An Energy-Efficient Region Source Routing Protocol for Lifetime Maximization in WSN

CHUAN XU¹, Zhengying Xiong¹, Guofeng Zhao¹ SHUI YU²(Senior Member, IEEE)

¹School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
²School of Software, University of Technology Sydney, Ultimo, NSW, Australia

Corresponding author: Chuan Xu (e-mail: xuchuan@cqupt.edu.cn).

This work was supported in part by the Chongqing Technology Innovation and Application Demonstration Project (cstc2018jzx-cyzdX0120), Innovation Funds of Graduate PhD (Grant No.BYJS201803), and Chongqing Graduate Research and Innovation Project (Grant No.CYB18167,CYB19176).

ABSTRACT As the sensor layer of Internet of Things (IOT), enormous amount of sensor nodes are densely deployed in a hostile environment to monitor and sense the changes in the physical space. Since sensor nodes are driven with limited power batteries, it is very difficult and expensive for wireless sensor networks (WSNs) to extend network lifetime. In order to achieve reliable data transmission in WSNs, energy efficient routing protocol is a crucial issue in extending the network lifetime of a network. However, traditional routing protocols usually propagate throughout the whole network to discover a reliable route or employ some cluster heads to undertake data transmission for other nodes, which both require large amount energy consumption. In this paper, to maximize the network lifetime of the WSN, we propose a novel energy efficient region source routing protocol (referred to ER-SR). In ER-SR, a distributed energy region algorithm is proposed to select the nodes with high residual energy in the network as source routing node dynamically. Then, the source routing nodes calculate the optimal source routing path for each common node, which enables partial nodes to participate in the routing process and balances the energy consumption of sensor nodes. Furthermore, to minimize the energy consumption of data transmission, we propose an effective distance-based ant colony optimization algorithm to search the global optimal transmission path for each node. Simulation results demonstrate that ER-SR exhibits higher energy efficiency, and has moderate performance improvements on network lifetime, packet delivery ratio, and delivery delay, compared with other routing protocols in WSNs.

INDEX TERMS Wireless sensor network, source routing, ER-SR, energy efficient, network lifetime.

I. INTRODUCTION

As the sensor layer of IoT, WSN consists of a large number of small and low-cost sensors, has been applied to a wide variety of applications with vastly varying requirements and characteristics [1]. However, the sensor nodes are usually powered only by batteries but expected to operate for a long period, and it is infeasible and costly to replace once sensor nodes have been deployed, so energy efficiency is always a primary concern in WSNs [2], [3]. Actually, previous studies prove that the consumption of the energy is major consumed in the data transmission, and the transmission performance depends largely on the routing strategy [4]. Therefore, it is necessary to design an energy-efficient routing protocol to greatly save energy and extend the network lifetime.

Many routing protocols have been applied for WSNs. RPL based routing has become a popular protocol for WSNs [5], which can be classified into three categories: 1) Optimizing the route discovery process [6], [7]. The route discovery process optimization methods try to reduce the number of nodes involved in route discovery, which will decrease network control overhead and system energy consumption. 2) Optimizing viable metric [8]–[10]. The path selection optimization methods improve the routing metric to select optimal path with the consideration of power, link quality and energy consumption to achieve energy balance. 3) Transmission power control [11], [12]. To reduce energy consumption of data transmission, the transmission power control methods measure and model
RSS (Received Signal Strength) to obtain optimal data transmission power, which is used to decrease channel interference among sensor nodes. Nevertheless, none of these methods mentioned above takes the solution of “the problem of single path” into account. When a path failure happens, the routing protocol needs to re-establish the network graph for each node, which is inefficient and leads to a serious energy loss.

Moreover, clustering strategy is another kind of effective methods to extend the network lifetime of WSN, which improves the network lifetime from two aspects: 1) Balancing node energy [13]–[15]. These schemes adopts the way in which the nodes turn into cluster heads and build the uneven size cluster according to the node energy conditions, which reduces the node mortality and prolongs the network lifetime. 2) Improving equilibrium distribution of cluster heads [16]–[18]. The authors adopt different cluster head selection methods, and use diverse auxiliary parameters to evenly distribute the cluster heads of the network to ensure the connectivity of the clusters, then the energy consumption of cluster heads for data transmission can be reduced. However, there is still a common problem in the clustering methods, the network is overly dependent on the cluster head, which results in excessive burden on the cluster heads and high energy consumption. Moreover, the frequent replacement of cluster head will cause the network turbulence and redundant computational overhead.

Recently, due to its high flexibility, source routing technology has been introduced to reduce control message overhead of wireless sensor networks. In works [19], [20], various multi-path transmission mechanisms based on source routing technology are proposed for wireless sensor networks. Motivated by these recent developments, in this paper, aiming to prolong the network lifetime, we propose a source routing based energy-efficient region routing protocol (ER-SR), to reduce data transmission energy consumption and balance energy consumption among different nodes jointly. And we construct simulations on NS3 to validate the performance of ER-SR, and the experimental results demonstrate that compared with ER-RPL, MSGR and PRD routing protocols, ER-SR can obtain a great performance superiority of energy consumption, network lifetime, packet delivery ratio, and delivery delay.

The main contributions of this paper include:

- Firstly, to prolong the network lifetime of a wireless sensor network, we try to reduce the data transmission energy consumption and balance the energy consumption of all nodes in the network jointly. Then, we integrate the joint optimization process into the ER-SR protocol to obtain the optimal transmission path, which can improve the energy efficiency of the network significantly.
- Secondly, to balance the energy consumption of the network, we exploit the idea of signaling and data separation into the design of ER-SR protocol. A distributed energy region algorithm is introduced to select the nodes with high energy as the optimal source routing nodes dynamically, which is used to compute the source route for other common nodes.
- Lastly, to minimize the energy consumption of node’s data transmission, we introduce the effective distance to estimate the optimal transmission distance, and extend the ant colony algorithm to find the optimal energy efficient transmission path for each node. Especially, we establish the ring domain and introduce single-hop optimal transmission distance to help the ants to find the optimal search ring domain more quickly, which can expedite the search of the global optimal relay node.

The rest of the paper is organized as follow. Section II reviews the related works. The system model and the ER-SR scheme are described in Section III. We present the details of the distributed energy region algorithm and the effective distance-based ant colony optimization algorithm In Section IV and V. The performance of ER-SR protocol is evaluated through simulations in Section VI, followed by conclusions in Section VII.

II. RELATED WORKS

In this section, we briefly outline the existing researches regarding to our topic, which can be categorized into two aspects: a) energy optimization based on RPL routing protocol. b) energy-efficient research based on clustering routing protocol.

A. ENERGY OPTIMIZATION BASED ON RPL ROUTING PROTOCOL

In a routing RPL, the sink establish a destination oriented directed acyclic graph (DODAG) at a high speed with the trickle algorithm [21], then, each sensor node chooses a parent to forward the data. In the view of energy consumption problem in WSNs composed of sensor nodes, researchers have made a lot of efforts to improve the RPL routing protocol.
Zhao et al. [7] propose a novel energy-efficient region-based RPL routing protocol (ER-RPL), which makes use of the region information to support efficient P2P communication and only requires a subset of nodes to discover the route. ETX [22] and ETT [24] are widely used in real WSNs. ETX reflects the expected transmission counts for a packet to reach its destination, whereas ETT is the expected transmission time of a packet, which employs the forwarding data delay between two nodes as the criterion of link quality to select the node with minimum delay as the parent node. However, neither ETX nor ETT take the residual energy of each sensor node into account, leading to the quick death of sensor nodes with low energy. Lai et al. [10] propose a novel link-delay aware energy efficient routing metric (PRD) for the routing path selection, which designs the routing metric with consideration of the link quality, the residual energy, the distances and the delay. T. G. Harshavardhana et al. [12] adopt the power control technique to address the delay and reliability requirements, which can achieve better packet transmission rate and smaller delay, and improve the performance in terms of power consumption.

However, in the existing methods, the consideration of single path problem is missed, and the path flexibility is not strong enough. When the path failure happens, the RPL protocol needs to re-establish the network graph through sending the control message, which is inefficient and may cause a large amount of signaling overhead.

B. ENERGY-EFFICIENT RESEARCHES BASED ON CLUSTERING ROUTING PROTOCOL

Sensor clustering method has been shown effective in prolonging the wireless sensor network lifetime. The basic idea of sensor clustering method is to categorize the sensors into a set of clusters, and in each cluster coverage, sensors transmit the collected data to their cluster heads (CHs). Each CH aggregates the data packets and forwards them to the sink node either directly or via relaying through other CHs [25].

Heinzelman et al. [26] develop a data aggregating cluster-based routing protocol, named LEACH. In LEACH, they assume a single-hop CH-to-sink connection and adopt the randomized rotation of CHs to ensure a balanced energy consumption. But such assumptions may not guarantee the network connectivity. Therefore, researchers introduce a series of improved protocols based on LEACH, i.e., the fuzzy logic-based clustering algorithm [27], the low energy adaptive clustering LEACH-S [28] and the routing algorithm based on multi-hop communication LEACH-M [29]. Mittal N et al. [14] propose a stable energy efficient clustering protocol, which balances the load among nodes using energy-aware heuristics and ensures higher stability period. Younis et al. [30] propose a hybrid energy-efficient distributed clustering routing (HEED) protocol where the CHs are selected probabilistically based on the primary and secondary parameters. The primary parameter depends on the residual energy of the node, which provides the node with high energy a greater chance to be selected as a cluster head. The secondary parameter depends on the cost of communication within the cluster, which allows the sensor to join the cluster with the lowest communication cost. HEED achieves a uniform distribution of cluster heads and prolongs network lifetime. To further extend network lifetime, Bozorgi et al. [16] propose a hybrid unequal energy efficient clustering protocol (HEEC) for wireless sensor networks. According to the distribution of nodes in the network, HEEC forms clusters of uneven size such that nodes closer to the base station (BS) have more energy to receive and relay data, which increases the lifetime of WSN. Sharma et al. [18] introduce a mode-switched grid-based sustainable routing protocol (MSGR) for WSNs, which divides the network area into several grids of virtual equal size, and selects one node in each grid as the grid head to establish a routing path. In this method, not all grid heads participates in the routing process at the same time, which saves network energy and improves network lifetime.

Unlike most of the previous clustering work that uses a two-layer hierarchy, based on a three-layer hierarchy, Lee et al. [32] proposed a hybrid hierarchical clustering approach (HHCA) by considering a hybrid of centralized gridding for the head selection at upper levels and distributed clustering for the head selection at lower levels, which selects the nodes with the higher residual energy to gather data and route the information. Moreover, in [33]and [34], El et al. propose the fuzzy logic (CFFL) approach to improve the lifetime of WSNs, which uses fuzzy logic for CHs selection and clusters formation processes by using residual energy and closeness to the sink as fuzzy inputs in terms of CH selection, and residual energy of CH and closeness to CHs as fuzzy inputs in terms of clusters formation. In [35], El et al. propose an enhanced clustering hierarchy (ECH) approach, which introduces a sleeping-waking mechanism for overlapping and neighboring nodes to minimize the data redundancy and maximize the network lifetime.

However, there still has a common problem in clustering routing algorithms, networks are overly dependent on the cluster head, which may cause the cluster head to be overburdened and be replaced frequently, which leads to network turbulence and high computational overhead.

III. SOURCE ROUTING BASED ENERGY-EFFICIENT REGION ROUTING PROTOCOL

A. SYSTEM MODEL

1) NETWORK MODEL

In a WSN, we assume that there are \( N \) stationary sensors nodes distributed in \( M \times M \) size area, and the transmission range of each node is \( R_x \). Any two nodes can communicate in multi-hop or single-hop, and the transmission power can be adjusted by the power control method. Sensor nodes in the network can be classified into common node and source routing node, which are denoted as \( c_i \in C = \{c_1, c_2, \ldots, c_m\} \) and \( s_i \in S = \{s_1, s_2, \ldots, s_n\} \) respec-
2) RADIO ENERGY DISSIPATION MODEL

We model the nodes in the network with four states: sending, receiving, sleeping and idling, and node’s energy is mainly consumed in the sending and receiving states. Consequently, the energy consumption can be expressed through the first order radio model [31], which has been widely used to measure the energy dissipation of WSNs [36], [38]. In the first order radio model, to operate the transmitter or receiver circuit, the radio consumption is $E_{elec}$, to achieve acceptable signal-to-noise ratio, the transmitting amplifier is $\varepsilon_{amp}$ and the propagation loss index is $\gamma$, where $\gamma \in [2, 4]$ [37]. The energy consumption for transmitting $k$-bits long packet with a distance $d$ is denoted as,

$$E_T(k,d) = kE_{elec} + k\varepsilon_{amp}d^\gamma$$

$$ = \begin{cases} kE_{elec} + k\varepsilon_{amp}d^2, & \forall d \leq d_0, \\ kE_{elec} + k\varepsilon_{amp}d^4, & \forall d > d_0. \end{cases}$$ (1)

To receive a $k$-bit packet, a sensor node needs additional $E_{DA}(nJ/bit/signal)$ amount of energy for data aggregation [39], and the energy expended is

$$E_R(k) = k(E_{elec} + E_{DA}).$$ (2)

B. THE SCHEME OF ER-SR PROTOCOL

To prolong the network lifetime, we propose an energy-efficient region source routing protocol (ER-SR). The scheme of ER-SR routing protocol is illustrated in FIGURE 3, and the data forwarding process based on ER-SR in WSN is illustrated in FIGURE 4, which includes three main steps: 1) region division; 2) information collection of source routing node; 3) data transmission.

1) REGION DIVISION

In traditional routing protocols, all nodes are involved in the routing discover process. Instead, ER-SR selects a part of nodes as the source routing nodes to achieve routing discovery and path selection for other common nodes, which can reduce the energy consumption of control overhead effectively. With the distributed energy region algorithm, ER-SR divides the network into multiple non-overlapping regions. The main steps of region division are as following.

Firstly, according to the ratio between the residual energy of the node and the average residual energy of its neighbor nodes, the node with high residual energy is selected as the candidate source routing node. Moreover, a time mechanism is introduced to select the node with high residual energy in the transmission range as a source routing node, and there are no "isolated nodes" in the network.

Secondly, setting the source routing node as a reference, the network can be divided into several non overlapping regions. The formation of the regions depends on the energy efficiency between the common node and the source routing node in the following cases.

- Case 1: there is only one source routing node in the range of the common node, so the common node joins
the region where the source routing node is located directly.
- Case 2: connecting with multiple source routing nodes, the common node calculates the communication cost with the distance and the remaining energy of the source routing node, and chooses to join the source routing node with the best energy efficiency.

2) INFORMATION COLLECTION OF SOURCE ROUTING NODES

For applying the source routing into the wireless sensor network, we only use the source routing nodes to establish the DODAG to obtain the information of all nodes. Each source routing node acts as a temporary root node and establishes a temporary DODAG with a minimum hop count as a routing metric after region division. When there are \( M \) source routing nodes in the wireless sensor network, \( M \) temporary DODAGs need to be constructed. If \( M - 1 \) source routing nodes are added to the DODAG established by any one of them, they will send the location information of the nodes in the area to the temporary root node through temporary DODAG.

3) DATA TRANSMISSION

To achieve energy efficient transmission, we introduce source routing and energy efficiency distance in ER-SR. Firstly, we establish a ring domain for each source routing node and propose effective distance as the best decision threshold to search the optimal transmission ring domain. Then, we extend the ant colony algorithm to search the global optimal relay node in the ring domain for each node. Especially, a local pheromone update strategy and a global pheromone update strategy are introduced into the ant colony algorithm to avoid it falling into local optimum, which can accelerate the optimal path searching process. Finally, the data from the common node is sent to the source routing node in the region to perform the path request, and the source routing node encapsulates the selected path information into the source routing header of the packet. Moreover, in the process of data transmission, the intermediate nodes along the path perform relay forwarding according to the source routing path information in the packet header without any communication to the source routing node, which can reduce the system energy consumption.

Accordingly, in the above three steps of ER-SR, we need to address these issues how to divide the region and how to select an energy efficiency route to extend the network lifetime.

IV. DISTRIBUTED ENERGY REGION ALGORITHM

In this section, to balance energy consumption, we divide the network into several non-overlapping areas with the reference of source routing nodes, and proposes a distributed energy region algorithm to achieve the region division and source routing node selection. The control messages used are illustrated in Table II.

A. CANDIDATE SOURCE ROUTING NODE SELECTION

In wireless sensor network, a node broadcasts \( \text{Node}_M\_\text{Msg} \) (including the location and residual energy) to its neighbor nodes within the transmission range \( R_c \) and receives \( \text{Node}_M\_\text{Msg} \) informations from its neighbor nodes. The average residual energy of its neighbor nodes can be obtained as following,

\[
E_{ia} = \frac{1}{K_i+1} \left( \sum_{j=1}^{K_i} E_{jr} + E_{ir} \right),
\]

(3)

where \( E_{ia} \) is the average residual energy of all nodes within node \( i \)'s transmission range \( R_c \), \( E_{ir} \) and \( E_{jr} \) are the residual energy of node \( i \) and node \( j \) respectively. \( K_i \) is the number of neighbor nodes of the node \( i \). If \( E_{ir} > E_{ia} \), the node is the candidate source routing node; not otherwise.

B. SOURCE ROUTING NODE SELECTION

We assign two periods for the process of source routing node selection: The first period for the candidate source routing node competition, the duration is \( T_1 \); The second period for the source routing node selection of the "isolated node", which refers to the node that does‘t receive \( \text{Source}_M\_\text{Msg} \) and is not the candidate node, and the duration is \( T_2 \).

In the candidate source routing node competition period, we introduce a time mechanism to select the source routing node.
node, and the time for waiting the candidate node to broadcast $Source_{Msg}$ is obtained as following,

$$t_{ca} = \frac{E_{ia}}{E_{ir}} T_{1} V_{r}, \ \forall E_{ir} > E_{ia},$$

(4)

where $t_{ca}$ is the time for waiting the candidate source routing node to broadcast its $Source_{Msg}$ information, $E_{ir}$ is the residual energy of the node $i$, and $V_{r}$ is a random real number between [0.9,1], which is used to reduce the collision of sending $Source_{Msg}$ information from multiple nodes simultaneously. When $t_{ca}$ expires, if a node hasn’t received any $Source_{Msg}$ from neighbor nodes, it will broadcast its own $Source_{Msg}$ information as a source routing node. Otherwise, if it receives the $Source_{Msg}$ from neighbor nodes, it will abandon the competition and become a common node.

When $T_{1}$ expires, the nodes that have not received the $Source_{Msg}$ from their neighbors will enter the source routing node selection period for the 'isolated nodes'. The time for waiting the "isolated nodes" to send $Source_{Msg}$ information is determined by,

$$t_{iso} = \frac{1}{E_{ir}} T_{2} V_{r}, \ \forall E_{ir} \leq E_{ia},$$

(5)

where $t_{iso}$ is the time for waiting the "isolated node" to broadcast $Source_{Msg}$ information.

Obviously, the node with high residual energy has a greater chance to become a source routing node. From (4) and (5), the higher the residual energy, the smaller the values of $t_{ca}$ and $t_{iso}$, and the energy of the final source routing node by our time mechanism is relatively high in the network. Moreover, we only select one source routing node within its transmission range, which ensures the equilibrium distribution of the source routing nodes.

C. REGION FORMATION CONSTRUCTION

The region formation is entirely depend on the energy efficiency between the source routing nodes and the common nodes. To extend the network lifetime, the energy efficiency determines which region of the source routing node is better for the common node. The region formation is performed through the procedure of control message exchange between source routing nodes and common nodes. The control messages are listed in Table 2, according to the number of source routing nodes in the transmission range of a common node, we classify the process of the region formation construction into two cases.

![FIGURE 4: ER-SR based data forwarding process in WSN.](image)

a) Multiple common nodes in the range of one source routing node. b) One common node in the range of three source routing nodes.

![FIGURE 5: The distribution of the source routing nodes.](image)
TABLE 2: List of control message used in the region formation process

<table>
<thead>
<tr>
<th>Message</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node_Msg</td>
<td>Broadcast node information by each node</td>
</tr>
<tr>
<td>Source_Msg</td>
<td>Notification of source routing node roles by source</td>
</tr>
<tr>
<td></td>
<td>routing nodes</td>
</tr>
<tr>
<td>Response_Msg</td>
<td>Receipt of Source_Msg informed by common nodes</td>
</tr>
<tr>
<td>Wait_Msg</td>
<td>Receipt of Response_Msg informed by source routing</td>
</tr>
<tr>
<td></td>
<td>nodes</td>
</tr>
<tr>
<td>Join_Msg</td>
<td>Region joining requested by common nodes</td>
</tr>
<tr>
<td>Allow_Msg</td>
<td>Allowance of region joining confirmed by source</td>
</tr>
<tr>
<td></td>
<td>routing nodes</td>
</tr>
</tbody>
</table>

FIGURE 6: Creating a region with n=1.

information, to inform the common nodes that the region formation process started. If these three nodes only receive one Source_Msg message from the source routing node s1, they will join the region where the source routing node s1 located without consideration of the energy efficiency. The process of the first case is illustrated in FIGURE 6, when the nodes c1, c2 and c3 decide to join the region of s1, they will send a Join_Msg message to the source routing node s1. The Join_Msg message contains the identification number, residual energy and position information to acknowledge the receipt of the Source_Msg message. If the Join_Msg message is received successfully by the source routing node s1 and node s1 accepts the joining request, it will send an Allow_Msg message to those three common nodes. After the source routing node updates its routing table, and region formation construction is completed. As illustrated in FIGURE 5 (b), in the second case, there are more than one source routing node in the communication range of the common node (i.e. n = 3), the common node c1 can communicate with the source routing nodes s1, s2 and s3. Similarly to the first case, source routing nodes s1, s2 and s3 broadcast a Source_Msg message to their neighbor common node c1 to notice their roles. Then, to feed back the response of the Source_Msg message from three source routing nodes, the common node c1 broadcasts a Response_Msg message. And a Wait_Msg message containing the residual energy and location information of the source routing node is sent from nodes s1, s2 and s3 respectively as a reply to c1. Based on the information of the Wait_Msg message, we propose $p_{c1}^{s_j}$ to describe the probability that common node $i$ chooses node $j$ as its source routing node, which can be obtained by

\[ p_{c1}^{s_j} = \frac{1}{e_{sr} \left( d_{c1,sj} \right)} \]

where $d_{c1,sj}$ is the distance between the common node $i$ and the source routing node $j$, $d_{c1,sj} = \sqrt{(x_{ci} - x_{sj})^2 + (y_{ci} - y_{sj})^2}$, $e_{sr}$ is the residual energy of the source routing node $j$. The common node c1 chooses to join the region of the source routing node with best $p_{c1}^{s_j}$ and send a Join_Msg message to inform nodes $s1, s2$ and $s3$. When the Join_Msg message is successfully received, an Allow_Msg will be sent by the source routing node as a region join permission. The process of creating a region for this case is summarized in FIGURE 7.

Moreover, the pseudo-code of the distributed energy region algorithm is presented in Algorithm 1. Obviously, the time complexity of the proposed distributed energy region algorithm is $O\left(N^2\right)$.

V. EFFECTIVE DISTANCE-BASED ANT COLONY OPTIMIZATION ALGORITHM

Next, aiming to find an optimal path, we propose an effective distance-based ant colony optimization algorithm in this section. We first introduce effective distance to describe the path priority, which can be calculated by the mathematical derivation. Moreover, the effective distance can be used as the optimal decision threshold to facilitate the ants to find the best next hop relay node for the optimal transmission path searching. Based on the effective distance, we propose

FIGURE 7: Creating a region with n>1.
Algorithm 1 Distributed energy region algorithm
1: Each node broadcasts Node_Msg;
2: Calculate $E_{ina}$ by (3);
3: Compare $E_{ina}$ and $E_{ira}$ of node $i$;
4: while $(E_{ira} > E_{ina})$ do
5: Calculate $t_{ca}$ by (4) for broadcasting Source_Msg;
6: while ($T_1$ hasn’t expired) do
7: if $t_{current} < t_{ca}$ then
8: if Receive Source_Msg from neighbor node then
9: Abandon sending Source_Msg;
10: else
11: Continue;
12: end if
13: else
14: Broadcasting Source_Msg;
15: end if
16: end while
17: while ($E_{ira} \leq E_{ina}$ and hasn’t received Source_Msg from neighbor nodes) do
19: Calculate $t_{ca}$ by (5) for broadcasting Source_Msg;
20: while ($T_2$ hasn’t expired) do
21: if $t_{current} < t_{isa}$ then
22: if Receive Source_Msg from neighbor node then
23: Abandon sending Source_Msg;
24: else
25: Continue;
26: end if
27: else
28: Broadcasting Source_Msg;
29: end if
30: end while
31: if receives only one Source_Msg then
32: Send Join_Msg directly;
33: else
34: Calculate $p_{j,i}^m$ by (6);
35: Send Join_Msg to the node with optimal $p_{j,i}^m$;
36: end if

Theorem 1 (Effective distance): In a multi-hop wireless transmission system, the smaller the difference in transmission distance between each hop, the best the energy efficiency of wireless transmission system can be obtained [40]. Therefore, we define the effective distance as the transmission distance between each hop, and assume that a data link between the source node and the destination node at distance $D$ is divided into $x$ hops by $(x-1)$ intervening relay nodes. Given the distance $D$ and the number of hops $x$, the total energy usage along the path can achieve the minimum when each hop shares the same transmission distance $d = D/x$.

Therefore, to determine the optimal relay, we calculate the effective distance $d$ by giving the total energy consumption $E_{total}$ of the path as follow.

$$E_{total} = xE_{R}(k) + (x - 1)E_{R}(k) = x(kele + k\varepsilon_{amp}d^\gamma) + (x - 1)(kele + E_{DA}) = (2x - 1)kele + x(k\varepsilon_{amp}d^\gamma + E_{DA}) - kE_{DA} = (2D/d - 1)kele + D(k\varepsilon_{amp}d^\gamma + E_{DA}) - kE_{DA}. \quad (7)$$

Lemma 1 (Convexity): The objective $E_{total}$ is convex. When all the hop distances $d$ are set equal to $D/x$, the total energy consumption $E_{total}$ can be minimized.

Proof 1: Based on the detailed analysis in Appendix, the convexity of $E_{total}$ in the case of equal-hop transmission can be proved.

To find the minimum $E_{total}$, we set the first derivative of (7) equal to zero, i.e.

$$Dk(\gamma - 1)\varepsilon_{amp}d^{\gamma - 2} - Dk(2E_{elec} + E_{DA})d^{-2} = 0. \quad (8)$$

Then, we can obtain the effective distance as following,

$$d_e = \sqrt{\frac{2E_{elec} + E_{DA}}{(\gamma - 1)\varepsilon_{amp}}}. \quad (9)$$

B. EFFECTIVE DISTANCE-BASED ANT COLONY OPTIMIZATION ALGORITHM

In ant colony optimization algorithm (ACO) [23], when an ant moves, it lays varying amount of pheromones, which are detectable by other ants along its path, thereby marking the path by a trail of such substances. As more ants pass by, more pheromones are deposited on the path. The richer the trail of pheromones in a path, the more likely it would be followed by other ants. Therefore, the ACO heuristic algorithm can be easily employed to find a optimal path.

In this section, we propose an effective distance-based ant colony optimization algorithm to find an energy-efficient transmission path for common nodes, which is deployed on source routing nodes. Especially, in each searching interval, a couple of ants are launched from a source routing node to find an energy-efficient path between the source node and the destination node.

As illustrated in FIGURE 2, the network area of a node is divide into multiple rings for seeking an optimal transmission distance. When an ant moves from ring $\varpi_m$ to $\varpi_n$ toward the destination node, it will create a corresponding path and the distance is

$$d_{mn} = (n - m)r. \quad (10)$$
Based on Theorem 1, the energy consumption will have a minimum when the transmission distance is close to the effective distance $d_e$. Therefore, we can determine the next hop ring area $\omega_n$ of the ant to move by

$$\min |d_{mn} - d_e|. \quad (11)$$

The ant selects the next hop node based on the same probabilistic rule at each node of ring area. For the $t$-th iteration, the transition probability of the ant from node $i$ to node $j$ is defined as

$$P_{ij}(t) = \begin{cases} 1, & \text{if } i \in S, j \text{ is a source node}, \\ \frac{(\eta_{ij}(t))^\alpha (\tau_{ij}(t))^\beta}{\sum_{z \in \omega_n} (\eta_{iz}(t))^\alpha (\tau_{iz}(t))^\beta}, & \text{otherwise}. \end{cases} \quad (12)$$

Especially, in order to find an energy efficient path between the source node and the destination node, we set the first hop of the ant as the source node $P_{ij}(t) = 1$, the node $i$ is source routing node and $j$ is source node. In (12), where $\eta_{ij}(t)$ is the desirability of node transition, $\tau_{ij}(t)$ is the amount of pheromone deposited for the transition from node $i$ to node $j$. The two parameters $\alpha$ and $\beta$ are constants, which are used to adjust the relative of the pheromone and the desirability respectively on the path selection of the ant.

In (12), the desirability of the path from node $i$ in $\omega_m$ to node $j$ in $\omega_n$ is defined as

$$\eta_{ij}(t) = \frac{1}{d_{ij}^\alpha + d_{j, dst}^\beta + \theta + \sum_{z \in \omega_n} E_{zr}^\alpha}, \quad (13)$$

where $d_{ij}$ is the distance between node $i$ and node $j$, $d_e$ is the effective distance, $d_{j, dst}$ is the distance between node $j$ and destination node. To ensure the denominator is not zero, we set the constant $\theta > 0$. $E_{zr}$ is the residual energy of node $j$, and the two parameters $\lambda_1$ and $\lambda_2$ are used to describe the weighting factors of the effective distance and the residual energy respectively.

We can find that in the direction of destination node, the closer the transmission distance between the two nodes to $d_e$ is, the better energy efficiency will be obtained. However, since the sensor nodes are always with very limited energy capacity, the optimal path between source node and the destination node should be determined not only in terms of the distance but also the energy level of the path. Therefore, in our effective distance-based ant colony optimization algorithm, the residual energy of nodes and energy efficiency distance are considered in the definition of desirability, which help the ant to select the path with high energy level that can achieve high energy efficiency to prolong the network lifetime.

Then, the pheromone intensity of the path is updated by

$$\tau_{ij}(t + 1) = (1 - \rho)\tau_{ij}(t) + \rho \Delta \tau_{ij}(t), \quad (14)$$

where $\tau_{ij}(t)$ is the amount of pheromone deposited for a node transition, $\rho \in (0, 1)$ is the pheromone evaporation coefficient. $\Delta \tau_{ij}(t)$ is the pheromone increment.

Moreover, in the ant colony algorithm, the pheromone is a continuously accumulated variable, the ant colony selects the path with more pheromone in the exploration of the path, and other optimal paths may be ignored, which may result in the final path found is a local optimal solution. Therefore, we propose two pheromone update mechanisms to avoid the problem of local optimization.

- **Local pheromone update mechanism.** When an ant moves from node $i$ to node $j$, the increased pheromone between two nodes is updated according to

$$\Delta \tau_{ij}(t) = \frac{Q E_{jr}}{E_i}, \quad (15)$$

where $Q$ is a constant, $E_i$ is the energy consumption of node $i$, which can be obtained from (7). It can be found that the number of pheromone increment is determined by the value of energy consumption on the path, and the node with high energy is selected as the next hop. Consequently, the node with low energy is protected and the network energy is balanced.

- **Global pheromone update mechanism.** When all ants arrive at the destination node, the backward ants along the reverse path go back to the node which it was created and update the increased pheromone of the path according to

$$\Delta \tau^{*}_{ij}(t) = \frac{Q}{E_{total}}. \quad (16)$$

We consider the total energy consumption $E_{total}$ as the evaluation criterion of the path

$$E_{total} = \sum_l E_l, \quad (17)$$

where $l$ is the node contained in the path. It can be found that the lower the energy consumption of the path, the more pheromones of the optimal path. Ants are inclined to choose the path with larger pheromone, so that ant’s search behavior may quickly concentrate on the optimal path and the efficiency of path searching can be improved.

The effective distance-based ant colony optimization algorithm employs the effective distance to achieve the optimal transmission with the local and global pheromone update mechanisms, the pseudo-code of which is illustrated in Algorithm 2. And according to the effective distance-based ant colony optimization algorithm, suppose that the number of nodes is $N$ and the number of ants is $k$. The maximum number of iterations is $\Omega_{max}$. In each iteration, one solution is constructed and the maximum of $k \times N$ pairs are examined. Therefore, the time complexity of one iteration is $O(k N^2)$. Additionally, the updating of the parameters (pheromone) can be performed in $O(N)$ time. For $\Omega_{max}$ iterations, the total time complexity of the effective distance-based ant colony optimization algorithm is $O(k N^2 \Omega_{max})$. Thus, according to the time complexity in distributed energy region algorithm and effective distance-based ant colony optimization algorithm, we can obtain the time complexity of the proposed ER-SR is $O(k N^2 \Omega_{max})$. 

VOLUME 4, 2016
Algorithm 2: Effective distance-based ACO algorithm

1: Set initial pheromone value of $\tau_{ij}(0)$;
2: Calculate the value of the effective distance $d_e$;
3: while (doesn’t reach the maximum number of iterations $\Omega_{\text{max}}$) do
4:   Ants move from the ring $\varpi_1$ to the outside one;
5:   for (Ants in sector $\varpi_m$) do
6:      Calculate the distance $d_{mn}$ to any other ring $\varpi_n$;
7:      Find min $|d_{mn} - d_e|$;
8:      Identify the sector $\varpi_n$ to be transferred;
9:      Calculate $p_{ij}(t)$;
10:     Calculate the transition probability $P_{ij}(t)$;
11:    Move from node $i$ in sector $\varpi_m$ to node $j$ in sector $\varpi_n$;
12:   Update the value of $E_i$ transmitted by node $i$;
13:   Calculate the local pheromone increment $\tau_{ij}(t)$;
14:   Update the local pheromone value of $\Delta_\tau_{ij}(t)$;
15: end for
16: Calculate the value of $E$ for each path;
17: if $E = \min E$ in all paths then
18:   Calculate the global pheromone increment $\Delta_\tau^*_ij(t)$;
19:   Update the global pheromone value of $\tau_{ij}(t)$;
20: else
21:    Break;
22: end if
23: end while
24: Find the optimal solution.

TABLE 3: Simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network coverage</td>
<td>$100m \times 100m$</td>
</tr>
<tr>
<td>Number of nodes(N)</td>
<td>$100 \sim 300$</td>
</tr>
<tr>
<td>Threshold distance(d0)</td>
<td>$75m$</td>
</tr>
<tr>
<td>Transmitter electronics(E_{elec})</td>
<td>50 mJ/bit</td>
</tr>
<tr>
<td>Transmit amplifier(e_{amp})</td>
<td>0.0013 P_{j}/bit/m^2</td>
</tr>
<tr>
<td>Packet size</td>
<td>$512bytes$</td>
</tr>
<tr>
<td>Control message size</td>
<td>$256bytes$</td>
</tr>
<tr>
<td>Initial energy</td>
<td>$1J$</td>
</tr>
<tr>
<td>Transmission radius of nodes (R_c)</td>
<td>20m</td>
</tr>
<tr>
<td>Data aggregation energy (E_{DA})</td>
<td>5nJ/bit/signal</td>
</tr>
</tbody>
</table>

VI. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our proposed ER-SR and its extension via simulation experiments. We first describe our simulation environments and performance metrics, and then evaluate the performance results. Finally, we show the comparison among ER-SR protocol, ER-RPL, PRD and MSGR routing protocols.

A. SIMULATION SETUP

We design experiments via NS3.28, which is used to evaluate the performance of ER-SR. We select IEEE802.11 as the MAC protocol. For fairly comparing, the network configuration is similar to that shown in [10] and [18]. The number of sensor nodes randomly distributed in a 2D area of $100m \times 100m$ is set to 100, and varies from 100 to 300. Each sensor node has an initial energy of $1J$, and source-destination pairs are randomly selected in the network. The maximum transmission range $R_c$ of sensor nodes is 20m. Moreover, the error erasure channel is used to model the lossy environment, and the wireless link is assigned with a random probability $P_{ij}$, where $0.3 < P_{ij} < 0.8$ [7].

The specific simulation environment parameters used in the experiments are listed in Table III. We conduct the experiments with random topology in each run, and repeat the simulations 50 times to obtain more accurate results.

B. PERFORMANCE METRICS

We investigate four performance metrics, including energy consumption, network lifetime, packet delivery ratio, and delivery delay, to evaluate our scheme.

- Energy consumption: That is defined as the total energy consumed by all sensor nodes which have participated
in data delivery.

• Network lifetime: That is defined as the time that 20 percent of nodes in the network die.

• Packet delivery ratio: That is defined as the ratio between the number of successfully delivered data packets and the number of data packets generated by the source node.

• Delivery delay: That is defined as the time delay from the generation of the packet to its delivery to the destination.

C. PERFORMANCE EVALUATION

An temporary topology obtained in a WSN with 100 nodes running in a $100m \times 100m$ region is shown in FIGURE 8. In FIGURE 8, the blue node is the source routing node and the red node is the common node. The connection in the figure represents the temporary DODAG topology formed when the source routing node with coordinates (16.6, 21.5) acts as the temporary root node, and the DODAG is established and maintained by transmitting DIO (DODAG Information Object). Moreover, it can be found that the uniform distribution of source routing nodes (blue nodes) can be obtained by the proposed ER-SR routing protocol.

FIGURE 9 depicts the value of network lifetime in considering the residual energy and effective distance, where the blue indicates the ER-SR with consideration of the effective distance and the residual energy, and the green and yellow indicate the routing scheme considering only the effective distance and the residual energy respectively. When the number of network nodes is small, the effective distance is not the main factor to improve the network lifetime since there are fewer spacing between nodes just equal to the effective distance. Therefore, in the routing protocol that only considers the effective distance, the network lifetime is obviously lower than others. As the number of nodes increases, the relay selection based on effective distance saves a lot of transmission energy consumption, which prolongs the network lifetime. As the number of nodes continues to increase, the choice of relay nodes increases. If only the remaining energy of the nodes is considered, a large amount of transmission energy consumption restricts the performance improvement; if only the effective distance is considered, the death of some fixed nodes is accelerated and network lifetime is reduced. Therefore, it can be seen that under different network density, the ER-SR performance with joint consideration of energy and distance is always
superior to others.

FIGURE 10 displays the total energy consumption of ER-SR, ER-RPL, MSGR and PDR for varying number of sensor nodes. It is obvious that the energy consumption decreases drastically as the number of nodes increases, because the increase in the number of nodes helps them find more energy efficient paths. ER-RPL establishes network-wide DODAG for routing discovery, which results in significantly higher energy consumption than other three routing protocols. Compared with MSGR and PRD, our proposed ER-SR only needs a part of nodes participated in routing process, which saves the control message overhead. Moreover, based on source routing mechanism, ER-SR doesn’t require signaling interaction in data transmission, which also reduces the signaling overhead and energy consumption. Therefore, our proposed ER-SR protocol is better than the other three routing protocols in terms of energy consumption.

FIGURE 11 demonstrates the network lifetime of ER-SR, ER-RPL, MSGR and PDR with different number of sensor nodes. It can be seen that when the network nodes increases, the network lifetime of all the routing protocols has a significant increase. The reason is that the increase in the number of nodes increases the choice of the optional energy-saving path, which can increase the average residual energy of the sensor nodes in WSN, thereby increasing the network lifetime. When the number of nodes is small, because the distance between nodes may be greater than the effective distance, which restricts the improvement of routing performance in ER-SR. However, since the source routing and the regional division are introduced to guarantee the balanced energy consumption of WSN effectively, the network lifetime of ER-SR is still better than the ER-RPL, MSGR and PDR.

FIGURE 12 demonstrates the packet delivery ratio under different number of sensor nodes. The results indicate that all four routing protocols have higher packet delivery ratio with more sensor nodes. This is because all of them can find more efficient paths for data delivery, and the probability of packet loss due to node death will decrease. Compared with ER-RPL, MSGR and PRD, our proposed ER-SR considers the residual energy of the node and the optimal effective distance to achieve optimal transmission, which ensures the reliability of the next hop node in terms of residual energy. Therefore, our proposed ER-SR outperforms the other three routing protocols in packet delivery ratio.

FIGURE 13 shows the delivery delay with varied number of sensor nodes. The results show that the delivery delay of our proposed ER-SR is lower than ER-RPL, MSGR and PRD as the number of sensor nodes increases. The main reasons are summarized as follows: Firstly, ER-SR introduces the energy efficiency distance as the optimal transmission distance, and establishes a ring domain, which helps the ant to quickly find the optimal path in effective distance-based ant colony optimization algorithm. Secondly, based on source routing mechanism, ER-SR doesn’t require signaling interaction during data transmission, the relay node only forwards the data packet according to the source routing path. Finally, the high packet delivery ratio reduces the probability of retransmission, which reduces the transmission delay.

VII. CONCLUSION

In this paper, we have addressed the issue of energy balance and energy consumption minimization in WSN. We design a distributed energy region algorithm to dynamically select nodes with high energy as source routing nodes, which is used to compute route for other common nodes. Moreover, we introduce the effective distance as a criterion to describe the optimal transmission distance and propose the effective distance-based ant colony optimization algorithm to find an optimal source routing path for each common node. By these innovations, The comparison experiments show that the proposed algorithm has good performance in the energy consumption, network lifetime, packet delivery ratio, and delivery delay. As this work mainly focuses on static networks, we plan to extend the region source routing protocol to mobile networks in our future work. Moreover, the further work will provide the theoretical analysis of network lifetime maximization including all the network overloads brought by the routing protocol.

APPENDIX[PROOF OF THE CONVEXITY]

We suppose that the total energy consumption $E_{\text{total}}$ is a function of the variable $d$.

\[ g(d) = f(d) - (kE_{\text{elec}} + kE_{\text{DA}}), \]
\[ f(d) = (2D \cdot kE_{\text{elec}} + D \cdot kE_{\text{DA}})d^{-1} + D \cdot k\varepsilon_{\text{amp}}d^\gamma. \]

Since the second part of $kE_{\text{elec}} + kE_{\text{DA}}$ is constant, which is both convex and concave. We only need to demonstrate the convexity of the part $f(d)$. Through introducing the parameter $t \in (0, 1)$, we have

\[ tf(d^1) = t \cdot D \cdot k(2E_{\text{elec}} + E_{\text{DA}})(d^1)^{-1} + t \cdot D \cdot k\varepsilon_{\text{amp}}(d^1)^{\gamma - 1}, \]
\[ (1-t)f(d^2) = (1-t)(2D \cdot kE_{\text{elec}} + D \cdot kE_{\text{DA}})(d^2)^{\gamma - 1} + (1-t)D \cdot k\varepsilon_{\text{amp}}(d^2)^{-1}. \]

\[ t f(d^1) + (1-t)f(d^2) = \left[ t(2DkE_{\text{elec}} + DE_{\text{DA}})(d^1)^{-1} + (1-t)(2DkE_{\text{elec}} + DE_{\text{DA}})(d^2)^{-1} \right] 
+ \left[ tDk\varepsilon_{\text{amp}}(d^1)^{\gamma - 1} + (1-t)Dk\varepsilon_{\text{amp}}(d^2)^{\gamma - 1} \right] \]

Also, we have

\[ t f(d^1) + (1-t)f(d^2) = (2DkE_{\text{elec}} + DE_{\text{DA}})(td^1 + (1-t)d^2)^{-1} + Dk\varepsilon_{\text{amp}}(td^1 + (1-t)d^2)^{\gamma - 1} \]
\[ \leq t(2DkE_{\text{elec}} + DE_{\text{DA}})(d^1)^{-1} + (1-t)(2DkE_{\text{elec}} + DE_{\text{DA}})(d^2)^{-1} + Dk\varepsilon_{\text{amp}}(d^1)^{\gamma - 1} + (1-t)Dk\varepsilon_{\text{amp}}(d^2)^{\gamma - 1} \]
\[ = tf(d^1) + (1-t)f(d^2). \]
Therefore, we can prove that both the \( f(d) \) is convex and \( E_{total} \) is also convex.

REFERENCES


