IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal
SPECIAL SECTION ON COLLABORATION FOR INTERNET OF THINGS

Received November 12, 2018, accepted November 21, 2018, date of publication December 7, 2018, date of current version February 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2885534

Mobile Sink-Based Path Optimization Strategy in Wireless Sensor Networks Using Artificial Bee Colony Algorithm

YONG LU, NA SUN, AND XIUQIN PAN
School of Information Engineering, Minzu University of China, Beijing 100081, China
Corresponding author: Yong Lu (yonglu@muc.edu.cn)

ABSTRACT In traditional static wireless sensor networks (WSNs), the unbalanced communication overhead in different regions will result in premature death of some monitoring nodes. The introduction of mobile sink in WSNs can not only balance the node traffic load, but also obtain even energy consumption of nodes, thus effectively avoiding the ''hot spot'' problem and prolonging the network lifetime. However, the mobility of the sink will lead to frequent changes in the aspect of network topology, which can aggravate the overhead of the node's reorganization in hierarchical WSNs. Therefore, it is essential to obtain the optimal trajectory design of the mobile sink so as to improve the ability of data gathering. In this paper, a mobile sink-based path optimization strategy in WSNs using artificial bee colony algorithm is proposed. First, the problem of overall energy consumption in the network can be transformed into the minimization of the total hops between all subnodes and the rendezvous points of the mobile sink. The objective function and the constraint criterion should be established. Second, an improved artificial bee colony algorithm is proposed to solve the problem. On the one hand, the cumulative factor is introduced to the position update of the employed bee stage to speed up the convergence of the algorithm. On the other hand, the Cauchy mutation operator is presented to increase the diversity of the feasible solution and enhance the global search ability of the algorithm. The simulation results show that the proposed algorithm is better than the traditional methods in the aspects of energy efficiency and the real-time performance of data collection.

INDEX TERMS Wireless sensor networks, mobile sink, path optimization, artificial bee colony algorithm.

I. INTRODUCTION

Wireless sensor networks (WSNs) are considered as a type of large-scale and distributed network, which consists of massive low-cost and battery-constrained sensor nodes deployed in the environment [1]. In traditional static wireless sensor networks, the sensor nodes will perceive the environmental information periodically, and transmit the collected data to the base station or the Sink node [2]. It is worth pointing out that all nodes are stationary and single-hop or multi-hop data transmission mode is exploited for data delivery to the destination. It leads to a rapid death of nodes near sink due to excessive forwarding of data, and the emergence of ''funnel effect'' or ''hot spot'' problem may be arisen, which will generate energy holes and affect network lifetime [3]. Subsequently, the mobile wireless sensor network is introduced to alleviate the energy consumption of fixed nodes in data fusion process. In such case, the length of communication path between sensor nodes and the destination is greatly reduced, and Sink nodes can communicate directly with sensing nodes or reduce the communication hops. In addition, mobile wireless sensor networks can connect multiple subnets organically to avoid the formation of isolated area.

Originally, Ma and Yang [4] introduced a mobile data observer, i. e. a mobile robot or vehicle with inexhaustible power capacity, to operate as a mobile sink for data collection in wireless sensor networks, which triggered a boom of investigation in the fields of the mobile wireless sensor networks. Many factors should be taken into consideration in the system architecture of mobile converging wireless sensor networks, such as the number of nodes, the data collection mode of the forwarding hops, Mobile path planning and the data flow of the Perceptive nodes, and so on [5], [6]. Compared to static networks, mobile nodes acting as MDC (mobile data collector) can make the set of nodes at each point of residence change at any time. Specifically, the node set optimization and reasonable path planning can balance the load of nodes, thereby alleviating the problem of premature death of some nodes. In addition, the mobility of sink node can also reduce the number of hops for data forwarding so as to reduce the energy consumption effectively. However, due to the limitation of sink mobile speed and the long period of data collection, the real-time performance will be poor. Meanwhile, the path optimization problem of mobile nodes should be solved according to the specific scenario and node distribution. In this paper, we focus on the optimization of mobile sink's trajectory design to improve the ability of data gathering, and propose the concept of rendezvous points (RPs) to formulate traveling salesman problem.

The rest of this paper is organized as follows: In Section 2, we briefly introduce related work. We describe the assumptions and explain the details of our method in Section 3. In section 4, the detailed algorithm is described. At next section, the experiments method is shown and the result is discussed regarding the performance evaluation of our method. Finally, we conclude this paper and discuss the future work in Section 6.

II. RELATED WORKS

Route planning for sink nodes is a hot topic in mobile wireless sensor networks [8]. In general, the methods of generating mobile sink node can be divided into categories, such as, random path, predictable path and controllable path.

The random mobile strategy is that the direction and speed of sink nodes in the monitoring area are random. In this manner, the sensor node stores the sensing data, and then sends the collected data to the sink when the Sink node arrives. This strategy can solve the "hot spot" problem to a certain extent, because it can maximize the probability of balancing the energy consumption of nodes in the network.

Munari *et al.* [9] proposed a stochastic mobile algorithm for mobile Sink path prediction. All nodes can obtain their own location information, and the location based routing and forwarding strategy is applied in the communication between mobile Sink and nearby sensors. This random mobile mode can reduce the energy consumption of relay nodes and extend the network lifetime. In view of the problem of Periodic broadcasting in traditional data collection process, Guo *et al.* [10] presented a stochastic compressive data collection protocol to reduce the amount of non-effective data. Some nodes can be selected as collector with a certain probability, and other common nodes will send the collected data to collector through orientated determination. When mobile Sink moves randomly in adjacent region, the collector can transfer the data to the Sink node. To achieve the shortest routes for data delivery, Chen *et al.* [11] proposed a convergecast algorithm with Virtual Circle Combined Straight Routing. Considering that the routes between the sink and the sensors being reconstructed dynamically, a spanning tree is constructed to collect data periodically on the basis of some cluster heads located near the virtual backbone. This method can reduce the cost of the reconstruction link and increase the data transmission rate. However, excessive nodes participating in the routing will increase the network load undoubtedly. In [12], the concept of the agent node, which locates between the mobile Sink and the source node, is introduced to track the location of mobile Sink. By minimizing the update overhead of the path, the energy consumption of all nodes for tracking Sink nodes with random movement can be reduced. Nevertheless, the uncertainty caused by random mobility is inevitable, which result in some blind areas and the real-time performance of data transmission not be guaranteed. By using distributed Mobile trajectory selection method, Lee *et al.* [13] established a mixed linear programming model, which combines some factors, such as, initial location, routing, residence time. When the residual energy of neighbor nodes is greater than average level, the sink will interact with its neighbors to complete data collection. In [14], a novel swarm Intelligence-based sensor selection algorithm is presented to meets predefined quality of service (QoS) constraint with uncontrollable sink's mobility. In [15], a random geometric graphs (RGG) model is introduced to deal with spatial proximity for wireless sensor network, and conduct the random walk with inertia to traverse distant neighbor nodes towards reducing area overlap as well as accelerate the coverage time.

In fact, the mobile problem of mobile sink can be transformed into a path selection problem. Compared with the random strategy, the Sink in the fixed mobile strategy can move along the pre set trajectory in the monitoring area, which can be considered in accordance with the data load and communication overhead of the nodes in the region, so as to get better benefits. In [16], the route planning problem with path length constraint is discussed, and the heuristic scheme of multi-path planning is proposed to extend network's lifetime by redundant coverage of Perceived nodes. In [17], a delay constrained data collection method based on fixed trajectory is proposed. According to the predetermined time sequence, the mobile sink can access fixed position and data collection by multi-hop manner. However, this method is only suitable for a uniform distributed network, and the multi-hop manner may cause energy imbalance. Han and Jeong [18] presented the minimum Wiener Index spanning tree method for wireless sensor networks in case of multiple mobile sinks. Multiple mobile Sink is used to collect data to improve efficiency, and reduce energy consumption by using the shortest path. But the disadvantage is that the combination of multi hop transmission and fixed trajectory will make the node near path deplete its energy faster. From the point of view of energy saving, the tour planning for mobile data-gathering mechanism is proposed [19], in which All nodes send the collected data to the mobile Sink in single hop manner. However, in large-scale wireless sensor networks, mobile Sink needs to access a large number of traversing points, which will increase the time delay of data collection and make it unsuitable for real-time aware applications. In [20], an adaptive optimization model of stop times for low latency data collection is introduced, of which mobile Sink moves along the boundary of the divided square areas in the region for data collection. As the length of the square is less than the single hop communication radius of the node, the mobile path of Sink will pass through the communication range of all nodes in the network, so that all nodes can send the data to the mobile Sink by single hop communication. However, due to the long path length of mobile Sink, it can easily lead to larger data transmission delay. Therefore, the algorithm is not suitable for applications which are more sensitive to time delay.

As mentioned above, fixed movement can solve the problem of obstacle collision in random mobile strategy, and achieve the balance of all nodes' time delay for data collection. Compared with the random mobile strategy, the fixed trajectory method has a poor effect on alleviating the ''hot spot'' problem. This is due to the fact that once the trajectory is fixed no longer, it will cause heavy burden on the nodes on both sides of the trajectory, resulting in the uneven energy consumption and the premature death of some nodes. In [2], a rendezvous-based approach enabling energy-efficient sensory data collection with mobile sinks is proposed. According to the distance from the mobile sink's trail, the whole region is divided into two parts, and the selection of cluster heads is executed in different ways. In addition, the nodes with sufficient energy are chosen as Rendezvous sensor node to keep the communication with the mobile Sink for data forwarding. The collected data is transmitted to the corresponding Rendezvous sensor node by cluster head, and finally the aggregated result will be sent to the Sink node.

However, the fixed path strategy is obviously not flexible to meet the real-time requirement of sensor networks. Therefore, the controllable mobile strategy [21] is proposed. In this mode, the Sink node can determine the moving direction and location of the next step according to the real time situation of the network and the timely feedback from the sensor nodes, and adjust the moving path and speed of the node to ensure the path optimization. In 2011, Zhao and Yang [22] proposed a data gathering algorithm with load balanced clustering, which mainly contribute to extend the network lifetime and improve data collection delay. All nodes in the network are divided into three layers: the lowest level is the ordinary node layer, the cluster head nodes are distributed in the middle layer, and the SenCar data collector is situated at the top level. There must be at least one cluster head node in its one hop range to ensure the single hop transmission of data, and mobile Sink uses multi-hop MIMO (multiple-input and multiple-output) to select traversal points in each cluster for mobile data collection. Salarian *et al.* [8] presented a weighted rendezvous planning protocol for mobile wireless sensor networks. By defining the Rendezvous Point (RP) as the final traversal position of the Sink, the selection of RPs is based on the weight with the maximum cost to reduce the load of the relay node in the multi hop transmission. When RPs is determined, the traveling salesman algorithm is exploited to calculate the shortest path, which can traverse all RPs and achieve the purpose of reducing data transmission delay.

In order to ensure the relative equilibrium of the energy consumption among the nodes, an energy-aware sink relocation method is proposed [23], in which the sensor nodes are divided into hierarchies and the data direction of the relay nodes is determined according to the residual energy. In [24], an intelligent mobile data gathering scheme is presented to implement dynamical changes of data gathering tour of cluster's neighbor information table (NIT). And the mobile data collector decides to choose the optimal path to traverse between cluster heads based on the NIT information. In [25], the problem of mobile sink's data gathering is formulated as general random walks, and a Markovian random-walk movement strategy is proposed for mobile collectors to move over a graph.

During the process of data collection, the primary goal of using the controllable mobile strategy is to improve the data collection rate, to balance the network energy consumption and to ensure the real-time data transmission. In general, the controllable mobile strategy is relatively flexible in aspect of path planning, but also more complex and challenging. Therefore, how to design the mobile trajectory to improve the performance of the network is particularly important.

III. SYSTEM MODEL AND PROBLEM DESCRIPTION

A. MODEL ANALYSIS

For simplicity, we assumed that the sensor nodes are randomly deployed in the square area to form a self-organizing network topology to meet the following conditions: (1) all sensor nodes are stationary and have unique identities. The sink node is capable of mobility, and the direction and speed of its movement can be regulated according to the settings. In addition, the moving process remains at a constant speed; (2) all sensor nodes are provided with same communication radius r and initial energy E_0; (3) some nodes are selected as cluster heads, which will be traversed in turn by mobile sink node in each round; (4) When the mobile sink node moves to the cluster head's position, the member nodes in the communication range will send the collected data to the sink in multi hop manner with the optimal route; (5) The data collection process allows a certain time delay; (6) as for the clustering in all grids, the cluster heads are selected by the traditional method LEACH [26].

B. PROBLEM DESCRIPTION

As shown in Figure 1, the geographical range of a cluster marked by a dotted line, and the sink moves according to the planned path. The RPs indicate the temporary stay set of mobile sink for data gathering according to the optimized path, of which the sensor nodes within the communicate range can send the data directly to the sink. When the sink node does not move to the communication range, the RPs can put the data into the cache and then wait for data transmission. This can reduce the total delay of data transmission. As shown in Figure 1, the monitoring area is divided into grids of uniform size with side length $L(L < R_c)$, where R_c is the

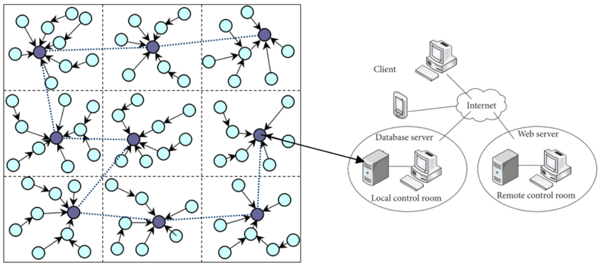

FIGURE 1. **Data collection process with mobile sink.**

communication radius for sensor nodes. Let VS represents the RPs set, and CH is the cluster heads set, and $|VS| = |CH| = m$. n sensor nodes are randomly distributed in the monitoring area, and the number of nodes in i-th grid is S_i. Accessing each sensor node will increase the path length of mobile Sink, which will lead to high time delay of data collection. Therefore, the rendezvous points set model is put forward to implement data gathering by mobile Sink's access a series to multi RPs. In order to reduce network energy consumption and data transmission delay, the establishment of RPs should take full consideration of the distribution of nodes and the number of hops to the sink.

In this paper, the energy consumption will be estimated by proposed model in [8]. For sending and receiving message, the energy consumption in each round by all nodes can be expressed as:

$$E_{total} = \sum_{i=1}^{n} (e_{tr}k_t^i + e_{rec}k_r^i) \tag{1}$$

where e_{tr} and e_{rec} represent the energy consumed by sending or receiving the data of per unit, and k_t^i and k_r^i indicate the amounts of data being received or transmitted at node i.

Suppose that the amount of data generated by sensor node during the process of monitoring per round is q, and the amount of data received and forwarded by node i from other nodes is k_r^i. Without considering the data fusion within, the amount of data that the node being forwarded can be expressed as: $k_t^i = k_r^i + q$. If the minimum hop count between node h_i and aggregate node is h_i, the energy consumption of data transmission and the hop count can be represented as the following relations:

$$\sum_{i=1}^{n} k_r^i = \sum_{i=1}^{n} h_i q \tag{2}$$

As a result, the overall energy consumption of the network can be given as:

$$\begin{aligned} E_{total} &= \sum_{i=1}^{n} (e_{tr}k_t^i + e_{rec}k_r^i) = \sum_{i=1}^{n} [e_{tr}(k_r^i + q) + e_{rec}k_r^i] \\ &= q[ne_{tr} + \sum_{i=1}^{n} (e_{tr} + e_{rec})h_i] \end{aligned} \tag{3}$$

From the above analysis, it can be deduced that the overall energy consumption of the whole network can reach minimum value as well as the minimization of the sum of hops of all nodes to RP. The number of hops is positively related to the distance from the node to the sink. Therefore, the path selection of the mobile sink will have an impact directly on the overall energy consumption of the network. The network energy consumption problem is equivalent to the Selection and path planning of RPs.

In the traditional way, the access to every sensor node in turn will increase the path length of the mobile Sink, and it will lead to great delay for data collection. Owing to the selection of the RPs, the sensor nodes can send the collected data in advance, and the mobile sink only needs to visit a series of RPs. It can improve the traverse efficiency of mobile sink in the premise of ensuring that static sensor nodes send data to destination under the condition of certain delay constraint.

Hence, the mathematical model of minimum energy consumption for data collection in mobile sink wireless sensor networks can be summarized as follows:

$$f = \min\{\sum_{i=1}^{m} H_i \times d_{TSP}\} \tag{4}$$

$$\begin{aligned} &s.t.\ dis(VS_i, CH_i) \le R_c, \quad i = 1, 2, \cdots, m; \\ &R_c \ge 0; \\ &\varepsilon t_i \le qS_i. \end{aligned} \tag{5}$$

where H_i represents the number of hops of gird i, and d_{TSP} is the path length of traversing the RPS in each grid by mobile sink. In addition, represents the data Collection Rate of the Mobile Sink, t_i represents the residence time of mobile sink during the rendezvous point VS_i.

Suppose that the number of hops $\{h_1, h_2, \cdots, h_n\}$ is a random variable and obeys Poisson distribution with the expectation μ and the standard deviation σ, they are independent and identically distributed. Consequently, the sum of hops from all child nodes to corresponding cluster heads, i.e. $\sum_{i=1}^{n} h_i$, which is approximately normal distribution. Assuming that H_{opt} is the sum of optimal hops, the probability that $\sum_{i=1}^{n} h_i$ is greater than H_{opt} can be derived from Lindburg-Levy central limit theorem:

$$\lim_{n\to\infty} P\{\frac{1}{\sigma\sqrt{n}}(\sum_{i=1}^{n} h_i - n\mu) \le \lambda\} = \Phi(\lambda) \tag{6}$$

where λ is any real number.

Using H_{sum} to represent the sum of all the sub nodes to their convergent points, we have

$$\begin{aligned} P\{H_{sum} > H_{opt}\} &= 1 - \prod_{i=1}^{n} P\{h_i \le H_{opt}\} \\ &= 1 - \Phi(\frac{H_{opt} - n\mu}{\sqrt{n\sigma^2}}) \end{aligned} \tag{7}$$

When $P\{H_{sum} > H_{opt}\}$ is smaller, the probability of H_{sum} less than H_{opt} is larger. At this time, the probability of the total hops to reach the optimal number is smaller. Conversely, when $P\{H_{sum} > H_{opt}\}$ is larger, the probability of total hops smaller than H_{opt} is low, which means the value of H_{sum} close to the optimal value. Therefore, if $\frac{H_{opt}-n\mu}{\sqrt{n\sigma^2}} = 0$, i.e. $H_{opt} = n\mu$, the total hops of the all sub nodes can obtain the minimum value. At this point, the number of hops between the child nodes to the RPs tends to the average path length in single cluster, and will achieve the minimum value of total path of traversing all RPs by mobile sink.

Therefore, the additional constraints can be expressed as

$$\forall i, \quad H_i \leq \frac{1}{n}\sum_{i=1}^{n} l_i, \quad 1 \leq i \leq m \tag{8}$$

IV. DATA GATHERING OPTIMIZATION

A. MODEL ANALYSIS

In the ABC algorithm, the bee searching for food can be regarded as the whole process of searching the global optimal solution [27]. The feasible solution can be expressed by the food source, and the quality of the feasible solution is determined by the quality of the food's source [28]. The artificial bee colony algorithm divides bees into three categories: scouts, onlookers, and employed bees. Classically, the bee colony initially contains only the employed bees and the onlookers, which have the same number of populations. Because the food source corresponds to the onlookers one by one, the number of the three groups, including the food source, onlookers and the employed bees, are equal as well. The ABC algorithm can be concluded into following steps. First of all, SN employed bees are randomly selected to search the whole feasible area to generate initial food sources. After that, the information of food source will be provided to all onlookers, and those bees were observed to determine the probability of every food source being selected based on the received information. Then, the food source is determined to search for by roulette method. Explicitly, the better the quality of food source is, the greater the probability of being selected. After the search process, the quality of the new food source and the current one can be compared, and the better food source will be reserved. Finally, the onlookers return to their nests, and the employed bees go back to search for their food sources near the optimal source's area. The above steps will iterate until the termination condition is satisfied, and the optimal solution will be obtained. Once employed bees or onlookers search for a certain number of times around a food source and Can't find a better new source of food, the employed bees or onlookers corresponding to the food source will be converted into a scout and looking for new food sources near the hive. After that, the mutated scouts will be turned into employed bees or onlookers to continue searching [29].

In the initialization stage, the specific parameters includes the number of food sources SN, the maximum cycle number MCN, and the failed times for continuous updating of food sources K. Let $X_i = (X_{i,1}, X_{i,2}, \cdots, X_{i,D})$ denote the location of the individual i in the population with D being the vector dimension of the optimization problem. The random generation of each individual in a population can be given as

$$X_{i,j} = X_j^{\min} + rand_{i,j}(X_j^{\max} - X_j^{\min}) \tag{9}$$

where i is the sequence number of the individual being initialized, j is the random value with the array of $\{1, 2, \cdots, D\}$. Besides, $X_j^{\min}$ and $X_j^{\max}$ are the upper and lower bounds of the entire search space, respectively. $rand_{i,j}$ represents a random number within the range of (0, 1).

After the initialization of the individuals being complete, their fitness should be determined as follows.

$$fit_i = \begin{cases} \dfrac{1}{1+f(X_i)}, & if\,(f(X_i) \geq 0) \\ 1 + abs(f(X_i)), & otherwise \end{cases} \tag{10}$$

where fit_i is the fitness value of the individual, and $f(X_i)$ is the function value of X_i relative to the optimization problem.

In this paper, the objective of the problem is to find the optimum value, namely, the smaller the function value is, and the better the fitness value will be. Since all employed bees have a unique individual and try to find a better individual around the individual, each new individual has only one dimension different from the original individual. Then, the population updating can be estimated by

$$V_{i,j} = X_{i,j} + \phi_{i,j}(X_{i,j} - X_{k,j}) \tag{11}$$

where $V_{i,j}$ represents the new value of the individual i with j-th dimension. k is a random value within the range of $\{1, 2, \cdots, SN\}$, and $k \neq i$. Besides, $\phi_{i,j}$ is a random real number between $[-1, 1]$.

Next, we need to determine the fitness of the new individual and compare with the original individual. If the new generated individual is provided with higher fitness, replace the original individual with the new on and record the counter $w(i)$ with the value of zero. Otherwise, the original individual will be reserved and execute $w(i) + 1$. After all the employed bees completed their search task, the individual related information will be shared by all of them and entered the onlookers stage. In the latter stage, the individual's probability of evolution will be estimated according to the fitness of the individual, and then the onlookers will choose the individuals based on Roulette method for further exploration. The individual's probability can be estimated by

$$p_i = fit_i / \sum_{i=1}^{SN} fit_i \tag{12}$$

It can be deduced that the higher the fitness of an individual, the greater the probability of being selected will be. Each scout will fly to the individual being chosen and use the Eq. (9) to generate a new individual. Similarly to the stage of employed bees, the fitness of V_i and X_i will be compared. If $V_i > X_i$, replace the original individual with the new on and rewrite the counter $w(i)$ with 0. Otherwise, the original individual can be reserved and execute $w(i) + 1$. When the search of the scouts is completed, the optimal individual can be selected from the counter set. If $w(i) > MCN$, the corresponding employed bee or the onlooker will be converted into scout. Thus, the new individual will be produced, and $w(i)$ is reset to 0. After that, the scouts turn into the role of employed bee or onlookers.

B. OPTIMIZATION STRATEGY

Due to the random functions being applied in location's update of bee groups, the basic artificial bee colony algorithm is easy to fall into local optimum and the convergence speed is limited [30]. For the optimization of nonlinear functions, the convergence rate of evolutionary strategies conducts slowly especially in solving some high-dimensional optimization problems. n order to improve the convergence speed, the specific methods include promoting the capability of mutation, and improving the efficiency of recombination or selection. Therefore, the speed of convergence of evolutionary strategies is closely related to variation.

In most practical applications, the probability distribution of physical quantities is either Gauss distribution or approximate Gauss's. Furthermore, the stochastic phenomena depicted by Gauss random variables are more common and conform to the process of cognition in human society. According to [31], Gauss mutation has strong local search ability and can maintain diversity of populations. When searching for food sources, the quality of food sources is proportional to the follow probability of the employed bees. On this account, the Gauss distribution [32] can be discussed in the location update stage of the algorithm. Therefore, the Gauss mutation operator is introduced in the employment bee stage to increase the convergence speed of the algorithm.

According to the description of Gauss distribution in the principle of probability, if the probability density function of random variable x is given as

$$\Pr(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{13}$$

where μ and σ denotes constants and $\sigma > 0$, then x obeys the Gauss distribution, and $x \sim N(\mu, \sigma^2)$.

If the food search by employed bees can be regarded as an event, using d_i to show whether there is a better event than the previous search result, the function $\Pr(d_i)$ can be defined as follows to represent the cumulative factor about the search result.

$$\Pr(d_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(d_i-\mu)^2}{2\sigma^2}) \tag{14}$$

Then, the cumulative factor is introduced into the position updating of employed bee as follows.

$$V_{i,j} = X_{i,j} \Pr(d_i) + \phi_{i,j}(X_{i,j} - X_{k,j}) \tag{15}$$

In this way, the convergence speed of the algorithm can be accelerated.

In addition, in order to increase the diversity of the feasible solution and enhance the global searching ability of the algorithm, it is necessary to tune the mutation operator properly. In general, the Cauchy distribution is easy to produce a random number away from the original point [33]. Generally, the peak value of Cauchy distribution is relatively small at the origin, but the distribution is longer at both ends. Therefore, Cauchy mutation can generate much more disturbance near the mutated individuals and result in making the variation range much wider [34]. If the Cauchy mutation is used to replace the mutation of the original evolutionary strategy to produce offspring, it means that the new solution of the mutation is likely to quickly jump from the local minimum area. To increase the diversity of the feasible solution, the Cauchy mutation operator is adopted in the phase of scout to avoid the local optimal solution of the algorithm.

The density function of Cauchy distribution is defined as:

$$C(x) = \frac{1}{\pi} \times \frac{\gamma}{\gamma^2 + x^2}, \quad -\infty < x < +\infty \tag{16}$$

If the variety of the optimized parameter exceeds the boundary of the search space, the parameter is equal to the boundary value. Here, the probability percent is introduced to describe the density function parameter γ of the Cauchy distribution quantificationally.

$$CQ(x) = \int_{x_j^{\min}}^{x_j^{\max}} \frac{1}{\pi} \frac{\gamma}{\gamma^2 + x^2} dx = \frac{1}{\pi} \arctan \frac{x}{\gamma} \Big|_{x_j^{\min}}^{x_j^{\max}} \tag{17}$$

where $X_j^{\min}$ and $X_j^{\max}$ represents the upper and lower bounds of the entire search space, respectively.

Therefore, the probability density function parameter of Cauchy distribution can be calculated as:

$$\gamma = \frac{x_j^{\max} - x_j^{\min}}{2\tan[\frac{\pi CQ(x)}{2}]} \tag{18}$$

As the global optimal fitness function value is in the latest N_0 iteration, if the absolute value of the change is less than the threshold value, then the mutation operations can be performed on the global optimal position in terms of the density function of Cauchy distribution probability.

$$RandCQ(x) = \int_{-\infty}^{X_i} \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - X_j^{gbest})} dx \tag{19}$$

where X_j^{gbest} denotes the global optimal solution found at present.

The location of the food source has not been updated after Maximum cycles, and it means that the location of the food source has fallen into local optimum. At this point, the onlookers should abandon the food source and turn into scout, and randomly generate a new food source position instead of the original one. Hence, the random number generated by the above Cauchy distribution function can be introduced into the generation of food source's location.

$$X_{i,j} = x_j^{\min} RandCQ(x) + (X_j^{\max} - X_j^{\min}) \tag{20}$$

Algorithm 1 Local Information Exchange Mechanism

Input: Population number SN; evaluation number K; Maximum cycle number MCN;
Output: The optimal individual in the current population.

for $i = 1$ to SN
    broadcast for adjacent nodes and place them into neighbor nodes set NS
    if $NS \neq \emptyset$
        for $k = 1$ to $|NS|$
            determine the best individual in the visual range
            $V_{i,j} = X_{i,j} \Pr(d_i) + \phi_{i,j}(X_{i,j} - X_{k,j})$
        end for
    else
        produce new individual V_i;
    end if
    evaluate the fitness value V_i;
    if $fit(V_i) < fit(X_i)$
        replace X_i with V_i, and set $w(i) = 0$;
    else
        $w(i) + +$;
    end if
    $K + +$;
    if $K = MCN$
        record the optimal solution so far
        jump out of the outermost while loop to end algorithm
    end if
end for
return $L(SN, K)$

Algorithm 2 Pseudo-Code of Proposed MSPO-ABC Algorithm

Input: Population number SN; Maximum cycle number MCN, Dimension of Vectors D, Lower bound and Upper bound of each element;
Output: The optimal individuals.

initialize population;
for $i = 1$ to SN
    evaluate the aggregate fitness function
end for;
while $K < MCN$
    for each employed bee
        obtain new solution using Eq. (15)
        evaluate the fitness value of new solution
        if $fit(X_i) < fit(V_i)$
            $X_i = V_i$
        end if
        $L(SN, K)$
    end for
    evaluate $RandCQ(x)$ for solution X_i using Eq. (19)
    for each onlooker bee
        Obtain a solution X_i based on Eq. (20)
    end for
    produce new solution V_i using Eq. (15)
    evaluate fitness value of new solution
    if $fit(X_i) < fit(V_i)$
        $X_i = V_i$
    else
        $K + +$
    end if
record the optimal solution so far
end while

The Cauchy mutation detection strategy based on the current optimal solution can effectively avoid the randomness of the solution, speed up the convergence speed, and improve the accuracy of the solution. Briefly, it can not only enhance the global searching ability of the algorithm, but also maintain the diversity of the population.

C. ILLUSTRATION OF PROPOSED ALGORITHM

Since the artificial bee colony optimization only changes on dimension at each time and the lack of information exchange between bees of the same kind, there is only one dimension between the new individual and the original individual, which is equivalent to a limited search in the vicinity of the original individual to search for individuals with better quality near the nest. After completing the search in turn, the employed bees provide the scouts with relevant information, including nectar content, the distance between the food source and hive, difficulty of mining, and so on. Then, the scouts can determine the probability of each individual being selected according to the above information, and selects the appropriate individual for further search.

Therefore, we implemented a local information exchange mechanism to achieve the exchange between the employed bees, which can be applied in both the employed bees and onlookers. By this means, it can not only enhance the communication between the same kinds of bees, but also guide the search by using high quality individual information to improve the search ability of the ABC algorithm. The details of local information exchange mechanism are presented in Algorithm 1.

The details of MSPO-ABC are presented in Algorithm 2. In algorithm MSPO-ABC, the problem of overall energy consumption in the network can be transformed into the minimization of the total hops between all sub nodes and the rendezvous points of mobile sink. The objective function and the Constraint criterion should be established. And then, the cumulative factor is introduced to the position update of the employed bee stage to speed up the convergence of the algorithm. Furthermore, the Cauchy mutation operator is presented to increase the diversity of the feasible solution and enhance the global search ability of the algorithm.

TABLE 1. Simulation parameters.

Parameter	Value
Network size	200 m × 200 m
Node deployment	Random
Number of nodes n	100~400
Side length L	50m
Number of grids m	16
Communication radius R_c	60 m
Maximum number of hops R	1~4
Moving Speed of the Mobile Sink	5~20 m/s
Initial energy	0.5 J
Residual energy threshold	0.1 J
Number of iterations N_0	500
q	1000 bits
e_{tr}	50 nJ/bit
e_{rec}	10 nJ/bit
ε	100 kbps

V. RESULTS AND ANALYSIS

To illustrate the effectiveness and performance of the proposed MSPO-ABC algorithm, we test and compare its performance with a number of competing clustering design protocols, namely, Rendezvous-based Data Collection algorithm (RDCA) [35] and Indegree-based Path Design for Mobile Sink algorithm (IPDMS) [36]. The parameter values for these approaches are selected in accordance with the values specified in [15] and [31], and the following simulation parameters are used in Table 1.

First, the convergence of the algorithm is verified. According to the optimization procedures, the cumulative factor is introduced to the position update of the employed bee stage to speed up the convergence of the algorithm. Moreover, the Cauchy mutation operator is presented to increase the diversity of the feasible solution and enhance the global search ability of the algorithm. Figure 2 demonstrates the relationship between the total path length and iteration times of MSPO-ABC and the traditional ABC algorithm. It can be seen from the results that MSPO-ABC algorithm has obvious improvement in aspect the accuracy and convergence speed compared with the traditional ABC algorithm.

Figure 3 shows the comparison of lifetime with different node's density and the threshold of maximum hops R. The FL is defined as the time elapsed until the first node in the network depletes its energy. It can be seen from the experimental result that FL decreases with the increase of node's density. In addition, when R is too large or too small, the appearance of first died node is relatively earlier. That is because the cluster heads as relay nodes will consume more energy, which increases significantly with the increase of node density. However, when R is small, the nodes far from the RP have to consume more energy for data delivery through single hop mode. Moreover, the hops between the RPs and the member nodes on the transmission path will increase as well as the value of R, and the transmission cost will be increased and result in more communication energy consumption. When the residual energy of nodes is insufficient, more hops will be used to send the data collected to RPs.

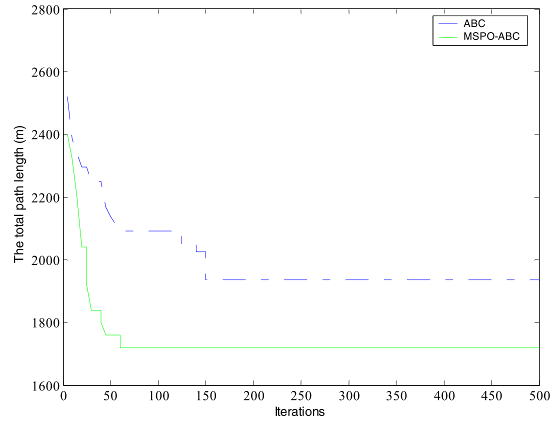

FIGURE 2. **The relationship between the total path length and iteration times.**

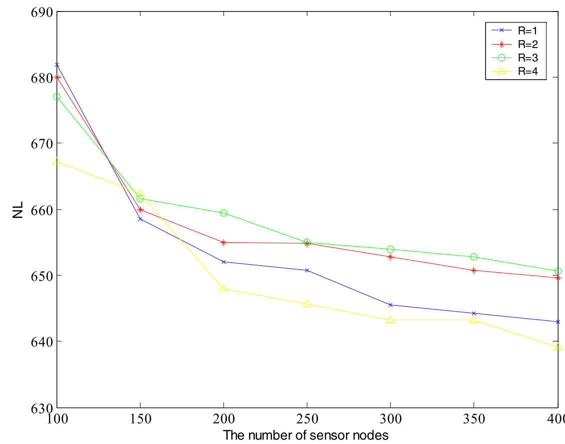

FIGURE 3. **The first node died in different node's density.**

Figure 4 shows the time required to traverse all grids by mobile sink per round. When the value of R is large, the length of Sink mobile path will be shortened. This is mainly reflected in the fact that the number of traverses on the horizontal side is less, and its required time is reduced. In addition, as shown in the results, the increase of Sink's moving speed has an obvious influence on the length of the single round data collection cycle.

Figure 5 illustrates the data collection efficiency. The data collection efficiency is defined as the ratio of the data collected by mobile Sink to the total data generated by all sensor nodes per round. From the results, when the R takes a larger value, the time of Sink traversing the whole network will be shortened, and the data collected by the node can be completely collected by the mobile Sink. Otherwise, when R is small, the node's capacity is limited and it may lead to buffer overflow. Thus, the data collection efficiency can't be guaranteed at a higher level.

Figure 6 shows the mean square deviation of the energy consumption of the nodes with the running time under

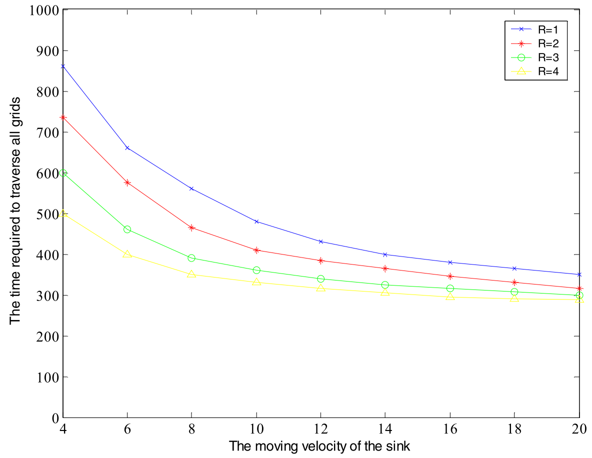

FIGURE 4. The time required to traverse all grids.

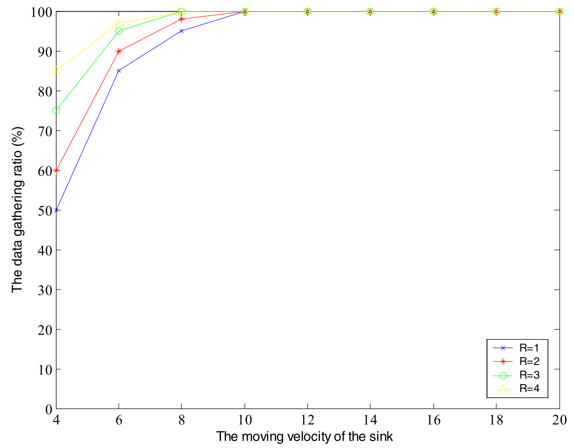

FIGURE 5. The data collection efficiency.

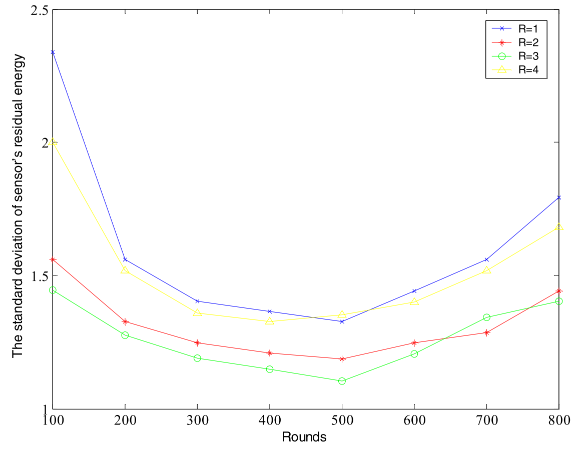

FIGURE 6. The standard deviation of sensor's residual energy.

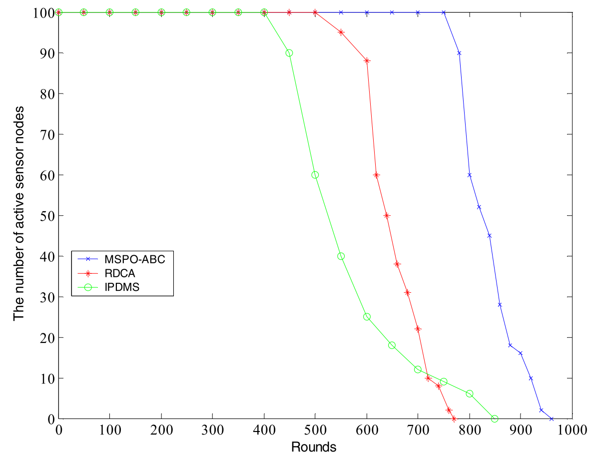

FIGURE 7. The number of active sensor nodes.

different node's densities. In all cases, the average variance of energy consumption in all cases is less than 0.3, and it demonstrates that our algorithm can achieve better energy balance. Comparatively, in the case of $R = 2$ or 3, the proposed schemes result into more balanced energy consumption. The reason is that the high value of R means the excessive communication cost of CHs. In addition, the nodes at the edge of the grid area will consume more energy in single hop mode. Thus, it will have a certain impact on the balance of energy consumption.

Next, we compared our algorithm with RDCA and IPDMS in aspect of network lifetime, the amount of data collection and time delay. Figure 7 shows the comparison of active nodes in different rounds. From the results, the round of first died node in MSPO-ABC algorithm is obviously later than the other algorithms. That is due to employ the grid area as the basic data collection unit of mobile Sink. In every round of data collection process, mobile Sink only needs to interact with less cluster heads, thus reducing the energy consumption of the whole network. With the operation of the network, the cluster head will be rotated, and the nodes with high residual energy and less communication cost in the grid will be possible to gain opportunities. Therefore, it will be further beneficial to the energy balance of nodes. Because of the limitation of cluster head selection in IPDMS, it is likely to cause cluster heads to be distributed in a certain area non-uniformly. Then, that is not conducive to reduce the communication cost of the nodes at the edge of the area or far from the cluster heads. In RDCA, the Sink's mobile trajectory is too fixed to fit for the changing of network topology. Then, the performance of energy efficiency is weaker than MSPO-ABC.

Figure 8 compares the amount of collected data by mobile sink between different algorithms. In this experiment, the rate of data collection of each node is 5bps. It can be seen from the results that the amount of data in RDCA and IPDMS increases slowly in the early stage of the network. Comparatively, MSPO-ABC shows a steady and rapid growth in aspect of the amount of collected data by mobile sink, which fully reflects the effect of mobile Sink on data collection efficiency based on the optimal path. In addition, the growth rate of data collection of RDCA and IPDMS began to converge earlier. This is due to the increasing number of death nodes appearing

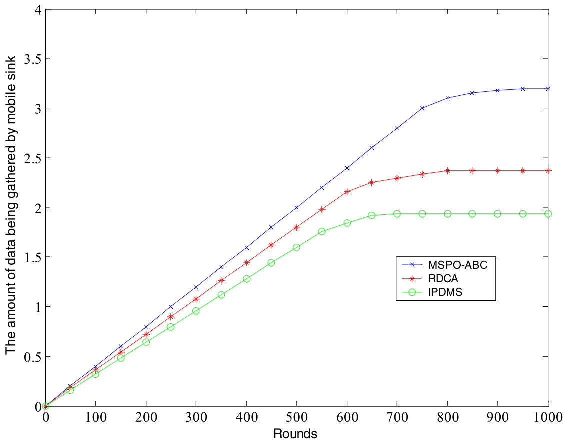

FIGURE 8. **The amount of data being gathered by mobile sink.**

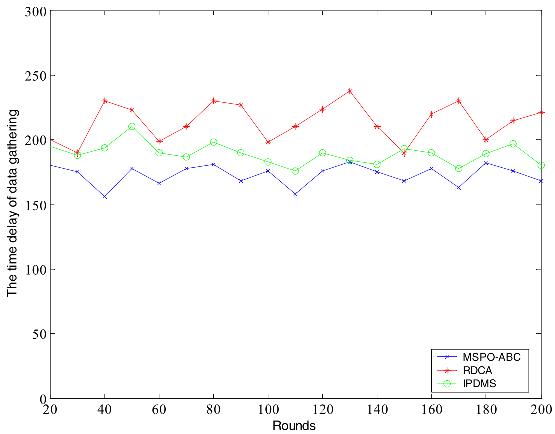

FIGURE 9. **The time delay of data gathering.**

in the network, which directly affects the length of time for data gathering.

Considering that the total delay of the network is not limited, the mobile sink stays at the RPs to receive the message and move to the next one only if all the monitoring data in the grid are sent completely. Figure 9 demonstrates the total time delay by different algorithms. Taking into account the different lifecycle of each algorithm, the sampled rounds without died node are selected in the experiment. Suppose that the velocity of mobile sink is equal to 10m/s, and ε is 100kbit/s. The simulation result proves that the proposed schemes result into less time delay of data gathering. In RDCA, too much number of nodes should be traversed directly, and the distribution of RPs in the network is too dispersive to implement the traversal of mobile sink. In IPDMS, the data collection tour is the perimeter of the sensing area and longer path length results in higher latency in data gathering.

VI. CONCLUSIONS

In this paper, a mobile sink-based path optimization strategy in wireless sensor networks using artificial bee colony algorithm (MSPO-ABC) is proposed. Firstly, the problem of overall energy consumption in the network can be transformed into the minimization of the total hops between all sub nodes and the rendezvous points of mobile sink. The objective function and the Constraint criterion should be established. Secondly, an improved artificial bee colony algorithm is proposed to solve the problem. On the one hand, the cumulative factor is introduced to the position update of the employed bee stage to speed up the convergence of the algorithm. On the other hand, the Cauchy mutation operator is presented to increase the diversity of the feasible solution and enhance the global search ability of the algorithm. In the next work, we will further study the sink mobile strategy under the time delay constraints, as well as the mobile data collection in the multi sink environment and the cooperative communication between nodes.

REFERENCES

[1] D. Kim, B. H. Abay, R. N. Uma, W. Wu, W. Wang, and A. O. Tokuta, "Minimizing data collection latency in wireless sensor network with multiple mobile elements," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 504–512.

[2] C. Konstantopoulos, G. Pantziou, D. Gavalas, A. Mpitziopoulos, and B. Mamalis, "A rendezvous-based approach enabling energy-efficient sensory data collection with mobile sinks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 5, pp. 809–817, May 2012.

[3] B. Tas and A. S. Tosun, "Coordinating robots for connectivity in wireless sensor networks," in *Proc. IEEE 11th Int. Conf. Mobile Ad Hoc Sensor Syst. (MASS)*, Oct. 2014, pp. 452–460.

[4] M. Ma and Y. Yang, "SenCar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 10, pp. 1476–1488, Oct. 2007.

[5] J. Tang, H. Huang, S. Guo, and Y. Yang, "Dellat: Delivery latency minimization in wireless sensor networks with mobile sink," *J. Parallel Distrib. Comput.*, vol. 83, pp. 133–142, Sep. 2015.

[6] M. E. Keskin, İ. K. Altınel, N. Aras, and C. Ersoy, "Wireless sensor network lifetime maximization by optimal sensor deployment, activity scheduling, data routing and sink mobility," *Ad Hoc Netw.*, vol. 17, pp. 18–36, Jun. 2014.

[7] F. Tashtarian, M. H. Y. Moghaddam, K. Sohraby, and S. Effati, "On maximizing the lifetime of wireless sensor networks in event-driven applications with mobile sinks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 3177–3189, Jul. 2015.

[8] H. Salarian, K.-W. Chin, and F. Naghdy, "An energy-efficient mobile-sink path selection strategy for wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2407–2419, Jun. 2014.

[9] A. Munari, W. Schott, and S. Krishnan, "Energy-efficient routing in mobile wireless sensor networks using mobility prediction," in *Proc. IEEE Conf. Local Comput. Netw.*, Zürich, Switzerland, Oct. 2009, pp. 514–521.

[10] L. Guo, R. Beyah, and Y. Li, "SMITE: A stochastic compressive data collection protocol for mobile wireless sensor network," in *Proc. IEEE Conf. Comput. Commun.*, Shanghai, China, Apr. 2011, pp. 1611–1619.

[11] T. S. Chen, H. W. Tsai, and Y. H. Chang, "Geographic convergecast using mobile sink in wireless sensor networks," *Comput. Commun.*, vol. 36, no. 4, pp. 445–458, 2013.

[12] J. Yu, E. Jeong, D.-Y. Seo, K. Park, and G. Jeon, "A dynamic multiagent-based local update strategy for mobile sinks in wireless sensor networks," in *Proc. Int. Conf. Comput. Sci. Appl. (ICCSA)*, Kuala Lumpur, Malaysia, Aug. 2011, pp. 185–196.

[13] K. Lee, Y.-H. Kim, H.-J. Kim, and S. Han, "A myopic mobile sink migration strategy for maximizing lifetime of wireless sensor networks," *Wireless Netw.*, vol. 20, no. 2, pp. 303–318, 2014.

[14] F. Restuccia and S. K. Das, "Optimizing the lifetime of sensor networks with uncontrollable mobile sinks and QoS constraints," *ACM Trans. Sensor Netw.*, vol. 12, no. 1, 2016, Art. no. 2.

[15] C. M. Angelopoulos, S. Nikoletseas, C. Raptopoulos, and D. Patroumpa, "Efficient collection of sensor data via a new accelerated random walk," *Concurrency Comput.-Pract. Exper.*, vol. 28, no. 6, pp. 1796–1811, 2016.

[16] D. Zhu, Y. Guo, and A. Tosun, "Multi-path planning for mobile element to prolong the lifetime of wireless sensor networks," in *Proc. 15th IEEE Int. Conf. Embedded Real-Time Comput. Syst. Appl.*, Aug. 2009, pp. 41–50.

[17] J. Luo and J.-P. Hubaux, "Joint sink mobility and routing to maximize the lifetime of wireless sensor networks: The case of constrained mobility," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 871–884, Jun. 2010.

[18] S.-W. Han, S. H. Kang, and I.-S. Jeong, "Low latency and energy efficient routing tree for wireless sensor networks with multiple mobile sink," *J. Netw. Comput. Appl.*, vol. 36, no. 1, pp. 156–166, 2013.

[19] M. Ma, Y. Yang, and M. Zhao, "Tour planning for mobile data-gathering mechanisms in wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1472–1483, May 2013.

[20] A. Kinalis, S. Nikoletseas, D. Patroumpa, and J. Rolim, "Biased sink mobility with adaptive stop times for low latency data collection in sensor networks," *Inf. Fusion*, vol. 15, no. 1, pp. 56–63, 2014.

[21] K. Akkaya, M. Younis, and M. Bangad, "Sink repositioning for enhanced performance in wireless sensor networks," *Comput. Netw.*, vol. 49, no. 4, pp. 512–534, 2009.

[22] M. Zhao and Y. Yang, "A framework for mobile data gathering with load balanced clustering and MIMO uploading," in *Proc. IEEE INFOCOM Conf.*, Shanghai, China, Apr. 2011, pp. 2759–2767.

[23] C.-F. Wang, J.-D. Shih, B.-H. Pan, and T.-Y. Wu, "A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks," *IEEE Sensors J.*, vol. 14, no. 6, pp. 1932–1943, Jun. 2014.

[24] U. Sharma, C. R. Krishna, and T. P. Sharma, "An efficient mobile data collector based data aggregation scheme for wireless sensor networks," in *Proc. IEEE Int. Conf. Comput. Intell. Commun. Technol. (CICT)*, Ghaziabad, India, Feb. 2015, pp. 292–298.

[25] C.-H. Lee, J. Kwak, and D. Y. Eun, "Towards distributed optimal movement strategy for data gathering in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 2, pp. 574–584, Feb. 2016.

[26] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Hawaii Int. Conf. Syst. Sci. (HICSS)*, 2000, pp. 3005–3014.

[27] A. Rajasekhar, A. Abraham, and M. Pant, "A hybrid differential artificial bee colony algorithm based tuning of fractional order controller for permanent magnet synchronous motor drive," *Int. J. Mach. Learn. Cybern.*, vol. 5, no. 3, pp. 327–337, 2014.

[28] C. M. Hsu, "An integrated portfolio optimisation procedure based on data envelopment analysis, artificial bee colony algorithm and genetic programming," *Int. J. Syst. Sci.*, vol. 45, no. 12, pp. 2645–2664, 2014.

[29] A. Basturk and R. Akay, "Parallel implementation of synchronous type artificial bee colony algorithm for global optimization," *J. Optim. Theory Appl.*, vol. 155, no. 3, pp. 1095–1104, 2012.

[30] J. Jayanth, A. Kumar, S. Koliwad, and S. Krishnashastry, "Artificial bee colony algorithm for classification of remote sensed data," in *Proc. Int. Conf. Ind. Instrum. Control (ICIC)*, May 2015, pp. 1512–1517.

[31] M. Aljuaid and H. Yanikomeroglu, "Investigating the Gaussian convergence of the distribution of the aggregate interference power in large wireless networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4418–4424, Nov. 2010.

[32] F. Yin, C. Fritsche, D. Jin, F. Gustafsson, and A. M. Zoubir, "Cooperative localization in WSNs using Gaussian mixture modeling: Distributed ECM algorithms," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1448–1463, Mar. 2015.

[33] C.-Y. Lee and Y.-J. Park, "Effect of multivariate Cauchy mutation in evolutionary programming," *IEICE Trans. Inf. Syst.*, vol. E97-D, no. 4, pp. 821–829, 2014.

[34] S. Venkatesan, N. Kamaraj, and S. Chellam, "Optimal wheeling transaction based on maximum allowable load of the buyer bus using HPSO with Cauchy mutation," *Int. Rev. Elect. Eng.*, vol. 6, no. 5, pp. 2440–2447, 2011.

[35] L. Bagheri and M. D. T. Fooladi, "A rendezvous-based data collection algorithm with mobile sink in wireless sensor networks," in *Proc. IEEE 4th Int. Conf. Comput. Knowl. Eng. (ICCKE)*. Mashhad, Iran: Ferdowsi Univ. Mashhad, Oct. 2014, pp. 758–762.

[36] P. Komal, K. Nitesh, and P. K. Jana, "Indegree-based path design for mobile sink in wireless sensor networks," in *Proc. IEEE 3rd Int. Conf. Recent Adv. Inf. Technol. (RAIT)*, Dhanbad, India, Mar. 2016, pp. 78–82.

YONG LU received the B.E. degree in mechanics from Zhengzhou University, Zhengzhou, China, in 1991, and the M.E. degree in national defense economy from the Military Economics Academy of PLA, Wuhan, China, in 2005. He is currently an Engineer with the School of Information Engineering, Minzu University of China. His research interests include parallel algorithms and intelligent systems.

NA SUN received the B.E. degree in electronic information engineering from the Hebei University of Technology, Tianjin, China, in 2004, the M.E. degree in communication and information system from Guizhou University, Guiyang, China, in 2007, and the Ph.D. degree in signal and information processing from the Beijing University of Posts and Telecommunications, in 2010. She is currently a Lecturer with the School of Information Engineering, Minzu University of China. Her current research interests include parallel algorithms and intelligent systems.

XIUQIN PAN received the B.E. degree in electric technology and the M.E. degree in power system and automation specialty from Zhengzhou University, Zhengzhou, China, in 1994 and 1999, respectively, and the Ph.D. degree in control theory and control engineering from the Beijing Institute of Technology, in 2002. She is currently a Professor with the School of Information Engineering, Minzu University of China. Her current research interest focuses on parallel algorithms and intelligent systems.