Received May 7, 2019, accepted May 20, 2019, date of publication May 23, 2019, date of current version June 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2918615

Mobile Intelligent Computing in Internet of Things: An Optimized Data Gathering Method Based on Compressive Sensing

ZEYU SUN[1,2], XIAOFEI XING[3], BIN SONG[4], YALIN NIE[1,2], AND HONGXIANG SHAO[1]

[1]School of Computer Science and Information Engineering, Luoyang Institute of Science and Technology, Luoyang 471023, China
[2]Key Laboratory of Intelligent of IoT, Luoyang Institute of Science and Technology, Luoyang 471023, China
[3]School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China
[4]School of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China

Corresponding author: Xiaofei Xing (xxfcsu@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant U1604149, in part by the Henan Province Education Department Cultivation Young Key Teachers in University under Grant 2016GGJS-158, in part by the Luoyang Institute of Science and Technology High-Level Research Start Foundation under Grant 2017BZ07, in part by the Henan Province Education Department Natural Science Foundation under Grant 19A520006 and Grant 19B510007, in part by the Guangzhou Education Bureau Science Foundation under Grant 1201430560, in part by the Science and Technology Planning Project of Guangzhou under Grant 201707010284, and in part by the Key Science and Technology Program of Henan Province under Grant 192102210249 and 182102210428.

ABSTRACT In order to alleviate the impacts of the rapid network energy exhaustion and the unreliable links on the data gathering in the Internet of Things (IoT), mobile intelligent computing based on compressive sensing date gathering (MIC-CSDG) algorithm is proposed in this paper, which could improve the data reconstruction accuracy. We conduct research from the following four links. First, this method employs mobile intelligent computing to derive the multi-hop function among sensor nodes, which is further utilized to determine the proportional relationship for the sensor nodes. Second, based on the sparse matrix, an observation matrix is designed with low correlation to mitigate the influences of the data packet loss on the entire IoT system and improve the data reconstruction accuracy for the sink node. Then, the acknowledge mechanism for the data forwarding strategy is employed to improve the reliability of the data transmission among clusters. Therefore, reliable data handover is accomplished for the multi-path routing data among different nodes. The results which are about the simulation shows that the loss rate of the packet is equal to 40%, the data reconstruction error of the MIC-CSDG algorithm still remains lower than 5%. Compared with other existing algorithms, the data forwarding time is reduced by 16.36%, while the average network energy consumption is reduced by 23.59%. Therefore, the validity and efficiency of the proposed method are verified.

INDEX TERMS Internet of Things, mobile intelligent computing, data gathering, compressive sensing.

I. INTRODUCTION

The Internet of Things (IoT) had riched the exchange of information for the objects in the physical world [1]–[5]. According to a predefined protocol, the IoT is established by the inter-connection of the data sensing terminals, e.g., sensor networks, RFID, ultra-red identifier, and the GPS system, to perform the information exchange and communication [6]–[9]. The IoT is widely employed for the smart identification, locating, tracking and managing of the objects [10], [11]. Usually, the IoT is organized in three layers, i.e., the application layer, the transmission layer and the sensing layer. The sensing layer riches the seamless sensing of the physical objects to identify the objects, obtain the basic properties of the objects, and gather data. The transmission layer achieves the reliable data transmission between the sensing layer and the application layer, to further reduce the network delay and packet loss, as well as improve the data gathering speed and accuracy. The application layer achieves the smart processing, i.e., the gathered data is analyzed and

The associate editor coordinating the review of this manuscript and approving it for publication was Xuxun Liu.

processed in the application layer for the effective check of the data [12]–[15]. The key to the application of IoT is the software, including the application software at the server end and the intermediate, as well as the software for data mining and analysis, and the embedded software in the transmission layer and the sensing terminals. The intermediate is the core of the IoT software, while the sensor network is a crucial part for the sensing layer in IoT. Sensor networks normally cannot be directly linked to the IoT [16]–[19]. Instead, the data is forwarded to the IoT via the Sink node. Therefore, the data gathered by the sensor nodes needs to be forwarded by the intermediate network. During the forwarding, the intermediate network and the choice of the gateway node significantly affects the data forwarding efficiency. The amount of data gathered by the sensing layer in IoT is huge, while the users are only interested in a particular part of the total data. Therefore, the excessive data transmission not only wastes a large amount of energy, but also leads to network congestion.

Recently, the emerging compressive sensing (CS) technique serves as a brand new method for the acquisition and processing of the signals, providing new solutions for the processing of the data in wireless sensor networks (WSNs) [20]–[23]. Research shows the data which is CS-based gathering methods could fully exploit the spatial characteristics of the network, reduce the network energy consumption, simplify the data compression procedure, and avoid the energy hole problem [24]–[26]. However, the gathering method of the CS-based data in WSNs is focused mainly on the reliable network environment, where packet loss is negligible on the network links. The packet loss in realistic WSNs is quite common. In addition, it is shown by existing results that the performance of the gathering methods of the CS-based data is strongly affected by the loss of the packet. For example, with extremely loss rate of the packet which is low, existing gathering methods of the data cannot guarantee the effective recovery of the of the network data. Although some solutions have been proposed based on the sparse random projection to alleviate the impacts of the packet loss, no realistic design is provided for the sparse expression base and the generality of these solutions cannot be guaranteed.

In order to address the problems mentioned above, a Mobile Intelligent Computing Based on Compressive Sensing Date Gathering algorithm (MIC-CSDG) is proposed in this work. This method employs the mobile intelligent computing to derive the multi-hop function among sensor nodes, which is further utilized to determine the proportional relationship for the sensor nodes. Based on the sparse matrix, an observation matrix is designed with low correlation to mitigate the influences of the data packet loss on the entire IoT system and improve the data reconstruction accuracy for the sink node. Then, the acknowledge mechanism for the data forwarding strategy is employed to improve the reliability of the data transmission among clusters. Therefore, the reliable data handover is accomplished for the multi-path routing data among different nodes. We summarized the work's major contributions as follows.

(1) The MIC-CSDG algorithm is proposed, which divides the entire WNS into different regions, and conducts independent sampling and measurement within each region. Therefore, the overload issue is solved for the central region, while the direct transmission and compressive sensing are effectively combined to reduce the transmission energy consumption.

(2) Based on the sparse matrix, an observation matrix is designed with low correlation to mitigate the influences of the data packet loss on the entire IoT system and improve the data reconstruction accuracy for the sink node. Therefore, the data compression performance of the proposed MIC-CSDG algorithm can be guaranteed. In addition, the low correlation of the sparse routing projection matrix is also maintained, which ensures the effective reconstruction of the sensing data.

(3) Based on (1), the mobile intelligent computing is employed for the clustering of the dynamic region division. The maximal hop number required for the intra-cluster communication is derived. Then, according to the probabilistic forwarding protocol, the decision process is provided for the intra-cluster data transmission to increase the data transmission reliability and reduce the packet loss.

(4) It is shown by the simulations that compared with existing algorithms; the proposed MIC-CSDG algorithm could reduce the reconstruction error, network delay and the consumption which is of the energy, even prolong lifetime of the network. Therefore, the validity and efficiency of the method which is proposed is verified.

II. RELATED WORKS

In realistic application scenarios, monitoring is often required for some key regions in the sensing area, and higher requirements are imposed on the timeliness and accuracy of the data reconstruction algorithms. For solving these problems, the area of the WSNs is divided into different regions and independent sampling is conducted in each region, which could effectively solve the overload issue of the central region. Meanwhile, the direct transmission is combined with the CS technique to reduce the transmission energy consumption. Then, an observation matrix is designed based on a sparse matrix, to curb the influence of packet loss on the entire IoT system and further improve the data reconstruction accuracy at the Sink node. With the emerging of the distributed CS theory, more and more researchers have paid attention to the application of the CS technique in WSNs.

A. CS RECONSTRUCTION ACCURACY

It was pointed out in paper [27] that distributed CS could exhibit the advantages of stronger error tolerance, higher security level and adaptivity, which lays the foundations for the application of CS techniques in WSNs. The CS technique was combined with the spatial and temporal correlation in paper [28] in order that it could decrease the whole number of communication and even prolong the lifetime of the network.

A fusion scheme of the layered data was put forward in paper [29] based on the compression field, where the parent node in each layer conducts the collection, compression and fusion of the data from lower-level layers. The data is transmitted to higher layers in turn, and finally reaches the cluster head node at the top level. This scheme could greatly reduce the redundant data within the network, the total amount of transmitted data, as well as the energy consumption. The iterative recurrent CS algorithm was proposed based on grids in paper [30] to solve the locating problem of multiple target nodes. In this algorithm, multiple target nodes can be located simultaneously and the locating accuracy can also be improved. For the locating problem of mobile target nodes, a CS based locating algorithm was proposed in paper [31], where the movement laws and the sparse expression base are matched to transform the locating problem of mobile targets into a reconstruction problem of the sparse signal. The CS technique is further combined with network coding in paper [32] and a compressive network coding scheme is proposed based on the temporal and spatial correlation. In this scheme, the data is compressed from the spatial dimension and temporal dimension to reduce the data amount and prolong the network lifetime. Exploiting the CS technique and the spatial-temporal correlation of the data, Energy consumption optimized compressive sensing in group sensor networks (ECO-CS) was proposed in paper [33]. Compared with the conventional data interpolation algorithm, the proposed algorithm is more advantageous in recovering lost packets. As for the application of CS in data gathering, the research topics mainly focus on the gathering efficiency, the matching between the observation matrix and the data gathering routing, as well as reducing the amount of transmitted data. Based on the temporal-spatial relation of the gathered data, a scheme of the data gathering was proposed in paper [34] to minimize the energy consumption. A gathering scheme of a sequence CS based data was raised in paper [35] for the WSN, which could simultaneously reduce the network transmission cost and guarantee the data recovery quality.

B. DATA GATHER ALGORITHM BASED ON CS

Compressive data gathering based on even clustering (CDG) was provided in paper [36] that the sparse projection could also achieve the effective recovery of the compressed data, and the relation was derived between the observation number and the sparsity degree of the projection matrix. The sparse projection was exploited in paper [37] to achieve the highly accurate reconstruction of the data sensed by a single sensor. In addition, according to the sparse projection, the sparse expression base was designed to satisfy different requirements under different application scenarios. The Random Walk method is employed in paper [38] to realize the data gathering in WSNs under distributed environment, and a theoretical analysis was provided on relevant performances of the data recovery. A data model estimation algorithm was proposed based on random projection in paper [39], where only the values of a quantity which is small of sensors are required in each measurement and the communication cost can be greatly reduced. A data gathering algorithm based on compressive sensing in lossy wireless sensor networks (CS-RTSC), the data gathering in paper [40] is conducted, and the quantity of sensor nodes involved in the gathering for each individual value can be greatly reduced. Therefore, the communication cost can be effectively reduced in the data gathering process. The gathering algorithm of a CS-based data with optimized routing protocol was proposed in paper [41]. This algorithm compresses the link data to prolong the lifetime of the network. However, this algorithm robustness is poor and the network energy consumption is not guaranteed minimized. A sparse projection matrix based on the gathering algorithm which is of data was put forward in paper [42], where the cost-minimizing sparse projection matrix is employed to achieve the data sampling of the node data. In addition, this algorithm also accomplishes the accurate reconstruction of the gathered data. A CS based data gathering scheme with optimized tree-splitting routing was proposed in paper [43]. This scheme can be divided into two steps, i.e., the sparsifying process of the node data and the optimizing of the data gathering routing, which could guarantee the accuracy of the gathered data.

Therefore, the data compression performance and the low correlation with the sparse routing projection matrix can be guaranteed simultaneously for the proposed algorithm. Henceforth, the sensing data can be effectively reconstructed. Based on the statements mentioned above, the mobile intelligent computing is employed in this paper to derive the clustering network model for the dynamic region division. Also provided is the analysis for the maximal number of required hops to finish the intra-cluster communication. In order that it could reduce the reliability of the data transmission and reduce the packet loss, the ant colony algorithm is utilized for the retransmission, check, and multi-path transmission within the cluster. Finally, simulations are performed based on the reconstruction error, network energy consumption, network lifetime, and network delay. It is shown that algorithm which is proposed outperforms the other algorithms, which proves the validity and effectiveness of the MIC-CSDG algorithm.

III. PROBLEM DESCRIPTION

A. DATE GATHERING IN COMPRESSIVE SENSING

Compressive Sensing is an effective method for the signal acquisition and compression, which is already widely applied in the simultaneous data sampling and compression. Assuming that X is a signal with dimension N, while $\mathbf{\Phi}$ is a matrix of measurement with size $M \times N$ with $M < N$. The received signal is $Y = \mathbf{\Phi} \times X$. When X is a combination of liner of k base vectors ($k << N$), we can reconstruct X with known Y and $\mathbf{\Phi}$ with a large probability. So as to reconstruct the target signal with higher accuracy, the row number M of $\mathbf{\Phi}$ should satisfy $M \geq c \cdot k$ in which k is the sparsity degree of the target signal and c is the sampling ratio function. When CS techniques are employed, if the sparse expression base of X

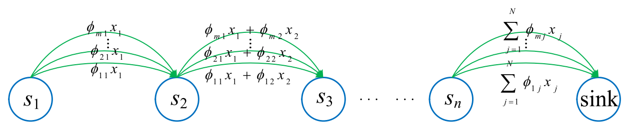


FIGURE 1. CS based data transmission process.

is Ψ, i.e., $X = \Psi \cdot \boldsymbol{\Theta}$, where $||\boldsymbol{\Theta}||_0 = k$ can be expressed as the 0-norm, then we denote the signal as k-sparse.

Assuming a WSN with N nodes, $s_1, s_2, s_3, \ldots, s_n$, and the sampling value of these nodes can be expressed as $x_1, x_2, x_3, \ldots, x_n$ at a certain time. In the gathering process of the CS based data, every node multiplies its sampling value by M weights and forwards the result to the node at the next hop. Finally, the Sink node receives M measurement values, which is a linear combination of the sensing data. Gathering process of the CS based data can be expressed as follows:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m1} & \phi_{m2} & \cdots & \phi_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

That is, $y = \boldsymbol{\Phi} \times x$, where $y = [y_1, y_2, \ldots, y_m]^{\mathrm{T}}$ is the value vector in received measurement at the Sink node and $x = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ is the sampling value at N nodes. $\boldsymbol{\Phi}$ is usually the Gaussian measurement matrix or the Bernoulli matrix where elements share the same distribution.

In order to further explain the CS based data gathering process in WSN with multi-hop transmission topology, the link structure for the data gathering process is illustrated as the following Fig. 1.

The sensor nodes set is $s_1, s_2, s_3, \ldots, s_n$, the generated sampling data is $x_1, x_2, \ldots, x_n$, and the data is transmitted to the node of sink in a multi-hop manner. If the conventional transmission mechanism is employed, then every node has to transmit the data passed from the downstream nodes and transmit data which are its own. Then the node of sink will get N uncoded data packets. The whole quantity of transmitted data packets in the network is $N(N+1)/2$, and the node energy consumption increases as the distance from the Sink node decreases. However, if the technique of CS is employed, the initial N data packets can be represented by M coded data packets, where $M << N$. Then the original data could be reconstructed by the Sink node with a large probability. The weights $\phi_i = \{\phi_{1\,i}, \phi_{2\,i}, \ldots, \phi_{mi}\}$ are imposed on each sampling value x_i, where ϕ_i is the element in the observation matrix $\boldsymbol{\Phi}_{M\times N}$. Therefore, the Sink node receives M data packets, and only $M \times N$ data packets are required for transmission.

In a dense matrix, each element in each row is non-zero. Therefore, the quantity of the data packets which is gathered in each measurement is $O(N)$, in which N is the quantity of sensor nodes which are in the network. The data fusion process, which is based on the matrix of the dense measurement, it is usually expressed in the form of a tree routing, where the network routing is the shortest spanning tree based on the cluster head [44]–[47]. Assuming that the amount of measurement is M, and then the amount of transmitted data packets for each node is M. The child nodes multiply the sensing data with their own weights and then transmit the result to the parent nodes. After the reception of the data from the child nodes, the parent node also multiplies its own data with the weights and performs the summation. Then the result is further transmitted to the nodes in the upper layer in the same manner, until the data from all the nodes reaches the cluster head.

According to the description above, in the conventional data gathering methods, the child nodes only needs to transmit their own data, e.g., the nodes far from the cluster heads. However, in the dense random data fusion methods, the child nodes need to transmit M data packets, which is larger than that of the conventional methods. Therefore, in Hybrid-CS based methods, only when the quantity of transmitted data packets is more than M is the CS operation adopted. The specific procedure is illustrated in Fig. 2(c), where the quantity on the transmission link represents the quantity of transmitted data packets for the corresponding node, while the number underlined on the links indicates that the corresponding nodes adopt the CS operation for the data fusion. The conventional data transmission method is shown in Fig. 2(a). It is shown in Fig. 2(a) that when the node is closer to the cluster head, it has to transmit more data packets. Therefore, the energy overhead for transmission is heavier, which could cause the imbalance of the network remaining energy. The dense random projection based method is illustrated in Fig. 2(b). It is shown that the number of the data packets which were transmitted is the same for each node, i.e., the consumption of the energy is balanced for different nodes. In addition, the total amount of transmission is smaller than that of the convention Non-CS based method. The Hybrid-CS based data fusion is shown in Fig. 2(c), where the node at the lower end of the fusion tree adopts the conventional transmission method while the CS method is employed only when the number of transmitted data packets exceeds the number of measurements. It can be observed in Fig. 2 that the method of the Hybrid-CS based data fusion could reduce the unnecessary resource waste for the nodes at the lower end, which greatly reduces the transmission amount as well as the communication load.

B. OTHER RECOMMENDATIONS

Assuming that, in the WSN area, we randomly deployed N nodes, and we denote the sensing data as $d = (d_1, d_2, \ldots, d_N)^{\mathrm{T}}$. If d is sparse under the sparse expression base $\Psi_{N\times N}$, then the measurement matrix is $\boldsymbol{\Phi} = (\phi)_{M\times N}$, and the observed vector is $\mathrm{Y}=(\mathrm{y}_i)_{M\times 1} = \boldsymbol{\Phi}\cdot\Psi^{\mathrm{T}}\cdot d$. According to (2) and (3), the Sink node could solve the optimization problem and the original data could be reconstructed with a certain accuracy degree.

$$\mathrm{Y} = \Phi \cdot \mathrm{S} = \Phi \cdot \Psi^T \cdot d = \Theta \cdot d \tag{2}$$

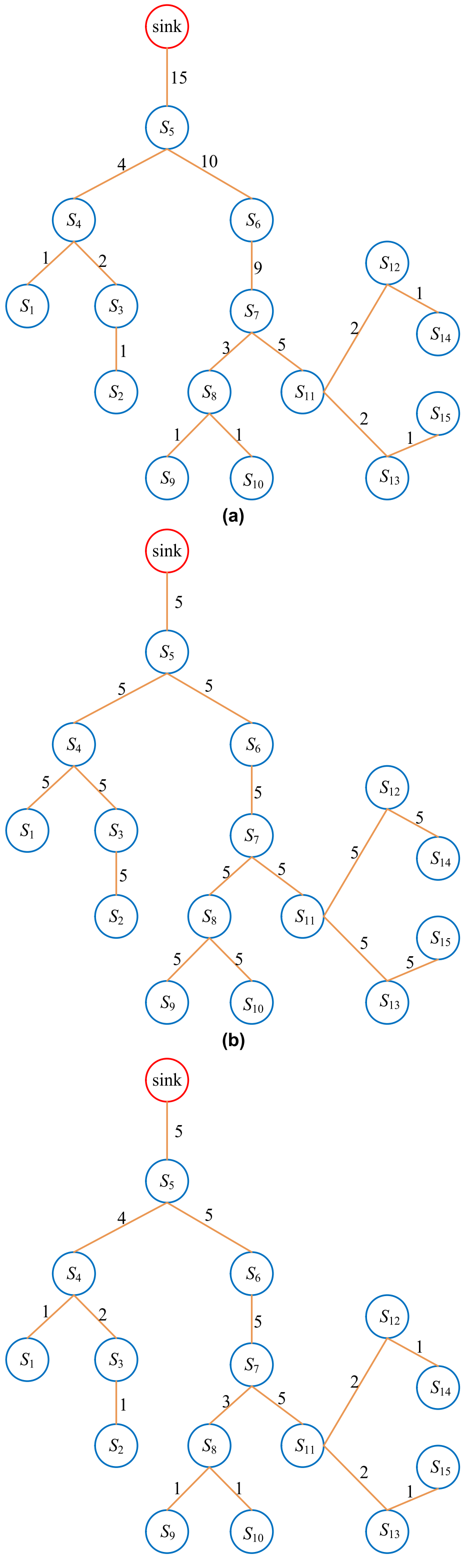


FIGURE 2. CS based on data gathering. (a) Non-CS which is based on data gathering. (b) Dense random projection based data gathering. (c) Hybrid CS based data gathering.

$$\vec{d} = \operatorname{argmin} \left\| \vec{d} \right\|_p \tag{3}$$

$$\delta_s(i,j) = \begin{cases} +1 & \textit{with prob.}\ (1-p)/2 \\ -1 & \textit{with prob.}\ (1-p)/2 \\ 0, & \textit{with prob.}\ p \end{cases} \tag{4}$$

where $\boldsymbol{\Theta} = \boldsymbol{\Phi} \cdot \Psi^{\mathrm{T}}$ is the matrix which is of sensing, and $\left\| \vec{d} \right\|_p$ is the l_p norm of the sensing data vector $\vec{d}$, as is shown as follows:

$$\left\| \vec{d} \right\|_p = \begin{cases} \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p} & 0 < p < +\infty \\ \max_{i=1,2\ldots,N} |x_i| \end{cases} \tag{5}$$

The scattered transformation base $\Psi_{N \times N}$ for the vector of data which is of sensing is obtained by the Discrete Fourier Transform (DFT), scattered transformation, and orthogonal sparse base, as shown in (6) and (7). The CS reconstruction accuracy is measured by the relative error η, and a smaller η indicates higher reconstruction accuracy.

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}kn} = \sum_{n=0}^{N-1} x(n) W_N^{kn} \tag{6}$$

$$\Psi_j(t) = 1 \Big/ \sqrt{N} \mathrm{e}^{i2\pi jt/N} \tag{7}$$

$$\eta = \left\| \hat{\mathbf{x}} - \mathbf{x} \right\|_2 / \left\| \mathbf{x} \right\|_2 = \sqrt{\sum_{n=0}^{N-1} (\hat{x}_n - x_n)^2} \Bigg/ \sqrt{\sum_{n=0}^{N-1} x_n^2} \tag{8}$$

Theorem 1: For matrix $\boldsymbol{\Phi}_s = (\xi_1, \xi_2, \ldots, \xi_M)^{\mathrm{T}}$, ξ_i is the discrete random sequence with independent and identical distribution. The random variable that constitutes the sequence follows the distribution of the Bernoulli. After that the matrix $\boldsymbol{\Phi}_s$ is full ranked with a probability approaching "1".

Proof: We can assume that with the conditions above, the matrix $\boldsymbol{\Phi}_s$ is not full rank. Therefore, for the i-th row of the matrix, a group of coefficients exist, satisfying the following equation:

$$x_i = a_1x_1 + a_2x_2 + \ldots + a_{i+1}x_{i+1} + \ldots . + a_Mx_M \tag{9}$$

where $a_1, a_2, a_3 \ldots, a_M$ are not all zeros. Make the process $\{X(n), n = 0, 1, \ldots, N\}$ which is random denote the row vector ξ_i, and then mean function and the variance function are:

$$EX\ (n) = [(+1)\ (1-p)\ /2] + [(-1)\ (1-p)\ /2] + 0 \times p = 0 \tag{10}$$

$$DX\ (n) = E\,[X\ (n) - EX\ (n)]^2 = E\,[X\ (n)]^2 = 1 - p \tag{11}$$

Denote a stochastic process $\{Y(n), n = 0, 1, \ldots, N\}$ as $a_1\xi_1 + a_2\xi_2 + \ldots + a_{i+1}\xi_{i+1} + \ldots + a_M\xi_M$, the mean and variance of this random variable are derived as follows:

$$EY\ (n) = E\left[\sum_{j \in [1,M], j \neq i} a_j \xi_j(n) \right] = \sum_{j \in [1,M], j \neq i} a_j E\xi_j(n) = 0 \tag{12}$$

$$DY(n) = E[Y(n) - EY(n)]^2 = E[Y(n)]^2$$
$$= \sum_{j\in[1,M],j\neq i} a_j^2 D\xi_j(n) = \sum_{j\in[1,M],j\neq i} a_j^2 (1-p) \quad (13)$$

Therefore, different stochastic processes are described by $X(n)$ and $Y(n)$. Because the discrete stochastic process $X(n)$, the value $X(i)$ of the corresponding variable $X(i)$ which is random takes from $x(i) \in \{+1, -1, 0\}$. Therefore, the state space length I_X is 3^N. Because the discrete stochastic process $Y(n)$, the value $Y(i)$ of the corresponding variable $Y(i)$ which is random takes from $-M+1 \leq y(i) \leq M-1, y(i) \in Z$. Therefore, the state space length I_Y is $(2M-1)^N$.

Assuming event A as the event which equation (12) holds, event of B as the event that the coefficients $a_1, a_2 \ldots, a_M$ are not whold zeros, event of C as the event that only one coefficient in $a_1, a_2, \ldots, a_M$ is non-zero, then we have:

$$P(A|B) < P(A|C) \quad (14)$$

Therefore, the derivation of $P(A|C)$ is transformed to addressing the probability that the independently identically distribution stochastic processes $X_1(n)$ and $X_2(n)$ have the same state. In the stochastic process state space $X(n)$, different states have respectively different probabilities. For the analytical convenience and without loss of any generality, we assume $p = 1/3$ in (5). Then,

$$P(A|B) < P(A|C) = (1/3)^N << 10^{-3} \quad (15)$$

Therefore, event A's probability is quite low, i.e., the assumption which is original did not hold and the matrix $\mathbf{\Phi}_s$ was full rank with a probability approaching 1.

IV. CLUSTER FORMATION AND OPTIMIZATION

A. DYNAMIC CLUSTERING

The unit operation time is taken as the basic number of rounds in WSNs. The dynamic clustering model is employed for a WSN which is with N sensor nodes in a certain operation period. N sensor nodes are separated into N_s clusters with the same cluster radius R_k, where $R_k = \lambda R_c$ and R_c is the radius, of the communication, of the sensing nodes and λ is the maximal number of communication hops within the cluster. N_s can be calculated as:

$$C_n = \delta\pi R_k^2 = \delta\pi\lambda^2 R_c^2 \quad (16)$$
$$N_s = \frac{N}{C_n} = \frac{N}{\delta\pi\lambda^2 R_c^2} \quad (17)$$

where C_n is the quantity of sensor nodes in the cluster, δ is the node density in the network. According to equations (16) and (17), it is shown that when the number of clusters N_s increases, the amount of communication within the cluster can be increased as well. The mathematical model for the communication gain is derived as follows:

$$f_{\max} = \omega_1 N_s + \omega_2 Q_c \quad (18)$$

where ω_1 and ω_2 are weighting factors and Q_c is the total number of communication hops in the cluster.

An appropriate scheduling of the clustering could reduce the size of the cluster and cluster number, equalize the number of hops which are among the sensor nodes, and reduce the network consumption of the communication energy. Theorem 2 is introduced based on the illustrations above.

Theorem2: The minimal amount of communication hops λ within a cluster should satisfy:

$$\lambda = \frac{\left(L + 9\omega_1^2\right)^{\frac{1}{3}} + \left(L - 9\omega_1^2\right)^{\frac{1}{3}}}{H^{\frac{2}{3}}} \quad (19)$$

where $H = \pi\delta\omega_2 R_c^2$, $L = 4\pi\delta\omega_1\omega_2 R_c^2$.

Proof: Within a cluster, the minimal number of hops for the communication among sensor nodes and the minimal amount of sensor nodes for the successful communication with the cluster head is:

$$C_\lambda = \delta\pi\lambda R_c^2 \quad (20)$$

For the minimal hop number β of a single cluster, the amount of upstream sensor nodes in that can be derived as:

$$C'_\beta = C_\beta - C_{\beta-1} = \delta\pi R_c^2 (2\beta - 1) a \quad (21)$$

The total number of communication hops in a single cluster is:

$$Q_{sc} = \sum_{\beta=1}^{\lambda} \beta C'_\beta = \sum_{\beta=1}^{\lambda} \pi\beta\delta R_c^2 (2\beta - 1) \quad (22)$$

where Q_{sc} is the number of total communication hops within a single cluster.

For the entire network, multiplying (17) with (22), we can get the total number of communication hops for the sensor nodes.

$$Q_c = N_s \times Q_{sc} = \frac{N}{\delta\pi\lambda^2 R_c^2} \times \sum_{\beta=1}^{\lambda} \pi\beta\delta R_c^2 (2\beta - 1) \quad (23)$$

Simplify (23) and we have:

$$Q_c = \frac{N}{\lambda^2}\left[\frac{2\lambda(\lambda+1)(2\lambda+1) - 3\lambda(\lambda+1)}{6}\right]$$
$$= N\left(\frac{2\lambda}{3} - \frac{1}{6\lambda} + \frac{1}{2}\right) \quad (24)$$

By substituting (17) and (24) into (18), we have:

$$f_{\max} = \omega_1 \frac{N}{\delta\pi\lambda^2 R_c^2} + \omega_2 N\left(\frac{2\lambda}{3} - \frac{1}{6\lambda} + \frac{1}{2}\right) \quad (25)$$

By calculating the partial derivative for (25), we have the approximate solution of λ:

$$\lambda = \frac{\left(4\pi\delta\omega_1\omega_2 R_c^2 + 9\omega_1^2\right)^{\frac{1}{3}} + \left(4\pi\delta\omega_1\omega_2 R_c^2 - 9\omega_1^2\right)^{\frac{1}{3}}}{\left(\pi\delta\omega_2 R_c^2\right)^{\frac{2}{3}}} \quad (26)$$

Therefore, the proof is completed.

B. PROBABILISTIC FORWARDING

The clustering structure is helpful for the transmission control for the data in large-scale WSNs, where a large amount of nodes which are of sensor are uniformly and randomly distributed in the target fields. After gathering of the data within the cluster, the sensor nodes transmit the data to the Sink node in a manner of multi-hop.

Definition 1: If the gradient of a sensor node is h, then this sensor node is termed h-hop node. The nodes that can directly transmit with the Sink node are termed the 1-hop node.

Definition 2: If the source node of a message is an h-hop node, then this message is termed the h-hop message.

Definition 3: If a h-hop node can directly communicate with a (h-1)-hop node, after that this (h-1)-hop node is termed the 1-hop downstream node of this h-hop node, meanwhile the h-hop node is termed the 1-hop upstream node of this (h-1)-hop node.

Definition 4: Unless s_i (h-1)-hop nodes are not the downstream nodes of the same h-hop node, these s_i nodes are termed homologous nodes.

Assuming that the required success probability for the message transmitted to the Sink node is p^*, while the forwarding probability is p for the sensor nodes, the relationship between p and p^* can be analyzed in the following four cases.

(1) The gradient of the source node is 1. In this case, the Sink is the 1-hop downstream node of this source node, i.e., the Sink node could transmitted with the source node directly and the data packets from the source node could be directly communicated to the Sink node. Therefore, no other node is required for the forwarding.

(2) The gradient of the source node is 2. In this case, the message is a 2-hop message and the source node cannot directly communicate with the Sink node, while the message needs to be forwarded via a 1-hop node to reach the Sink node. Assuming that there are K_1 1-hop downstream nodes for the source node and each 1-hop downstream node forwards this message with the same probability p_1, then the probability for the Sink node to receive the 2-hop message is:

$$P = 1 - (1 - p_1)^{K_1} \tag{27}$$

Apparently, this probability has to satisfy:

$$p^* \leq 1 - (1 - p_1)^{K_1} \leq 1 \tag{28}$$

Therefore, we have:

$$1 - \left(1 - p^*\right)^{\frac{1}{K_1}} \leq p_1 \leq 1 \tag{29}$$

Let $\alpha_1 = 1 - (1 - p_1)^{K_1}$ where α_1 denotes the probability that a 2-hop message could successfully reach the Sink node after the forwarding of a 1-hop node. Equation (29) describes the upper bound and the lower bound for this forwarding probability of the 1-hop node. Since the lower bound is strongly dependent on K_1, p_1 is dependent on K_1.

(3) The gradient of the source node is 3. In this case, the 3-hop message has to be forwarded twice, through a 1-hop node and a 2-hop node, to reach the Sink node. Assuming that there are K_2 1-hop downstream nodes for the source node, i.e., $s_1, s_2, \ldots, s_K$. In addition, each 1-hop downstream node forwards this message with the same probability p_2. These K_2 nodes have K_1 independent 1-hop nodes. The 1-hop downstream node j of the source node forwarding this message, the message will be received by node j and K_1 1-hop downstream nodes. According to the analysis in the second case above, this node will forward this message with probability p_1, and p_1 satisfies (29). Then, the probability that the message could be forwarded to the Sink node via node j is $\alpha_1 p_2$, i.e., after the forwarding of node j and the 1-hop node, the message can successfully reach the Sink node. The probability that the 3-hop message cannot reach the Sink node is:

$$P_1 = \prod_{j=1}^{K_2} \left(1 - \alpha_1^j p_2\right) \tag{30}$$

The probability that the message can successfully reach the Sink node is:

$$p = 1 - P_1 = \prod_{j=1}^{K_2} \left(1 - \alpha_1^j p_2\right) \tag{31}$$

Therefore, we could derive that the forwarding probability p_2 of the 2-hop node should satisfy:

$$\begin{cases} p^* \leq 1 - \prod\limits_{j=1}^{K_2} \left(1 - \alpha_1^j p_2\right) \leq 1 \\ \alpha_1^j = 1 - \left(1 - p_1^j\right)^{K_1^j} \end{cases} \tag{32}$$

Equation (32) indicates that the forwarding probability p_2 of the 2-hop node is dependent on p_1. This dependence is harmful for the calculation of p_2. We can calculate from formulas (29) and (30).

$$\alpha_1^j = 1 - \left(1 - p_1^j\right)^{K_1^j} \geq p^* \tag{33}$$

which holds for an arbitrary 2-hop node j. Substitute (33) into (32), we have

$$1 - \prod_{j=1}^{K_2} \left(1 - \alpha_1^j p_2\right) \geq 1 - \left(1 - pp^*\right)^{K_2} \tag{34}$$

Further simplify (34), we have the constraint condition for p_2.

$$\frac{1 - (1 - p^*)^{\frac{1}{K_2}}}{p^*} \leq p_2 \leq 1 \tag{35}$$

When p_2 satisfies (35), equation (32) will definitely hold. Therefore, the calculation of p_2 is no longer dependent on p_1.

(4) The gradient of the source node is $m + 1$. In this case, a $(m + 1)$-hop message should be forwarded by $K_m m$-hop nodes. Every m-hop node forwards the message with the same probability p_m. Similar to the analysis in the third case, we can

derive that the forwarding probability of the m-hop nodes should satisfy:

$$\frac{1-(1-p^*)^{\frac{1}{K_m}}}{p^*} \leq p_m \leq 1 \tag{36}$$

Based on the analysis above, in order to ensure that a $(m+1)$-hop message could eventually reach the Sink node with probability p^*, the forwarding probability of the$i$-hop node should satisfy the following conditions:

$$\begin{cases} 1-(1-p^*)^{\frac{1}{K_i}} \leq p_i \leq 1 & i=1 \\ \dfrac{1-(1-p^*)^{\frac{1}{K_i}}}{p^*} \leq p_m \leq 1 & i=2,3,\cdots,m \end{cases} \tag{37}$$

Equation (37) is only dependent on the number of homologous nodes and the expectation of the successful transmission probabilityp^*. In this way, the bound for the forwarding probability is:

$$p_i = \begin{cases} 1-(1-p^*)^{\frac{1}{K_i}} & i=1 \\ \dfrac{1-(1-p^*)^{\frac{1}{K_i}}}{p^*} & i=2,3,\cdots,m \end{cases} \tag{38}$$

According to (38), p_i decreases as K_i increases, which could be illustrated by the fact that more homologous nodes can lead to more message transmission paths and lower forwarding probability. In particular, if each forwarding node has only one homologous node, i.e., one transmission path to the Sink node, then according to (38), the last 1-hop node will forward with probability p^* while the remaining nodes will forward with probability 1. Apparently, this can be one way to satisfy the forwarding probability p^*.

C. ALGORITHM DESCRIPTION

Step1: Initialization of the WSN. Firstly all the nodes are activated by the Sink node to establish Minimum Spanning Tree (MST) routing and finish the networking of the node.

Step2: The Sink node broadcast the data packet which is the heartbeat to the network. Upon the reception of the packet of the heartbeat data, the nodes which are in the network transmit their heartbeat data packet in period T_1. This transmission is conducted along the network routing and in the conventional transmission-forwarding data gathering method.

Step3: Each node performs the real-time counting of the received data packets on their own reception links. The counting results are stored in the memory, and further serve as the information of the priori for the sliding window-aided packet loss prediction mechanism.

Step4: At the end of the time period T_1, the received sequence for each node is denoted as $\{X_i\}$. The Sink node firstly broadcasts the seed ξ which is random to the network. Upon the reception of the seed which is random, the node i combines ξ with its ID and generates (ξ, IDi). Then, a unique observation matrix $(\delta_{i1}, \delta_{i2}, \ldots, \delta_{iM})^{\mathrm{T}}$ is generated and stored in tis own memory.

Step5: According to the routing, every node multiplies the data d_i which is gathered with the observation coefficients δ_{ij}. The results are summed up and further transmitted to the Sink node. According to (13), the re-transmission scheme is selected to recover the lost packets. The maximal re-transmission number is set as *max_num*. Otherwise; the prediction is taken as the received the packet of data and communicated through the hop which is the next. Therefore, this round of CS-based observation and sampling is finished.

Step6: After one round of data gathering and according to M observation values, the Sink node constructs the observed vector $\mathbf{Y} = (y_1, y_2, \ldots, y_M)^{\mathrm{T}}$. The Sink node further reconstructs the observation matrix $(\delta_{ij})_{M\times N}$ according to the ξ which is a seed which is random and the ID of the network nodes. Based on the sparse expression base $\Psi_{N\times N}$, CS algorithms are utilized to reconstruct the sparse matrix $\mathbf{S}$. By calculating $d = \Psi\cdot\mathbf{S}$, the original signal vector d can be reconstructed.

V. PERFORMANCE EVALUATION

So as to verify the effectiveness and the validity of the proposed MIC-CSDG algorithm, we perform the simulations on Matlab and simulation results are illustrated in this section. Related configurations are descried as follows. The square monitoring area is 200m $\times$ 200m and $N = 550$ sensor nodes are deployed densely. The energy for every sensor node is 10J. A Sink node is deployed outside the monitoring area. Parameters for data gathering are: $\omega_1 \in [0.1, 0.4], \omega_2 \in [0.5, 0.9], \lambda \in [1, 10], R_c = 10\text{m}, t = 600\text{s}$.

The comparison on the reconstruction error with different parameters is illustrated in Fig. 3 to Fig. 5, for the proposed MIC-CSDG algorithm and the other three algorithms. According to Fig. 3, the improvement of the proposed MIC-CSDG algorithm is not very obvious with the network running time.

However, the other three algorithms exhibit obvious improvement as the network running time increases.When the network running time is 300s, the performances are significantly improved for the other three algorithms. By contrast, the performance of the proposed MIC-CSDG algorithm is almost steady and the variation is almost smaller than 3%. According to Fig. 4, when the network running time t∈[200s, 400s], the performances are improved for all of the four algorithms. In addition, these four algorithms show a balanced improvement rate. At the beginning of the running time of the network, the error performances of the other three algorithms are worse than the proposed MIC-CSDG algorithm. As the running time increases, the CS-RTSC algorithm, the CDG algorithm, and the ECO-CS algorithm exhibit faster improvement, while the performance of the MIC-CSDG algorithm remains steady. The parameters for Fig. 5 are $\omega_1 = 0.5$, $\omega_2 = 0.9$, $\omega_1 = 0.3$, $\omega_2 = 0.7$, $\lambda = 10$. Similarly, the performance of the proposed MIC-CSDG algorithm also

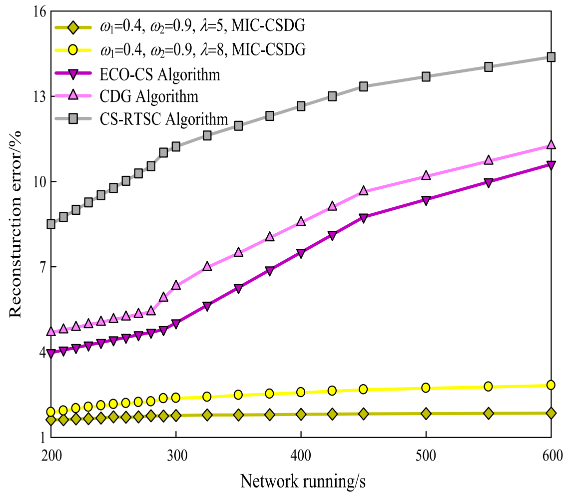


FIGURE 3. Comparison on reconstruction error with different parameters ($\omega_1 = 0.4$, $\omega_2 = 0.9$, $\lambda_1 = 5$, $\lambda_2 = 8$).

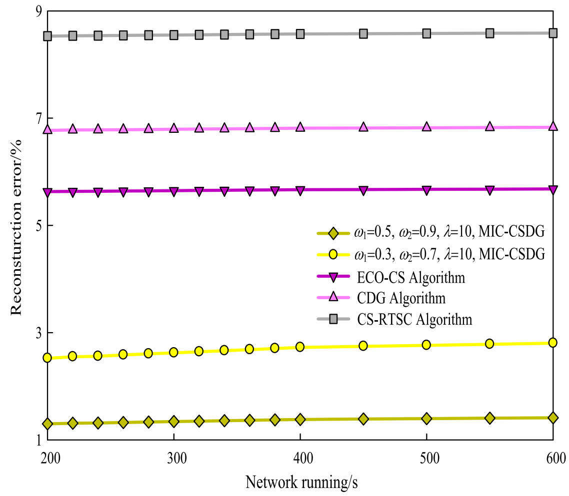


FIGURE 5. Comparison on reconstruction error with different parameters ($\omega_1 = 0.5$, $\omega_2 = 0.9$, $\omega_1 = 0.3$, $\omega_2 = 0.7$, $\lambda = 10$).

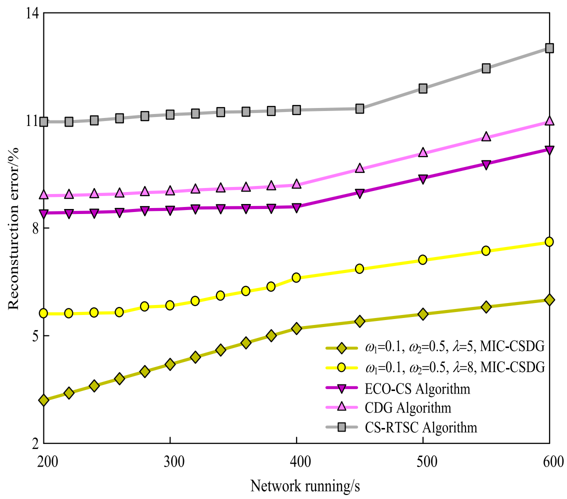


FIGURE 4. Comparison on reconstruction error with different parameters ($\omega_1 = 0.14$, $\omega_2 = 0.5$, $\lambda_1 = 5$, $\lambda_2 = 8$).

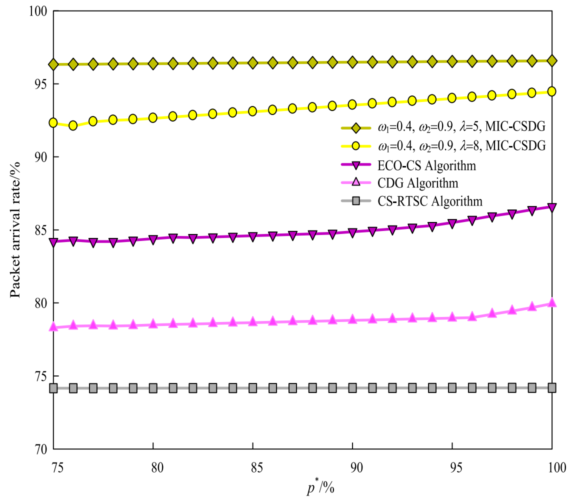


FIGURE 6. Comparison on pack arrrival rate with different parameters ($\omega_1 = 0.4$, $\omega_2 = 0.9$, $\lambda_1 = 5$, $\lambda_2 = 8$).

remains steady. By contrast, the other three algorithms shows worse performances and their performances are not obviously improved as the network running time increases. The main reason which is behind this result is that the algorithm of CS-RTSC adopts the extremely sparse observation matrix in order to reconstruct the data which is original. On the one hand, the number of samples is not sufficient. On the other hand, the loss of the packet will decrease the quantity of observations, which further causes deteriorated reconstruction accuracy due to insufficient observations. Although the CDG algorithm and the ECO-CS could alleviate the problem of misjudgment for the reception of the packet situation, the correlated packet loss which is in the gathering of the CS-based data would undermine the accuracy of the reconstruction seriously. By contrast, the number of observations is maximized in the proposed MIC-CSDG algorithm. In addition, the misjudgment problem at the Sink node is alleviated. The correlated packet loss problem is also addressed by the MIC-CSDG algorithm. In this way, the accuracy, of the reconstruction, of the CS-based data gathering processing is guaranteed from different perspectives. Furthermore, it is illustrated by the results of simulation that with the loss rate of a packet of 40%, the reconstruction error of the proposed MIC-CSDG algorithm could still remain lower than 5%. Therefore, the effective reconstruction in CS-based data gathering is guaranteed.

The comparisons on the packet arrival rate are illustrated in Fig. 6, Fig. 8, and Fig. 10 for different parameters. In addition, the comparisons on the average forwarding time are illustrated in Fig. 7, Fig. 9, and Fig. 11, respectively. 550 sensor nodes are assumed for the simulations, while each sensor node will transmit 55 messages. The proposed MIC-CSDG algorithm is compared with the CS-RTSC algorithm, the CDG algorithm, and the ECO-CS algorithm with

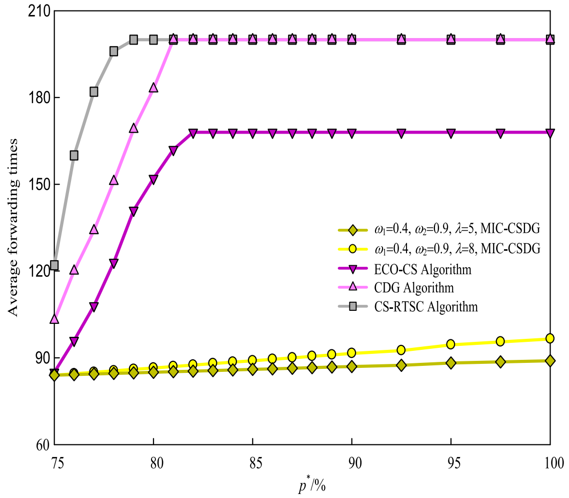


FIGURE 7. Comparison on the average forwarding times with different parameters ($\omega_1 = 0.4$, $\omega_2 = 0.9$, $\lambda_1 = 5$, $\lambda_2 = 8$).

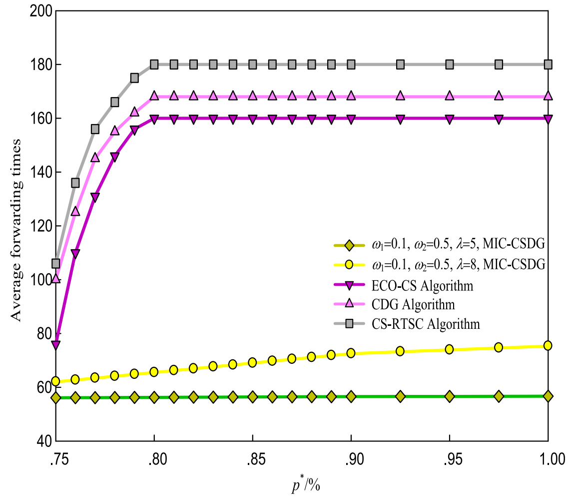


FIGURE 9. Comparison on the average forwarding times with different parameters ($\omega_1 = 0.1$, $\omega_2 = 0.5$, $\lambda_1 = 5$, $\lambda_2 = 8$).

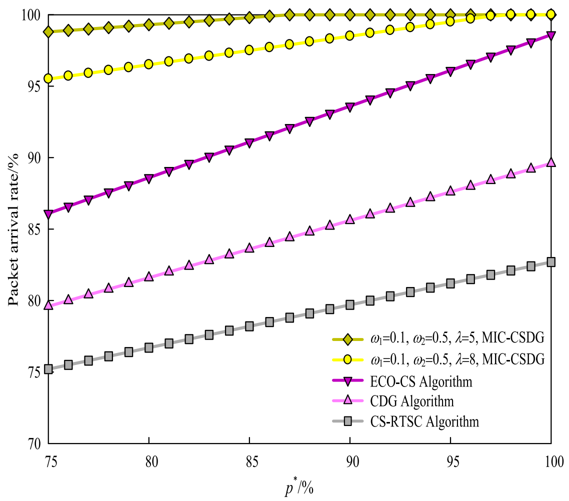


FIGURE 8. Comparison on pack arrrival rate with different parameters ($\omega_1 = 0.1$, $\omega_2 = 0.5$, $\lambda_1 = 5$, $\lambda_2 = 8$).

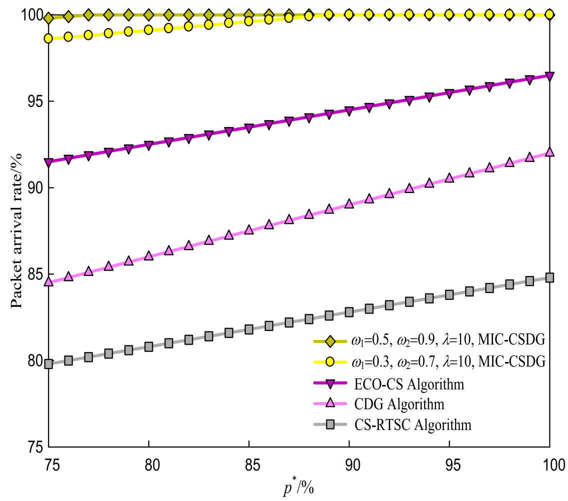


FIGURE 10. Comparison on pack arrrival rate with different parameters ($\omega_1 = 0.5$, $\omega_2 = 0.9$, $\omega_1 = 0.3$, $\omega_2 = 0.7$, $\lambda = 10$).

different p^*. The simulation results in Fig. 6, Fig. 8, and Fig. 10 are obtained by counting the received data packets at the Sink node. According to the results of the simulation, the proposed MIC-CSDG algorithm could satisfy the requirement for p^*. As p^* increases, the packet arrival rate could closely approach the theoretical upper bound for p^*, i.e., 1. According to Fig. 7, Fig. 9, and Fig. 11, the average forwarding time for each message of the proposed MIC-CSDG algorithm also increases. When p^* approaches 1, the average forwarding time of the MIC-CSDG algorithm still remains steady. Actually, this property guarantees that the proposed MIC-CSDG algorithm could improve the packet arrival rate with a small number forwarding and further save the energy consumption. By contrast, the other three algorithms require more message forwarding times and exhibit lower packet arrival rate. The reason is that for each algorithm, if the data sparsity increases, the data correlation will be weaker. Therefore, according to the CS theory, the message forwarding time will increase and the packet arrival rate will decline. However, when the data sparsity is 9, the proposed MIC-CSDG algorithm could outperform the CS-RTSC algorithm, the CDG algorithm, and the ECO-CS algorithm in terms of the packet arrival rate performance and the average forwarding time performance. But when the data sparsity is 17, the proposed MIC-CSDG algorithm is more advantageous due to weaker data correlation. The CS-RTSC fails to effectively evaluate the network node data which is based on a limited number of samples which are compressed. In addition, the CDG algorithm and the ECO-CS algorithm maximize the number of observations for the node data to improve the packet arrival rate and average forwarding time. However, these two algorithms fail to outperform the proposed MIC-CSDG algorithm. Over the unreliable links, the MIC-CSDG algorithm could alleviate the heavy dependence of the gathering

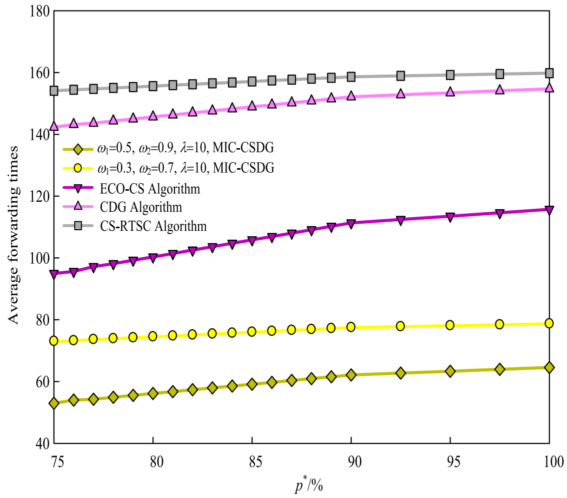


FIGURE 11. Comparison on the average forwarding times with different parameters ($\omega_1 = 0.5$, $\omega_2 = 0.9$, $\omega_1 = 0.3$, $\omega_2 = 0.7$, $\lambda = 10$).

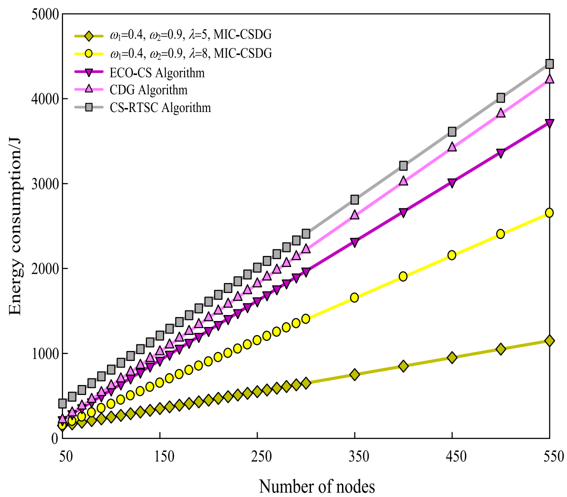


FIGURE 12. Comparison on networks energy consumption with different parameters ($\omega_1 = 0.4$, $\omega_2 = 0.9$, $\lambda_1 = 5$, $\lambda_2 = 8$).

algorithms of the CS-based data effectively on the data correlation. Therefore, the packet arrival rate performance and the message forwarding time performance could be improved for correlated data sets.

According to Fig. 12 and Fig. 13, as the amount of sensor nodes increases, the consumption of the energy also increases for the four algorithms. This is due to the fact that in order to obtain higher packet arrival rate, the number of sampling nodes has to be increased for all the algorithms to improve the network throughput. Accordingly, the energy consumption is increased. However, the proposed MIC-CSDG algorithm exhibits the lowest energy consumption among all the algorithms. This is due to the clustering network structure, which could balance the intra-cluster node energy. In addition, only nodes in the same cluster are able to communicate in the MIC-CSDG algorithm. Therefore, no other nodes are required to transmit with the Sink node and the network load can be further balanced.

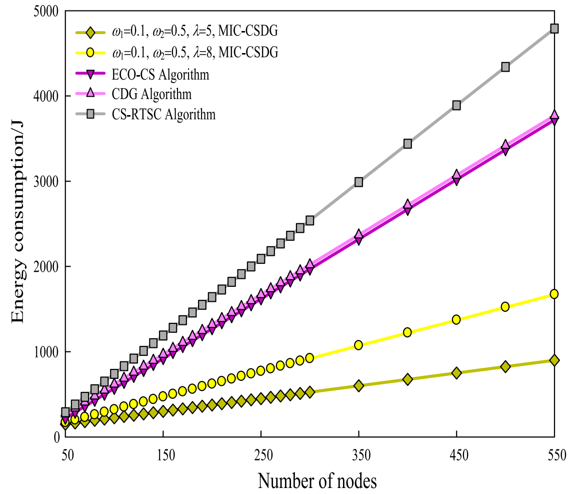


FIGURE 13. Comparison on networks energy consumption with different parameters ($\omega_1 = 0.1$, $\omega_2 = 0.5$, $\lambda_1 = 5$, $\lambda_2 = 8$).

VI. CONCLUSIONS

The MIC-CSDG algorithm was proposed in the paper in order to facilitate the energy-efficient sampling in WSNs. According to this algorithm, the CS-based data gathering process is independently conducted within each cluster, which could exploit the advantages of CS techniques and avoid excessively large data center in the data transmission process. To further verify the efficiency of the algorithm which is proposed, we perform simulations on the reconstruction, packet arrival rate, message forwarding time, and the consumption of the energy. It is shown that the algorithm which is proposed could achieve the accurate reconstruction of the signal with unknown sparsity. In addition, this algorithm exhibits strong global searching ability to improve the overall performance. Especially for unstable networks, the data reconstruction accuracy shows prominent improvement over existing data gathering algorithms. In addition, the network coverage optimization mechanism is introduced with the network clustering. Therefore, the amount of sampling data could be reduced for each cluster member node. Meanwhile, the nodes are only required to perform communication within the cluster, without communicating with the Sink node. Compared with the conventional data gathering algorithms in WSNs, the proposed MIC-CSDG algorithm could effectively decrease the energy consumption for the nodes in the cluster, curb the rapid energy exhaustion for the network, and therefore prolong the network lifetime.

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010. doi: 10.1016/j.comnet.2010.05.010.

[2] X. Liu and P. Zhang, "Data drainage: A novel load balancing strategy for wireless sensor networks," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 125–128, Jan. 2018. doi: 10.1109/LCOMM.2017.2751601.

[3] T. Wang, J. Zhou, X. Chen, G. Wang, A. Liu, and Y. Liu, "A three-layer privacy preserving cloud storage scheme based on computational intelligence in fog computing," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 1, pp. 3–12, Feb. 2018. doi: 10.1109/TETCI.2017.2764109.

[4] Z. Sun, H. Wang, B. Liu, C. Li, X. Pan, and Y. Nie, "CS-FCDA: A compressed sensing-based on fault-tolerant data aggregation in sensor networks," *Sensors*, vol. 18, no. 11, pp. 1–17, Nov. 2018. doi: 10.3390/s18113749.

[5] X. Liu, "Atypical hierarchical routing protocols for wireless sensor networks: A review," *IEEE Sensors J.*, vol. 15, no. 10, pp. 5372–5383, Oct. 2015. doi: 10.1109/JSEN.2015.2445796.

[6] M. A. Saleem, Z. Shijie, and A. Sharif, "Data transmission using IoT in vehicular ad-hoc networks in smart city congestion," *Mobile Netw. Appl.*, vol. 24, no. 1, pp. 248–258, Feb. 2019. doi: 10.1007/s11036-018-1205-x.

[7] T. Wang, M. Z. A. Bhuiyan, G. Wang, M. A. Rahman, J. Wu, and J. Cao, "Big data reduction for a smart city's critical infrastructural health monitoring," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 128–133, Mar. 2018. doi: 10.1109/MCOM.2018.1700303.

[8] A. Liu, M. Huang, M. Zhao, and T. Wang, "A smart high-speed backbone path construction approach for energy and delay optimization in WSNs," *IEEE Access*, vol. 6, pp. 13836–13854, 2018. doi: 10.1109/ACCESS.2018.2809556.

[9] M. Mangia, A. Marchioni, F. Pareschi, R. Rovatti, and G. Setti, "Administering quality-energy trade-off in IoT sensing applications by means of adapted compressed sensing," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 8, no. 4, pp. 895–907, Dec. 2018. doi: 10.1109/JETCAS.2018.2846884.

[10] X. Liu, "Routing protocols based on ant colony optimization in wireless sensor networks: A survey," *IEEE Access*, vol. 5, pp. 26303–26317, 2017. doi: 10.1109/ACCESS.2017.2769663.

[11] B. Huang, W. Liu, T. Wang, X. Li, H. Song, and A. Liu, "Deployment optimization of data centers in vehicular networks," *IEEE Access*, vol. 7, pp. 20644–20663, 2019. doi: 10.1109/ACCESS.2019.2897615.

[12] Z. Sun, G. Zhao, and X. Xing, "ENCP: A new energy-efficient nonlinear coverage control protocol in mobile sensor networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2018, pp. 1–15, Jan. 2018. doi: 10.1186/s13638-018-1023-7.

[13] X. Liu, T. Qiu, and T. Wang, "Load-balanced data dissemination for wireless sensor networks: A nature-inspired approach," *IEEE Internet Things J.*, to be published. doi: 10.1109/jiot.2019.2900763.

[14] T. Wang, J. Zeng, Y. Lai, Y. Cai, H. Tian, Y. Chen, and B. Wang, "Data collection from WSNs to the cloud based on mobile fog elements," *Future Gener. Comput. Syst.*, to be published. doi: 10.1016/j.future.2017.07.031.

[15] Y. Lai, X. Gao, M. Liao, J. Xie, Z. Lin, and H. Zhang, "Data gathering and offloading in delay tolerant mobile networks," *Wireless Netw.*, vol. 22, no. 3, pp. 959–973, Apr. 2016. doi: 10.1007/s11276-015-0995-z.

[16] X. Liu, "Node deployment based on extra path creation for wireless sensor networks on mountain roads," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2376–2379, Nov. 2017. doi: 10.1109/LCOMM.2017.2739727.

[17] Z. Sun, Y. Zhang, Y. Nie, W. Wei, J. Lloret, and H. Song, "CASMOC: A novel complex alliance strategy with multi-objective optimization of coverage in wireless sensor networks," *Wireless Netw.*, vol. 23, no. 4, pp. 1201–1222, May 2017. doi: 10.1007/s11276-016-1213-3.

[18] T. Wang, G. Zhang, A. Liu, B. M. Z. Alam, and Q. Jin, "A secure IoT service architecture with an efficient balance dynamics based on cloud and edge computing," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2870288.

[19] W. Zhang, L. Wei, W. Tian, L. Anfeng, Z. Zhiwen, S. Houbing, Z. Shaobo, "Adaption resizing communication buffer to maximize lifetime and reduce delay for WVSNs," *IEEE Access*, vol. 7, pp. 48266–48287, 2019. doi: 10.1109/ACCESS.2019.2909869.

[20] G. Azarnia, M. A. Tinati, and T. Y. Rezaii, "Cooperative and distributed algorithm for compressed sensing recovery in WSNs," *IET Signal Process.*, vol. 12, no. 3, pp. 346–357, May 2018. doi: 10.1049/iet-spr.2017.0093.

[21] W. Tian, G. Zhang, Z. A. Bhuiyan, A. Liu, W. Jia, and M. Xie, "A novel trust mechanism based on fog computing in sensor–cloud system," *Future Gener. Comput. Syst.* doi: 10.1016/j.future.2018.05.049.

[22] Y. Chen, Q. Zhao, X. Hu, and B. Hu, "Multi-resolution parallel magnetic resonance image reconstruction in mobile computing-based IoT," *IEEE Access*, vol. 7, pp. 15623–15633, 2019. doi: 10.1109/ACCESS.2019.2894694.

[23] Y. Liu, M. Dong, K. Ota, and A. Liu, "Activetrust: Secure and trustable routing in wireless sensor networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2013–2027, Sep. 2016. doi: 10.1109/TIFS.2016.2570740.

[24] M. Dong, K. Ota, and A. Liu, "RMER: Reliable and energy-efficient data collectoin for large-scale wireless sensor networks," *IEEE Internet Things J.*, vol. 3, no. 4, pp. 511–519, Aug. 2016. doi: 10.1109/JIOT.2016.2517405.

[25] Y. Liu, A. Liu, X. Liu, and X. Huang, "A statistical approach to participant selection in location-based social networks for offline event marketing," *Inf. Sci.*, vol. 480, pp. 90–108, Apr. 2019. doi: 10.1016/j.ins.2018.12.028.

[26] L. Zhang and Y.-C. Liang, "Average throughput analysis and optimization in cooperative IoT networks with short packet communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11549–11562, Dec. 2018. doi: 10.1109/TVT.2018.2871259.

[27] C. Zhang, O. Li, G. Liu, and M. Li, "A practical data-gathering algorithm for lossy wireless sensor networks employing distributed data storage and compressive sensing," *Sensors*, vol. 18, no. 10, pp. 1–24, Oct. 2018. doi: 10.3390/s18103221.

[28] T. Wang, J. Zhou, A. Liu, M. Z. A. Bhuiyan, G. Wang, and W. Jia, "Fog-based computing and storage offloading for data synchronization in IoT," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2875915.

[29] X. Liu, "A novel transmission range adjustment strategy for energy hole avoiding in wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 67, pp. 43–52, May 2016. doi: 10.1016/j.jnca.2016.02.018.

[30] Z. Sun, X. Xing, T. Wang, Z. Lv, and B. Yan, "An optimized clustering communication protocol based on intelligent computing in information-centric Internet of Things," *IEEE Access*, vol. 7, pp. 28238–28249, 2019. doi: 10.1109/ACCESS.2019.2896250.

[31] W. Tian, L. Yuzhu, J. Weijia, A. Muhammad, L. Anfeng, and X. Mande, "Coupling resource management based on fog computing in smart city systems," *J. Netw. Comput. Appl.*, vol. 135, pp. 11–19, Jun. 2019

[32] X. Xue, S. Xiao, and L. Quan, "Efficient measurement method for spatiotemporal compressive data gathering in wireless sensor networks," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 4, pp. 1618–1637, Apr. 2018. doi: 10.3837/tiis.2018.04.012.

[33] H. Yang and X. Wang, "ECO–CS: Energy consumption optimized compressive sensing in group sensor networks," *Comput. Netw.*, vol. 146, pp. 159–166, Dec. 2018. doi: 10.1016/j.comnet.2018.07.003.

[34] H. Zhe, Z. Xia, L. Ou, Z. Ce, and Z. Dalong, "Data gathering algorithm based on compressive sensing under lossy WSN," *Ruan Jian Xue Bao/J. Softw.*, vol. 28, no. 12, pp. 3257–3273, Dec. 2017.

[35] J. Fan, Y. Hu, T. H. Luan, and M. Dong, "DisLoc: A convex partitioning based approach for distributed 3-D localization in wireless sensor networks," *IEEE Sensors J.*, vol. 17, no. 24, pp. 8412–8423, Dec. 2017. doi: 10.1109/JSEN.2017.2763155.

[36] J. Qiao and X. Zhang, "Compressive data gathering based on even clustering for wireless sensor networks," *IEEE Access*, vol. 6, pp. 24391–24410, May 2018. doi: 10.1109/ACCESS.2018.2832626.

[37] C. Lv, Q. Wang, W. Yan, and J. Li, "A sparsity feedback-based data gathering algorithm for wireless sensor networks," *Comput. Netw.*, vol. 141, pp. 145–156, Aug. 2018. doi: 10.1016/j.comnet.2018.05.022.

[38] X. Liu, "A deployment strategy for multiple types of requirements in wireless sensor networks," *IEEE Trans. Cybern.*, vol. 45, no. 10, pp. 2364–2376, Oct. 2015. doi: 10.1109/TCYB.2015.2443062.

[39] Z. Sun, Y. Yun, H. Song, and H. Wang, "CCAJS: A novel connect coverage algorithm based on joint sensing model for wireless sensor networks," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 10, pp. 5014–5034, Oct. 2016. doi: 10.3837/tiis.2016.10.020.

[40] Z. Han, X. Zhang, D. Zhang, C. Zhang, and S. Ding, "A data gathering algorithm based on compressive sensing in lossy wireless sensor networks," in *Proc. 2nd Int. Conf. Frontiers Sensors Technol. (ICFST)*, Shenzhen, China, Apr. 2017, pp. 146–153. doi: 10.1109/ICFST.2017.8210492.

[41] Z. Sun, X. Xing, B. Yan, and Z. Lv, "CMTN-SP: A novel coverage-control algorithm for moving-target nodes based on sensing probability model in sensor networks," *Sensors*, vol. 19, no. 2, pp. 1–21, Jan. 2019. doi: 10.3390/s19020257.

[42] X. Liu, T. Qiu, X. Zhou, T. Wang, L. Yang, and V. Chang, "Latency-aware anchor-point deployment for disconnected sensor networks with mobile sinks," *IEEE Trans. Ind. Informat.*, to be published. doi: 10.1109/TII.2019.2916300.

[43] Y. Mostofi, "Compressive cooperative sensing and mapping in mobile networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 12, pp. 1769–1784, Dec. 2011. doi: 10.1109/TMC.2011.31.

[44] M. Mangia, F. Pareschi, R. Varma, R. Rovatti, J. Kovaćević, and G. Setti, "Rakeness-based compressed sensing of multiple-graph signals for IoT applications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 5, pp. 682–686, May 2018. doi: 10.1109/TCSII.2018.2821241.

[45] B. Xue, L. Zhang, W. Zhu, and Y. Yu, "A new sensor selection scheme for Bayesian learning based sparse signal recovery in WSNs," *J. Franklin Inst.*, vol. 355, no. 4, pp. 1798–1818, Mar. 2018. doi: 10.1016/j.jfranklin.2017.06.009.

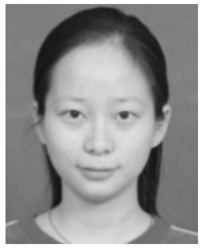
[46] G. Li, S. Peng, C. Wang, J. Niu, and Y. Yuan, "A energy-efficient data collection scheme using denoising autoencoder in wireless sensor networks," *Tsinghua Sci. Technol.*, vol. 24, no. 1, pp. 86–96, Feb. 2019. doi: 10.26599/TST.2018.9010002.

[47] Z. Guangxue, W. Tian, W. Guojun, L. Anfeng, and J. Weijia, "Detection of hidden data attacks combined fog computing and trust evaluation method in sensor-cloud system," *Concurrency Comput. Pract. Exper.*, to be published. doi: 10.1002/cpe.5109.

ZEYU SUN received the B.S. degree in computer science and technology from the Henan University of Science and Technology, in 2003, and the M.Sc. degree from Lanzhou University, in 2010, and the Ph.D. degree from Xi'an Jiaotong University, in 2017. He is currently an Assistant Professor with the School of Computer and Information Engineering, Luoyang Institute of Science and Technology, Luoyang, Henan, China. His research interests include wireless sensor networks, mobile computing, and the Internet of Things. He is a member of the China Computer Federation.

XIAOFEI XING received the B.S. degree in computer science and technology from the Henan University of Science and Technology, in 2003, and the Ph.D. degree from Central South University, China, in 2012. He is currently an Associate Professor with Guangzhou University and High-Level Talents of Guangzhou. He has been a Research Fellow with the University of Tsukuba, Japan, and a Postdoctoral Fellow in applied mathematics. His research interests include the Internet of Things, wireless sensor networks, and mobile computing.

BIN SONG received the B.S. and M.Sc. degrees in computer science from Zhengzhou University, in 2003 and 2007, respectively, and the Ph.D. degree from Korea University, in 2016. He is currently a Lecturer with the Henan University of Science and Technology. His research interests include the Internet of Things, wireless sensor networks, and mobile computing.



YALIN NIE received the B.S. degree in computer science and technology and the M.S. degree in computer application technology from the School of Computer and Communication, Hunan University, Changsha, Hunan, China, in 2004 and 2007, respectively, and the Ph.D. degree in computer application technology from the School of Computer Science and Technology, Xidian University, in 2015. Since 2007, she has been a Lecturer with the Luoyang Institute of Science and Technology. Her research interests include routing protocols, mobile computation, data aggregation, and wireless sensor networks.

HONGXIANG SHAO received the Ph.D. degree in communications and information system from the PLA University of Science and Technology. He is currently a Lecturer with the Luoyang Institute of Science and Technology. His research interests include resource allocation in small cell networks, cognitive radio networks, and optimization techniques.