Vol. 17, No. 5, pp. 879-887, October, 2016.

[27] Z. Y. Sun, Y. X. Shu, X. F. Xing, W. Wei, H. B. Song, W. Li, LPOCS: A Novel Linear Programming Optimization Coverage Scheme in Wireless Sensor Networks, *Ad Hoc & Sensor Wireless Networks*, Vol. 33, No. 1-4, pp. 173-197, October, 2016.

[28] F. Xia, R. N. Hao, J. Li, N. X. Xiong, L. T. Yang, Y. Zhang, Adaptive GTS Allocation in IEEE 802.15.4 for Real-Time Wireless Sensor Networks, *Journal of Systems Architecture*, Vol. 59, No. 10, pp. 1231-1242, November, 2013.

[29] H. J. Cheng, Y. Z. Chen, N. X. Xiong, F. F. Li, Layer-Based Data Aggregation and Performance Analysis in Wireless Sensor Networks, *Journal of Applied Mathematics*, Vol. 2013, No. 1, pp. 1-12, June, 2013.

[30] T. Wang, Z. Peng, J. B. Liang, S. Wen, MD Z. A. Bhuiyan, Y. Q. Cai, J. N. Cao, Follow Targets for Mobile Tracking in Wireless Sensor Networks, *ACM Transactions on Sensor Networks*, Vol. 12, No. 4, pp. 31.1-31.24, April, 2016.

[31] Z. Y. Sun, Y. L. Yun, H. B. Song, H. H. Wang, CCAJS: A Novel Connect Coverage Algorithm Based on Joint Sensing Modes for Wireless Sensor Networks, *KSII Transactions on Internet and Information Systems*, Vol. 10, No. 10, pp. 5014-5033, October, 2016.

[32] P. F. Wu, F. Xiao, C. Sha, H. P. Huang, R. C. Wang, N. X. Xiong, Node Scheduling Strategies for Achieving Full-view Area Coverage in Camera Sensor Networks, *Sensors*, ID:191528, Vol. 17, No. 6, pp. 1303, June, 2017.

[33] N. X. Xiong, R. W. Liu, M. H. Liang, D. Wu, Z. Liu, H. S. Wu, Effective Alternating Direction Optimization Methods for Sparsity-Constrained Blind Image Deblurring, *Sensors*, Vol. 17, No. 1, pp. 174, January, 2017.

[34] T. Wang, W. J. Jia, G. L. Xing, M. M. Li, Exploiting Statistical Mobility Models for Efficient WiFi Deployment, *IEEE Transactions on Vehicular Technology*, Vol. 62, No. 1, pp. 360-373, January, 2013.

[35] X. F. Xing, G. J. Wang, J. Li, Polytype Target Coverage Scheme for Heterogeneous Wireless Sensor Networks Using Linear Programming, *Wireless Communications and Mobile Computing*, Vol. 14, No. 14, pp. 1397-1408, September, 2014.

[36] W. Fang, X. Liang, S. Li, L. Chiaraviglio, N. Xiong, VMPlanner: Optimizing Virtual Machine Placement and Traffic Flow Routing to Reduce Network Power Costs in Cloud Data Centers, *Computer Networks*, Vol. 57, No. 1, pp. 179-196, January, 2013.

[37] C. F. Li, Z. Y. Sun, H. H. Wang, H. B. Song, A Novel Energy-efficient *k*-Coverage Algorithm Based on Probability Driven Mechanism of Wireless Sensor Networks, *International Journal of Distributed Sensor Networks*, Vol. 2016, No. 1, pp. 1-9, January, 2016.

[38] Z. Y. Sun, W. G. Wu, H. Z. Wang, H. Chen, X. F. Xing, A Novel Coverage Algorithm Based on Event-probability-driven Mechanism in Wireless Sensor Network, *EURASIP Journal on Wireless Communications and Networking*, Vol. 2014, No. 1, pp. 1-17, April, 2014.

[39] T. Wang, W. J. jia, B. N. Zhong, H. Tian, G. L. Zhang, BlueCat: An Infrastructure-free System for Relative Mobile

Localization, *Ad Hoc & Sensor Wireless Networks*, Vol. 29, No. 1-4, pp. 133-152, February, 2015.

Biographies

Zeyu Sun received Ph.D. degree in Xi'an Jiaotong University, China, in 2017. He is an assistant professor in the School of Computer and Information Engineering, Luoyang Institute of Science and Technology, Luoyang, Henan, China. He is a member of China Computer Federation. His research interests lie in wireless sensor networks, parallel computing and Internet of Things.

Longxing Li received the Ph.D. degree from University of Science and Technology Beijing, Beijing, China, in 2006. He is a professor in Luoyang Institute of Science and Technology, Luoyang, Henan, China. His research interests lie in wireless sensor networks, parallel computing and Internet of Things.

Xiaofei Xing received Ph.D. degree in Central South University, China, in 2012, respectively. He has been a Research Fellow at University of Tsukuba, Japan, and he had been also a Post-doctorate researcher in Applied Mathematics. His research interests include wireless sensor networks, mobile computing, network performance analysis.

Zhiguo Lv is currently a Ph.D. candidate at Xidian University, Xi'an, China. His research interests include Compressive Sensing, channel estimation and MIMO techniques.

Neal N. Xiong is currently an Associate Professor at Department of Mathematics and Computer Science, Northeastern State University, OK, USA. He received his both Ph.D. degrees in Wuhan University (about sensor system engineering), and the Japan Advanced Institute of Science and Technology (about dependable sensor networks), respectively. His research interests include Cloud Computing, Security and Dependability, Parallel and Distributed Computing, Networks, and Optimization Theory.