

# Efficient group-based discovery for wireless sensor networks

Pengpeng Chen, Ying Chen, Shouwan Gao, Qiang Niu and Jun Gu

## Abstract

Due to the combination of constrained power, low duty cycle, and high mobility, neighbor discovery is one of the most challenging problems in wireless sensor networks. Existing discovery designs can be divided into two types: pairwise-based and group-based. The former schemes suffer from high discovery delay, while the latter ones accelerate the discovery process but incur too much energy overhead, far from practical. In this article, we propose a novel efficient group-based discovery method based on relative distance, which makes a delicate trade-off between discovery delay and energy consumption. Instead of directly referring to the wake-up schedules of a whole group of nodes, efficient group-based discovery selectively recommends nodes that are most likely to be neighbors, in which the probability is calculated based on the nodes' relative distance. Moreover, the sequence of received signal strengths are employed to estimate the relative distance for avoiding the effect of the node distribution. Extensive simulations are conducted to verify the effectiveness of the design. The results indicate that efficient group-based discovery statistically achieves a good trade-off between energy cost and discovery latency. Efficient group-based discovery also shows one order of magnitude reduction in discovery delay with a maximum of 6.5% increase in energy consumption compared with typical discovery methods.

## Keywords

Wireless sensor networks, neighbor discovery, group-based discovery, relative distance

Date received: 8 January 2017; accepted: 12 June 2017

Academic Editor: Miltiadis Lytras

## Introduction

Wireless sensor networks (WSNs) are promising technologies that can be widely applied to many fields,<sup>1</sup> such as emergency rescue, biological detection, environmental measurement, and real-time monitoring.<sup>2</sup> Due to the energy constraints of nodes, the low duty-cycle mechanism is often adopted in WSNs, which saves the energy by listening briefly and shutting down radio most of the time. Such a mechanism makes it difficult for nodes to discover their neighbors because nodes have to listen the channel in an asynchronous manner. The problem will become even more challenging if nodes are mobile in the network.

Neighbor discovery in low duty-cycle WSNs has attracted extensive research attention. In such networks, each node periodically wakes up and sleeps.

If nodes are within the communication range of each other, then they can communicate, such that a neighbor discovery is completed. Currently, there exist two types of discovery methods: pairwise-based and group-based. The pairwise discovery algorithms are achieved only when two nodes are simultaneously pre-scheduled to wake up. In group-based discovery methods, nodes can share their known schedules with each other, and the discovery process is a proactive and fast manner with

School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China

### Corresponding author:

Shouwan Gao, School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China.  
Email: gaoshouwan@cumt.edu.cn



Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<http://www.uk.sagepub.com/aboutus/openaccess.htm>).

energy overhead. In other words, traditional pairwise discovery algorithms focus on reducing energy consumption, while group-based discovery methods reduce discovery delay with increased energy consumption. The aforementioned algorithms don't make a good trade-off between energy consumption and discovery delay, which are not suitable for practical applications. Thus, the need for a novel neighbor discovery method arises.

Facing the requirements mentioned above, we propose a relative distance based on group discovery scheme for leveraging the schedule reference mechanism such that the speed of neighbor discovery can be accelerated. Our relative distance design is motivated by calculating neighbor probability by means of the number of common neighbors two nodes have, which may underutilize the mobility of node. Although received signal strength (RSS) is not a good choice for physical distance estimation in many scenarios,<sup>3</sup> it does provide certain useful distance-related information in addition to indicating connectivity among neighboring nodes. Subsequently, we can determine the near-far relationship between two nodes and dynamically adjust duty cycle according to neighbor probability. Technically, the key novelty of efficient group-based discovery (EGD) is that it leverages knowledge in the neighbor tables of known neighbors to accelerate discovery of unknown neighbors, while effectively avoiding excessive reference operations.

The main contributions of our work are as follows:

1. Most of the previous works focus on scheduling designs for pairwise discovery with high discovery delay. We propose the EGD based on relative distance, which can reduce discovery delay with small overhead.
2. Using relative distance to indicate the proximity between two neighboring nodes, we can adjust the duty cycle such that the neighbor discovery efficiency can be improved and the discovery delay can be reduced.
3. The performance of the proposed design is evaluated by extensive simulation studies. The results show that EGD achieves better trade-off between discovery delay and energy consumption compared with existing works.

The remainder of this article is organized as follows: the related works are briefly discussed in "Related work," while section "Network model" introduces the network model. The main design and theoretical analysis are described in section "System design." The simulation results are presented in section "Evaluation." Finally, section "Conclusion" concludes the whole article.

## Related work

As a critical component of WSNs, neighbor discovery is a protocol that allows different nodes on the same link to advertise their existence to neighbors and learn about the existence of neighbors. Neighbor discovery algorithms of WSNs in low duty cycle have been studied,<sup>4</sup> and many results have been found. Typical low-duty-cycle neighbor discovery algorithms can be categorized into one of the following two classes: pairwise-based<sup>5–14</sup> and group-based.<sup>15,16</sup> In pairwise-based discovery algorithms, we can further divide into three types: stochastic-based protocols,<sup>5,6</sup> quorum-based protocols,<sup>7–11</sup> and deterministic protocols.<sup>12–14</sup>

Stochastic-based protocols, such as Birthday protocol,<sup>17</sup> set listening, transmitting, or sleeping in three states for sensor nodes. The probability principle of the birthday paradox guarantees that neighbor nodes can be found in a continuous time slot with high probability. However, the worst-case discovery latency cannot be guaranteed owing to the probabilistic characteristic of algorithm,<sup>18</sup> thereby causing a long tail on discovery probabilities over time.

Quorum-based protocols<sup>7–11</sup> address unbounded discovery latency issues by ensuring overlapping active durations between pairwise nodes within a bounded time. These schemes construct  $m \times m$  grid matrix within continuous slots, and each node casually selects one row and one column of entries from the matrix as an active slot for transmitting and receiving. No matter which row and column one node chooses, at least two common active slots are guaranteed with other nodes. The main drawback of quorum-based protocols is the global parameter of  $m$ , that is, all nodes in the network have to operate in a symmetric duty-cycle setting.<sup>19,20</sup>

Deterministic protocols have been recently proposed to support asymmetric duty-cycle patterns and handle the global parameter problem, as these schemes allow sensor nodes to select one or multiple prime numbers to present its duty cycle. Disco<sup>12</sup> presents a new neighbor discovery method based on the Chinese Remainder Theorem.<sup>21</sup> In this protocol, each node chooses a pair of dissimilar primes, such that the sum of their reciprocals is equal to the desired radio duty cycle. Disco achieves discovery faster than quorum-based protocols for a given duty cycle, allows nodes to independently select their own duty cycle, and offers a provable upper bound on discovery latency. U-Connect<sup>13</sup> proposes a unified solution to the neighbor discovery problem, and this approach addresses the symmetric and asymmetric asynchronous neighbor discovery problems. U-Connect also presents the power-latency product metric and uses this metric to measure the performance of the system.

Group-based neighbor discovery proposes a selective reference mechanism based on spatiotemporal

properties of neighborhood to significantly reduce discovery delay by proactively referring wake-up schedules among a group of nodes.<sup>15</sup> Nevertheless, only a small number of known schedules exist in the initialization phase of group formation. Nodes therefore cannot proactively wake up in time to verify the neighbor nodes, thereby leading to long discovery latency. An improved group-based neighbor discovery algorithm dynamically adjusts the active time of nodes by fully extracting the spatial properties, and it avoids unnecessary references.<sup>16</sup> The position distribution of nodes in some scenarios can be obtained or estimated in advance; thus, authors estimate the total number of nodes in a certain area to adjust the active time of these nodes.

Although neighbor discovery algorithms present good results and excellent performance, their energy efficiency can still be further improved and investigated.<sup>22,23</sup> Given that they are not making good performance trade-offs, EGD presents the design of relative distance between neighboring nodes from their RSS orderings, which can effectively reflect the distance between two nodes. This unique design effectively reduces the reference overhead of the referring nodes.

## Network model

We first define the network model and then generalize the assumptions on mobile sensor network such that nodes can discover each other without any infrastructure support in low-duty-cycle WSNs.

We assume that nodes are distributed with identical disk-shaped communication range. If a node is always in an active state, then the frequency transceiver of the node is in the state of listening. Supposing that the node does not receive any packets, the node can be in an idle listening.<sup>24,25</sup> In particular, no information exists for broadcasting or transmission. Idle listening consumes a large amount of unnecessary energy. Thus, a node remains dormant most of the time at more than 95% and be in active state only briefly to sense and communicate. During the entire discovery process, the working status of sensor nodes is constantly switching between the active and dormant states with low-energy consumption. In the active state, a node can broadcast a packet and receive packets sent by neighbors, while the node turns off the module of receiving packets except for transmitting a packet in the dormant state. This way can effectively reduce the energy consumption of nodes, thereby prolonging the lifetime of nodes and the network.<sup>26</sup> According to the Chinese Remainder Theorem, every node selects two different primes to solve the issue of two nodes of the same prime. Consequently, the actual duty cycle is equal to the sum of two reciprocal prime numbers. For convenience, we

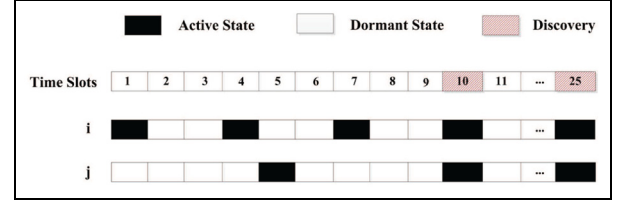


Figure 1. Network model.

assume that a node chooses one prime to present the duty cycle of this node. The maximal discovery latency between two nodes is equal to the product of two prime numbers chosen by the two nodes, and wake-up time slots of the two nodes will periodically or partially overlap. Thus, it ensures that any two nodes in a physical connection state can find each other in a certain time.

An example of a discovery process is shown in Figure 1. For arranging the working schedule of nodes, the working schedule of a mobile node is divided into continuous fixed length time slots in active or dormant state. The prime of node  $i$  is 3 and that of node  $j$  is 5, and their duty cycles are approximately 33% and 20%, respectively. When  $t = 1$ , node  $i$  is in the active state, whereas  $j$  is in the dormant state. Thus, they cannot find each other although they are in physical neighborhood. At time 10, nodes  $i$  and  $j$  wake up with partial overlap, if they are within the communication range of each other, they can discover each other and become aware of each other's working schedule.

The hypotheses of the network model are as follows:

1. All sensor nodes have a unique ID number and a local counter based on time slot  $\tau$  as unit.
2. Nodes are assumed to have the same transmission energy and channel frequency, and the communication coverage radius of the node is presumed consistent.
3. The switching energy consumption between different states is assumed to be small; thus, we can ignore the energy consumed by node state switching.
4. When nodes have common wake-up time slots and lie in each communication range, they can discover each other.

## System design

We introduce our detailed design for acceleration of neighbor discovery in mobile applications.

### Schedule reference mechanism of group-based discovery

For legacy pairwise discovery methods, neighbor discovery is passively achieved when two nodes are

pre-scheduled to wake up at the same time and two nodes are all within the communication range of each other. Pairwise discovery methods focus on a pair of nodes. Different from pairwise discovery methods, nodes in group-based discovery actively share the working schedule of their existing neighbors with new node that they have recently discovered. Consequently, the newly discovered node can quickly become aware of the time scheduling of its surrounding nodes and actively verify whether it can communicate with those nodes at their wake-up time.

The principle of the schedule reference mechanism is simple. In the initialization phase of group formation, no node has one neighbor and knows the working schedule of its neighbor. Similar to pairwise discovery algorithms, the node is periodically switched between the active and dormant states according to its own working schedule. In the active state, sensor nodes can broadcast and receive messages sent by neighbors; meanwhile, the sensor nodes will turn off the broadcasting and receiving functions in the dormant state. When two nodes are in the active state and within the communication range of each other, they can discover each other and form a group. For example, when the third and first nodes wake up simultaneously according to their working schedule, the third node finds the first node and adds the first node to its neighbor table. Then, the first node will share its known neighbors (i.e. the second node) with the third node. The third node has been informed of the working schedule of the second node through the first node. Accordingly, the wake-up time of the third and second nodes becomes synchronous. In this condition, the third node proactively wakes up to determine whether the second node is the neighbor node. If yes, then they add each other to the neighbor table and form a larger group than before. This group of neighbor nodes then constantly expands, thereby completing the discovery of neighbor nodes.

Figure 2 shows the group-based discovery process. We consider sensor network with  $n$  mobile sensor

nodes and take nodes  $i, j$ , and  $k$  as examples, where the node  $k$  is the discovery node. The three nodes are scheduled according to their own work arrangement. The group-based discovery process for nodes  $i, j$ , and  $k$  is as follows:

1. At time  $t = 1$ , node  $j$  and node  $i$  discover each other and become aware of other's working schedule. They add each other to its neighbor table and form a virtual group.
2. At time  $t = 4$ , nodes  $k$  and  $j$  simultaneously wake up and become neighbors. After node  $j$  discovers nodes  $i$  and  $k$ , node  $j$  proactively wakes up at time 11 when node  $k$  is scheduled to be active and notifies the working schedule of its neighbor nodes (including node  $i$ ) to node  $k$ . According to the scheduling information of node  $i$ , node  $k$  will proactively wake up to find node  $i$ .
3. At time 12, node  $i$  is in the wake-up state and node  $k$  proactively wakes up and identifies whether node  $i$  is its neighboring node. After time 12, node  $k$  has discovered nodes  $j$  and  $i$ . With regard to pairwise (such as Disco<sup>12</sup>), node  $k$  cannot discover nodes  $j$  and  $i$  until time 67. Compared with pairwise methods, group-based discovery significantly reduces the discovery delay.

Schedule reference mechanism speeds up the discovery process by proactively referring working schedules among a group of nodes. The acceleration is mainly implemented from a temporal perspective. Node  $j$  announces all its discovered neighbor information to node  $k$  that it has just discovered. In this way, the addition of node  $k$  into the group of node  $j$  will have a complete picture of nodes in the surrounding area. However, this design would still waste energy unnecessarily because not all known neighbors of node  $j$  are neighbors to node  $k$ . In the following, we introduce the recommendation mechanism based on relative distance,

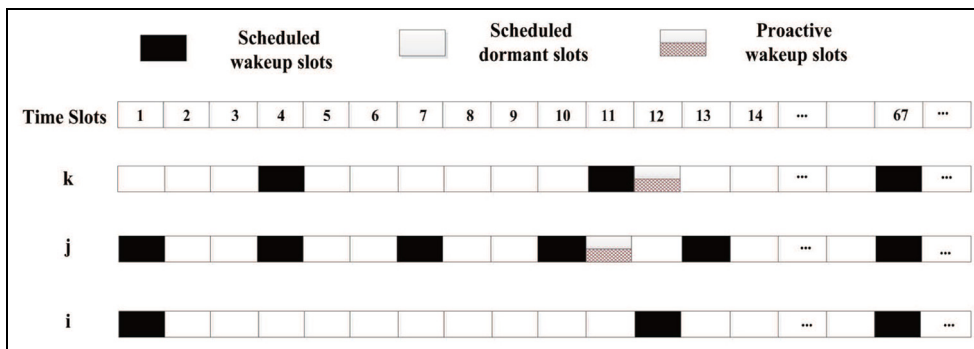
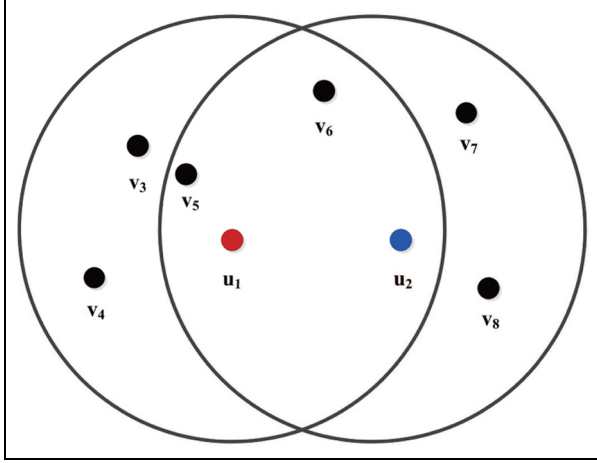


Figure 2. Group-based discovery process.

**Table 1.** RSS values from all their neighbors.

Obj \ Nei	$u_1$	$u_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$
$u_1$	0	-57	-76	-81	-54	-58	Inf	Inf
$u_2$	-63	0	Inf	Inf	-73	-55	-68	-66

**Figure 3.** Example of evaluating the distance between two nodes.

which can successfully reduce invalid recommended nodes.

### Design: relative distance estimation

We present the main idea of calculating relative distance among neighboring nodes. Many existing methods only directly use RSS to estimate physical distance in several scenarios. Without loss of generality, we can evaluate the distance between objects  $u_1$  and  $u_2$  using the data set  $S = (U, V, R)$ , where  $U = \{u_1, u_2\}$ ,  $u_1$  and  $u_2$  are the target objects for distance quantization;  $V = \{v_1(u_1), v_2(u_2), v_3, \dots, v_N\}$ , where  $V$  is a set that comprises  $u_1$ ,  $u_2$  and their neighbors. A total of  $C_N^2$  neighbor pairs are obtained for nodes  $u_1$  and  $u_2$ ;  $R$  is  $2 \times N$  matrix, and  $R_{ij}$  represents the RSS from node  $v_j$  for node  $u_i$ .

A simple example is shown in Figure 3 for ease of expression. In this example, the problem is to evaluate the distance between objects  $u_1$  and  $u_2$ . Table 1 lists the RSS values from all their neighbors. If  $R_{ij}$  is set to be Inf, then node  $u_i$  cannot receive the signal from node  $v_j$ . And we set  $R_{11}$  and  $R_{22}$  to be zero, which means that the node  $u_1$  and  $u_2$  have the strongest signal from themselves.

The neighbor pairs of objects  $u_1$  and  $u_2$  can be categorized through the following method. By comparing the RSS values,  $u_1$  can determine whether it is closer to  $v_5$  than  $v_6$ . We define  $v_5$  and  $v_6$  as an effective neighbor pair and use  $N_{eff}$  to denote the number of effective neighbor pairs. By contrast,  $u_1$  cannot determine

whether it is closer to  $v_7$  than  $v_8$  because it cannot receive signals from the two nodes. Thus, we call  $v_7$  and  $v_8$  as an ineffective neighbor pair. Similarly,  $u_2$  cannot determine whether it is closer to  $v_3$  than  $v_4$ . It is easy to determine that,  $v_3$  and  $v_4$  are also an ineffective neighbor pair. If  $N_{inv}$  is used to represent the number of ineffective neighbor pairs. The sum of  $N_{eff}$  and  $N_{inv}$  should be equal to  $C_N^2$ . The distance evaluation between objects  $u_1$  and  $u_2$  is actually based on the  $N_{eff}$  effective neighbor pairs. Any effective neighbor pair  $v_j$  and  $v_k$  are a concordant pair if  $R_{1j} > R_{1k}$  and  $R_{2j} > R_{2k}$ . In contrast, if  $R_{1j} > R_{1k}$  and  $R_{2j} < R_{2k}$  (or  $R_{1j} < R_{1k}$  and  $R_{2j} > R_{2k}$ ), then  $v_j$  and  $v_k$  are defined as a discordant pair. In the given example, nodes  $v_5$  and  $v_7$  is a discordant pair, while  $v_3$  and  $v_5$  is a concordant pair. In this paper, we use symbols  $N_{con}$  and  $N_{dis}$  to represent the number of concordant pairs and discordant pairs, respectively.

On the basis of the above-mentioned classification and explanation, we can evaluate and quantify the relative distance between objects  $u_1$  and  $u_2$  as follows.

The relative distance between two neighboring nodes is equal to the ratio between the number of discordant pairs and the number of effective neighbor pairs, that is

$$d_{seq}(u, v) = \frac{N_{dis}}{N_{eff}} \quad (1)$$

where  $N_{dis} = N_{eff} - N_{con}$ , and  $N_{eff} = C_N^2 - N_{ine}$ . The value of  $d_{seq}(u, v)$  is the relative distance of neighboring objects  $u_1$  and  $u_2$ ,  $d_{seq}(u, v)$  is equivalent to physical distance; thus,  $d_{seq}(u, v)$  can be used as a relative distance for the case of sensor nodes under unevenly distributed network.

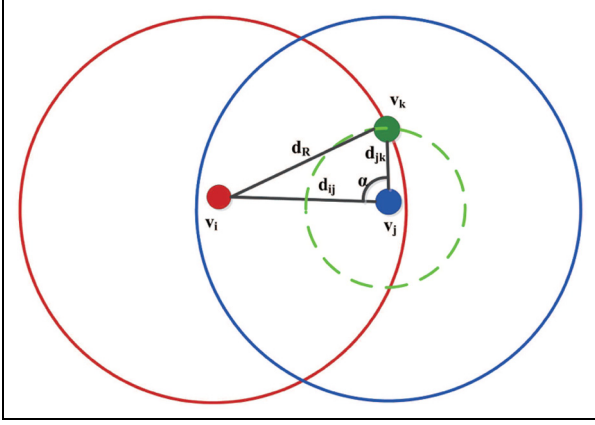
Taking objects  $u_1$  and  $u_2$  in Figure 3 as an example, we can obtain 10 discordant pairs ( $(v_1, v_2)$ ,  $(v_1, v_6)$ ,  $(v_2, v_5)$ ,  $(v_3, v_7)$ ,  $(v_3, v_8)$ ,  $(v_4, v_7)$ ,  $(v_4, v_8)$ ,  $(v_5, v_6)$ ,  $(v_5, v_7)$ , and  $(v_5, v_8)$ ) and two ineffective neighboring pairs ( $(v_3, v_4)$  and  $(v_7, v_8)$ ). Thus, the relative distance between objects  $u_1$  and  $u_2$  is  $10/24 = 0.418$ .

### Neighbor node recommendation using relative distance

We introduce an algorithm combining the above method and relative distance for computing the probability that the neighboring nodes  $v_i$  and  $v_k$  of node  $v_j$  are within the communication range of each other.

From equation (1), we obtain the relative distance  $d_{jk}$  between recommending node  $v_j$  and referring node





**Figure 4.** Neighbor probability of nodes  $v_i$  and  $v_k$ .

$v_k$  and  $d_{ij}$  between node  $v_j$  and recommended node  $v_i$ . When  $d_{jk}$  is fixed, then node  $v_k$  is located on the dashed circle. In calculating the probability of two neighbors  $v_i$  and  $v_k$  of node  $v_j$ , we assume that the relative distance of node communication radius is 1. Figure 4 shows that, when node  $v_k$  falls in the overlapping area between nodes  $v_i$  and  $v_j$ , node  $v_k$  is the common neighbor of nodes  $v_i$  and  $v_j$ . If  $d_{ij} + d_{jk} \leq 1$ , then node  $v_k$  must be in the overlapped region of nodes  $v_i$  and  $v_j$ . Therefore, the probability that node  $v_k$  is the common neighbor of nodes  $v_i$  and  $v_j$  is 100%. When  $d_{ij} + d_{jk} > 1$ , node  $v_k$  is also the common neighbor of two nodes, if and only if node  $v_k$  is located at the dashed circle inside  $v_i$ . On the basis of cosine theorem, the angle  $\alpha$  in Figure 4 can be represented as

$$\alpha = \arccos\left(\frac{d_{ij}^2 + d_{jk}^2 - 1}{2d_{ij}d_{jk}}\right) \quad (2)$$

The neighbor probability of nodes  $v_i$  and  $v_k$  can be expressed as

$$\frac{2\alpha}{2\pi} = \frac{1}{\pi} \arccos\left(\frac{d_{ij}^2 + d_{jk}^2 - 1}{2d_{ij}d_{jk}}\right) \quad (3)$$

Combining the above two assumptions,<sup>15</sup> we can obtain the formula to present the probability of the two neighbor nodes  $v_i$  and  $v_k$  of node  $v_j$  are also neighbor as

$$P_{j,ik}(d_{ij}, d_{jk}) = \begin{cases} \frac{1}{\pi} \arccos\left(\frac{d_{ij}^2 + d_{jk}^2 - 1}{2d_{ij}d_{jk}}\right), & d_{ij} + d_{jk} > 1 \\ 1, & d_{ij} + d_{jk} \leq 1 \end{cases} \quad (4)$$

Our work presents a different design to calculate neighbor probability. By setting different threshold values for  $P_{j,ik}$ , a node can selectively reference its neighboring nodes to a newly discovered node; therefore, it can achieve a better trade-off between reducing discovery

delay and energy consumption.<sup>27</sup> The advantage of this design is that estimating distance between two mobile nodes with relative distance can provide heuristic information about which neighboring node is closer and which is further.

### Qualitative comparison: EGD versus Disco

EGD is a protocol based on schedule reference mechanism, in which nodes actively share the working schedule of their neighbors. In this scenario, upon discovering a new node, members of the group would gossip about their own neighbors with the new node, which would help the new node discover its neighbors more quickly. In other words, EGD usually has shorter discovery latency than Disco. In the following, we quantitatively compare the difference of discovery delay between EGD and Disco.

For Disco, it ensures that a pair of neighboring nodes  $i$  and  $j$  discovering each other within a bounded time  $T_{i,j}$ . Therefore, if  $t \geq T_{i,j}$  and they are within in each other's communication range, nodes  $i$  and  $j$  can find each other with the 100% probability. If  $t < T_{i,j}$ , we assume that discovery probability of two nodes still discover each other with the function  $f(i, j, t)$ . The class of neighbor discovery probability  $P_{i,j}(t)$  is given by<sup>15</sup>

$$P_{i,j}(t) = \begin{cases} 1, & t \geq T_{i,j} \\ f(i, j, t), & t \in [0, T_{i,j}) \end{cases} \quad (5)$$

Then for Disco, the probability distribution function for node  $i$  discovering all  $n$  nodes in this scenario before time  $t$  can be presented as  $P_d^i(t) = \prod_{j=0}^{n-1} P_{i,j}(t)$ . Consequently, the corresponding probability density function is  $[P_d^i(t)]'$ . The expected time for node  $i$  discovering all  $n$  nodes is given by

$$t_d^i = E_d^i(t) = \int_0^{T_{\max}^i} t [P_d^i(t)]' dt \quad (6)$$

where  $T_{\max}^i = \text{Max}(T_{i,0}, T_{i,1}, \dots, T_{i,n-1})$ .

For EGD, the probability distribution function of a new node  $i$  discovers one of the nodes in the group before time  $t$  is  $P_e^i(t) = 1 - \prod_{j=0}^{n-1} [1 - P_{i,j}(t)]$ . The corresponding expected time for node  $i$  discovering at least one node in the group is, therefore, given by

$$E_e^i(t) = \int_0^{T_{\min}^i} t [P_e^i(t)]' dt \quad (7)$$

where  $T_{\min}^i = \text{Min}(T_{i,0}, T_{i,1}, \dots, T_{i,n-1})$ .

After node  $i$  discovers node  $j$  which would share working schedules of all nodes in its group with node  $i$ . Node  $i$  proactively wakes up at the wake-up state of non-discovered nodes and verifies whether it can

**Table 2.** Parameters of simulation environment.

Parameter	Value
Number of nodes	78
Grid length	500 m
Grid width	500 m
Initial duty cycle	1%
Communication	100 m
Node speed	1 m/s

communicate with those nodes in the group. The maximal time  $t_e^i$  for node  $i$  discovering all  $n$  nodes in the group is at most  $E_e^i(t) + 2\text{Max}(T_{\text{gap}})$ .  $T_{\text{gap}}$  is the time gap between two consecutive wake-ups of a node pair. After the first discovery, in the worst case, a node takes another two  $\text{Max}(T_{\text{gap}})$  delay to finish reference and verification.

## Evaluation

### Simulation setup

In the simulation, mobile nodes are distributed in an area of  $500\text{m} \times 500\text{m}$ , and the communication radius of the sensor node is 100 m. For the purpose of investigating the impact of node velocity on system performances, node velocity changes from 1 to 9 m/s. When exploring the number of nodes, the total number of nodes changes from 26 to 143. In order to evaluate the impact of different radio irregularity of node, the value of that varies from 0 to 0.9. Unless otherwise mentioned, Table 2 lists the default simulation configurations for the following sections.

### Performance evaluation

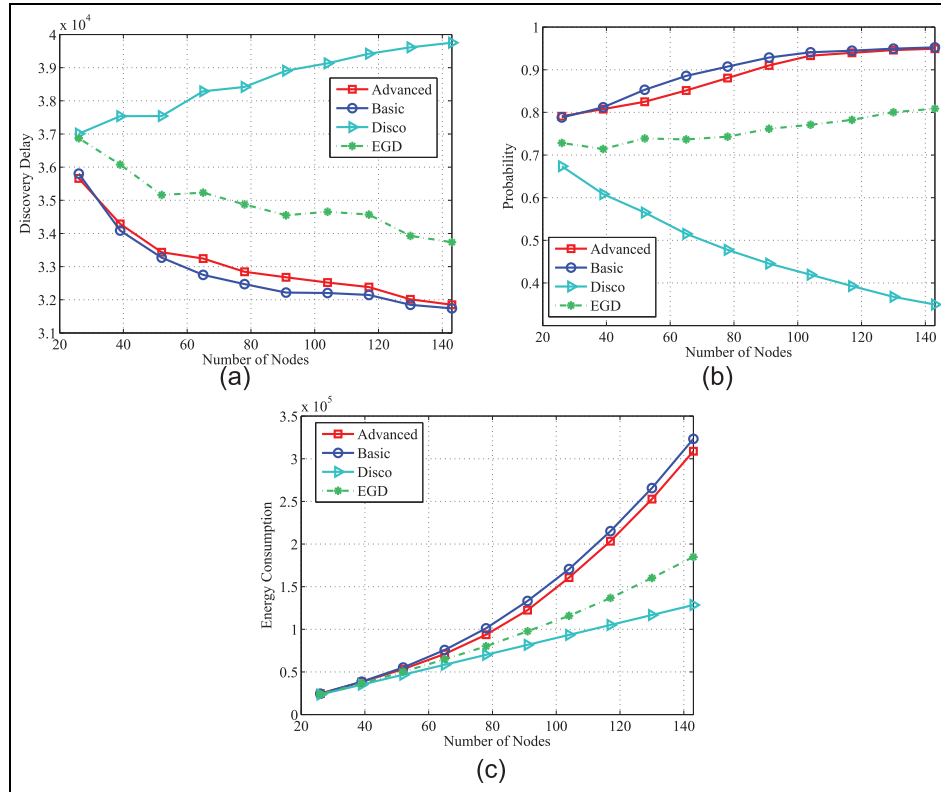
To evaluate the system performance under different number of nodes, node velocities, and radio irregularities, we mainly use three metrics, namely, discovery probability, discovery delay, and energy consumption.

**Effect of the number of nodes on system performance.** The first set of simulation aims to study the effect of the number of sensor nodes on the system performance. Figure 5(a) shows that, as the number of nodes increases, the discovery delay of group-based discovery and EGD decreases whereas that of Disco increases. The reason is that group-based discovery and EGD profit from the increase in number of nodes as a large number of working schedules are shared among neighbor nodes. For the same reason, the discovery probability of group-based discovery and EGD steadily increases whereas that of Disco decreases as number of nodes increases (see Figure 5(b)). As shown in Figure 5(c), when number of nodes increases, the energy

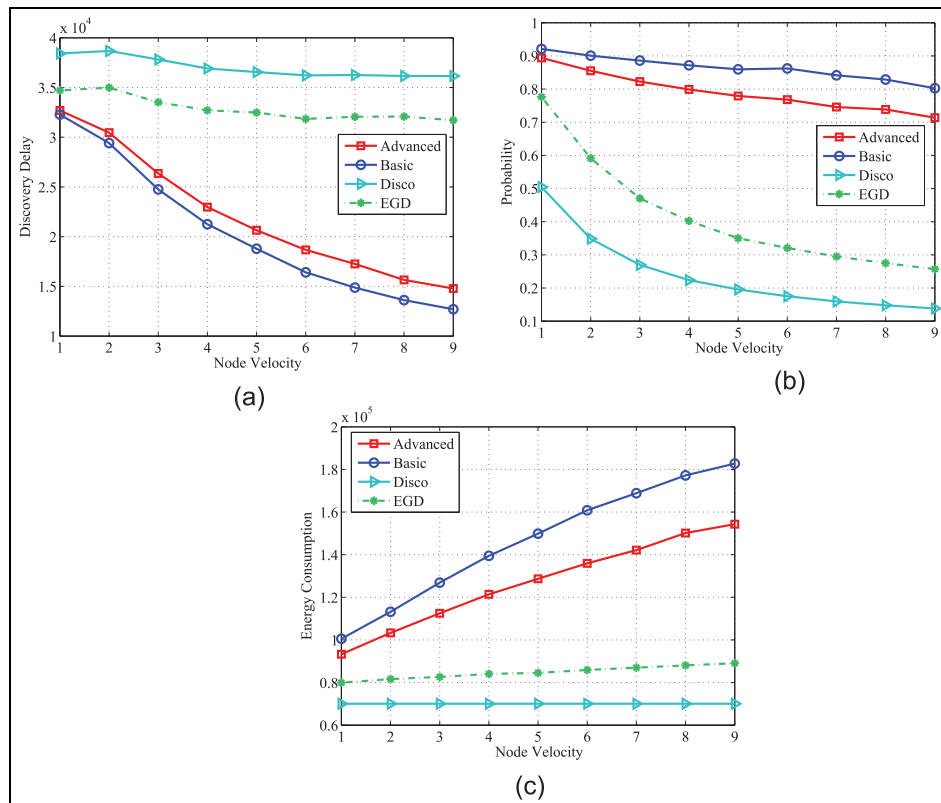
consumption of the three methods increases. More nodes means a large number of nodes proactively wake up to verify the recommended nodes. The trend of the numerical value reflects that the energy consumption of EGD is obviously lower than that of group-based discovery but significantly higher than that of Disco. Nevertheless, EGD can save more energy than group-based discovery under the same number of nodes. For example, with 91 nodes, EGD obtains a 5.72% higher discovery delay with a 20.25% lower energy consumption compared with advanced group design.

**Effect of average node velocity on system performance.** Velocity of nodes has significant effects on the system performance in mobile sensor networks.<sup>28</sup> Hence, the second simulation attempts to investigate the influence of the average velocity on the system performance. Figure 6(a) and (b) shows that the discovery probability and discovery delay of the three designs decrease as node velocity increases. This result is attributed to the duration that two nodes are physically within the communication range of each other and shortens when node velocity increases. Consequently, the discovery delay of the three designs decreases. With the increase in velocity, the number of common neighbor nodes of EGD changes faster than that of group-based discovery. Therefore, the discovery delay of EGD is significantly higher than that of group-based discovery. Figure 6(c) shows that the energy consumption of the three algorithms increases with the increase in node velocity. The energy loss of Disco and EGD design remains stable whereas that of group-based discovery significantly increases with increasing node velocity. This result is due to the fact that EGD discards considerable useless reference information as nodes move far from the group. At about 2 m/s, the discovery delay for EGD is about 1.15 times higher than advanced group discovery, while the discovery probability is about 30.84% shorter. Additionally, the energy consumption for EGD is about 20.98% shorter than advanced group discovery. In general, EGD achieves good energy-delay trade-off.

**Effect of radio irregularity on system performance.** For convenience of description, we assume that the communication radius of node is a unit disk. We then study the effect of radio range irregularity on the system performance.<sup>29</sup> In the simulation, the radio range irregularity of node broadcast changes from 0% to 90%. Radio range irregularity is defined as the maximum range variation per unit degree change in the direction of radio propagation. When radio range irregularity is 90%, the communication range of a node can belong to any value between 10 and 100 m.

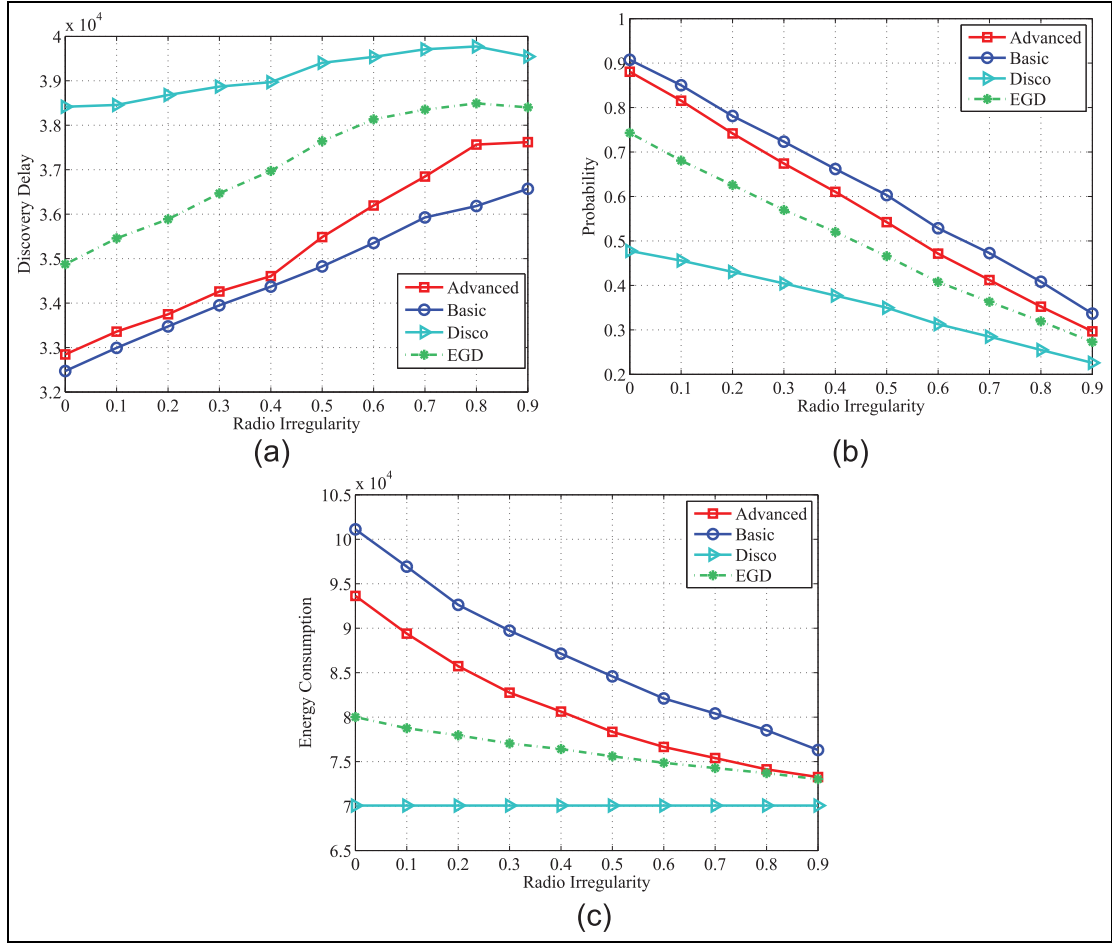


**Figure 5.** Effect of the number of nodes: (a) number of nodes versus discovery delay, (b) number of nodes versus probability, and (c) number of nodes versus energy consumption.



**Figure 6.** Effect of average node velocities: (a) node velocity versus discovery delay, (b) node velocity versus probability, and (c) node velocity versus energy consumption.





**Figure 7.** Effect of radio range irregularity: (a) radio irregularity versus discovery delay, (b) radio irregularity versus probability, and (c) radio irregularity versus energy consumption.

Figure 7(a) shows that the discovery delay of all the three schemes increases with the increase in degree of radio irregularity. However, the discovery delay of Disco is higher than that of group-based discovery and EGD. The reason is that, when a referring message in group-based discovery and EGD is lost, other referring messages informed by neighbor nodes in a short period still exist. However, for Disco, nodes have to wait for the next wake-up slots to discover neighbor nodes.

As shown in Figure 7(b), the discovery probability of the three discovery methods decreases as the degree of radio irregularity increases. This result is due to the fact that such increase decreases the accuracy of neighbor estimation using a fixed communication radio between two nodes. Regarding energy consumption, Figure 7(c) shows that the energy consumption of group-based discovery and EGD decreases whereas that of Disco remains stable when the degree of radio irregularity increases. Compared with group-based discovery, EGD has a few redundant reference messages; hence, its energy loss is significantly lower than that of group-based discovery. The simulation results prove

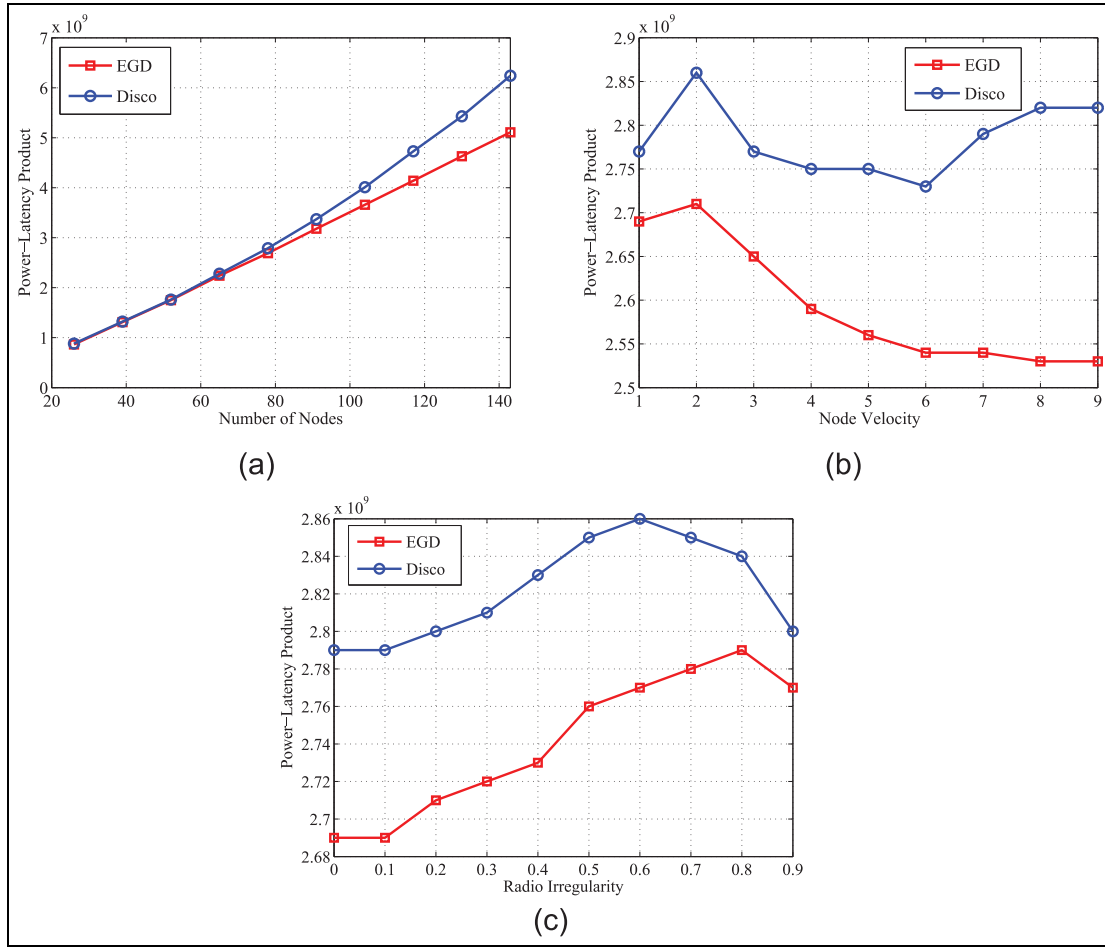
that EGD can achieve good results in reducing energy consumption while improving discovery probability.

**Evaluation of designs.** Energy efficiency and low discovery latency are the key metrics involved in evaluating neighbor discovery protocols in the context of mobile sensor networks. In order to perform a holistic evaluation of EGD and Disco, we consider the power-latency ( $PL$ ) product  $I$  as the performance metric.<sup>13</sup> The  $PL$  product  $I$  is

$$I = PL \quad (8)$$

The power-latency product has been widely adopted to evaluate many neighbor discovery algorithms like Quorum,<sup>7</sup> Disco,<sup>12</sup> and U-connect.<sup>13</sup> Similar to these classical works, we adopt the above metric to evaluate the performance.

Figure 8(a)–(c) shows that EGD improves the system performance with small overhead compared with Disco. When the number of nodes is 104, the value of  $I$  of EGD has been declined 8.90% with a 12.92%

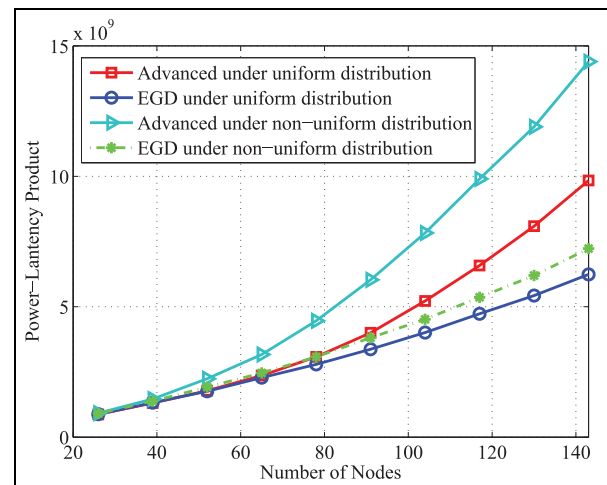


**Figure 8.** Comparison with Disco: (a) number of nodes versus power-latency product, (b) node velocity versus power-latency product, and (c) radio irregularity versus power-latency product.

shorter discovery delay at the cost of 19.32% more energy consumption. This result is attributed to the fact that EGD has reference mechanism, more neighbor information is shared, the network consumes more energy, and costs less time.

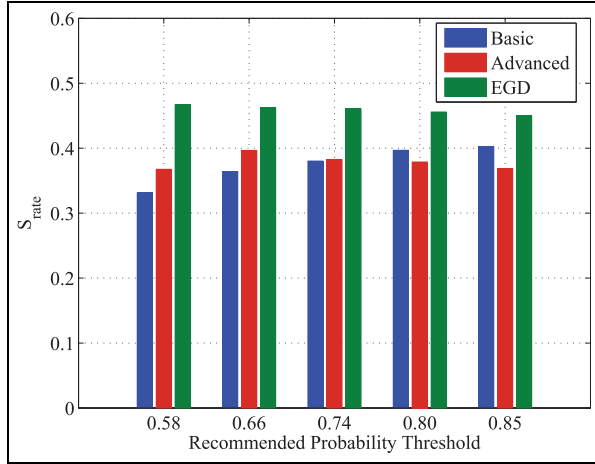
**Evaluation of node distribution.** In order to investigate the impact of different node distributions, in Figure 9, we show the system performance under both uniform node distribution (random waypoint mobility model) and nonuniform node distribution (hotspot mobility model). For hotspot mobility model, it mimics the behavior of people moving from a point of interest to another and stopping for a certain time at each point of interest before moving to the next one.

**Evaluation of recommended probability thresholds.** EGD and advanced group discovery accelerate the discovery process by proactively referring working schedules among a group of nodes. We study the recommended success rate of these nodes under different recommended



**Figure 9.** Impact of node distribution on system performance.

probability thresholds. Threshold reflects how many neighbors are required to be recommended. A high value of threshold indicates a large number of neighbor



**Figure 10.** Recommended success rate under different thresholds.

nodes need to be recommended. However, this condition does not mean the recommended success rate is high, as not all recommended nodes are neighbors of referring nodes. Thus, studying the recommended success rate in group discovery algorithms is necessary.

In low-power mobile sensor networks,<sup>30</sup> one node recommends all its neighbor nodes to the newly found neighbor node; however, recommended nodes are not all neighbors to the new node. Therefore, we use the recommended success rate  $S_{rate}$  weights of various designs

$$S_{rate} = \frac{T_S}{T_S + T_F} \quad (9)$$

where  $T_S$  is the sum of all sensor nodes successfully recommending nodes to their newly discovered neighbor node, and  $T_F$  is the total number of recommending failures. Figure 10 shows that the success rate of advanced group design and EGD slightly decreases whereas that of basic group design increases with the increase in threshold. However, the recommended success rate of EGD and advanced group design is higher than that of basic group design under different recommended probability thresholds. The main reason is that advanced group design and EGD leverage temporal selection to filter the neighbor nodes of the recommending nodes and spatial selection,<sup>15</sup> such that these recommended nodes are abandoned with a small recommended probability. EGD combines the spatiotemporal properties with the signal strength of sensor node to estimate the relative distance; therefore, the active time of nodes can be adjusted in time. For example, at a threshold of 0.66, EGD obtains 21.31% and 14.26% higher recommended success rate compared with basic group and advanced group. We can conclude that EGD recommends neighbors more accurate than group-based discovery.

## Conclusion

In this article, we introduce EGD, a novel scheme based on group discovery schemes for the acceleration of neighbor discovery process. The key insight into accelerating the discovery process is that known neighbors can help a node learn unknown neighbors indirectly by schedule reference mechanism. To avoid energy overhead caused by redundant reference messages, neighbor probability calculated by relative distance enables nodes to adjust their own duty cycle with more accurate value. We design the relative distance on the basis of RSS ordering instead of directly mapping the RSS values to distances. Numerous simulations proved that EGD method improves recommended accuracy and further increases energy efficiency.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported, in part, by the Fundamental Research Funds for the Central Universities (Grant No. 2017XKQY080).

## References

1. Ouafaa I, Mustapha E, Salah-Ddine K, et al. An advanced analysis on secure hierarchical routing protocols in wireless sensor network. In: *Proceedings of the international conference on engineering and MIS (ICE-MIS)*, Agadir, Morocco, 22–24 September 2016, pp.1–6. New York: IEEE.
2. Min KT, Forys A and Schmid T. Demonstration abstract: airfeed: indoor real time interactive air quality monitoring system. In: *Proceedings of the international symposium on information processing in sensor networks*, Berlin, 15–17 April 2014, pp.325–326. New York: IEEE.
3. Zhong Z and He T. Achieving range-free localization beyond connectivity. In: *Proceedings of the international conference on embedded networked sensor systems*, Berkeley, CA, 4–6 November 2009, pp.281–294. New York: ACM.
4. Vasudevan S, Towsley D, Goeckel D, et al. Neighbor discovery in wireless networks and the coupon collector's problem. In: *Proceedings of the annual international conference on mobile computing and networking (MOBI-COM)*, Beijing, China, 20–25 September 2009, pp.181–192. New York: ACM.
5. Mcglynn MJ and Borbash SA. Birthday protocols for low energy deployment and flexible neighbor discovery in ad hoc wireless networks. In: *Proceedings of the ACM international symposium on mobile ad hoc networking and*

- computing (*MOBIHOC*), Long Beach, CA, 4–5 October 2001, pp.137–145. New York: ACM.
6. Borbash SA, Ephremides A and Mcglynn MJ. An asynchronous neighbor discovery algorithm for wireless sensor networks. *Ad Hoc Netw* 2007; 5(7): 998–1016.
  7. Tseng YC, Hsu CS and Hsieh TY. Power-saving protocols for IEEE 802.11-based multi-hop ad hoc networks. *Comput Netw* 2003; 43(3): 317–337.
  8. Zheng R, Hou JC and Sha L. Asynchronous wakeup for ad hoc networks. In: *Proceedings of the ACM international symposium on mobile ad hoc networking and computing (MOBIHOC)*, Annapolis, MD, 1–3 June 2003, pp.35–45. New York: ACM.
  9. Lai S, Zhang B, Ravindran B, et al. Cqs-pair: cyclic quorum system pair for wakeup scheduling in wireless sensor networks. In: *Proceedings of the international conference on principles of distributed systems (OPODIS)*, Luxor, Egypt, 15–18 December 2008, pp.295–310. Berlin: Springer Verlag.
  10. Lai S, Ravindran B and Cho H. Heterogenous quorum-based wake-up scheduling in wireless sensor networks. *IEEE T Comput* 2010; 59(11): 1562–1575.
  11. Link JgB, Wollgarten C, Schupp S, et al. Perfect difference sets for neighbor discovery: energy efficient and fair. In: *Proceedings of the extreme conference on communication: the amazon expedition (ExtremeCom)*, Manaus, Brazil, 26–30 September 2011, pp.342–349. New York: ACM.
  12. Dutta P and Culler D. Practical asynchronous neighbor discovery and rendezvous for mobile sensing applications. In: *Proceedings of the ACM conference on embedded networked sensor systems (SenSys)*, Raleigh, NC, 5–7 November 2008, pp.71–83. New York: ACM.
  13. Kandhalu A, Lakshmanan K and Rajkumar R. U-connect: a low-latency energy-efficient asynchronous neighbor discovery protocol. In: *Proceedings of the international conference on information processing in sensor networks (IPSN)*, Stockholm, 12–16 April 2010, pp.350–361. New York: ACM.
  14. Purohit A, Priyantha B and Liu J. Wiflock: collaborative group discovery and maintenance in mobile sensor networks. In: *Proceedings of the international conference on information processing in sensor networks (IPSN)*, Chicago, IL, 12–14 April 2011, pp.37–48. New York: IEEE.
  15. Chen L, Shu Y, Gu Y, et al. Group-based neighbor discovery in low-duty-cycle mobile sensor networks. *IEEE T Mob Comput* 2016; 15(8): 1996–2009.
  16. Niu Q, Bao W and Xia S. An improved group-based neighbor discovery algorithm for mobile sensor networks. *Int J Distrib Sens N* 2014; 2014(2): 1–7.
  17. Qi Z and Ping H. The topology construction power control algorithm based on birthday paradox. In: *Proceedings of the international conference on computer sciences and applications (CSA)*, Wuhan, China, 14–15 December 2013, pp.223–228. New York: IEEE.
  18. Kandhalu A, Xhafa AE and Hosur S. Towards bounded-latency bluetooth low energy for in-vehicle network cable replacement. In: *Proceedings of the international conference on connected vehicles and expo (ICCVEx)*, Las Vegas, NV, 2–6 December 2013, pp.635–640. New York: IEEE.
  19. Lupu AT, Benisty H and Lavrinenko AV. Tailoring spectral properties of binary PT-symmetric gratings by duty-cycle methods. *IEEE J Sel Topic Quantum Electron* 2016; 22(5): 35–41.
  20. Al-Mhana T, Zahawi B and Pickert V. Symmetrical duty cycle control for FCSC converter for wave energy applications. In: *Proceedings of the international symposium on communication systems, networks and digital signal processing (CSNDSP)*, Manchester, 23–25 July 2014, pp.56–60. New York: IEEE.
  21. Hardy GH and Wright EM. *An introduction to the theory of numbers*. London: Institute of Physics, 1979.
  22. Lin K, Lai CF, Liu X, et al. Energy efficiency routing with node compromised resistance in wireless sensor networks. *Mobile Netw Appl* 2012; 17(1): 75–89.
  23. Qiu Y, Li S, Xu X, et al. Talk more listen less: energy-efficient neighbor discovery in wireless sensor networks. In: *Proceedings of the IEEE international conference on computer communications (INFOCOM)*, San Francisco, CA, 10–14 April 2016, pp.1–9. New York: IEEE.
  24. Polastre J, Hill J and Culler D. Versatile low power media access for wireless sensor networks. In: *Proceedings of the international conference on embedded networked sensor systems*, Baltimore, MD, 3–5 November 2004, pp.95–107. New York: ACM.
  25. Kohvakka M, Suhonen J, Kuorilehto M, et al. Energy-efficient neighbor discovery protocol for mobile wireless sensor networks. *Ad Hoc Netw* 2009; 7(1): 24–41.
  26. Zhang L, Chen S, Jian Y, et al. Maximizing lifetime vector in wireless sensor networks. *IEEE ACM T Netw* 2013; 21(4): 1187–1200.
  27. Wang Y, Vuran MC and Goddard S. Stochastic performance trade-offs in the design of real-time wireless sensor networks. In: *Proceedings of the international conference on computing, networking and communications (ICNC)*, Garden Grove, CA, 16–19 February 2015, pp.931–937. New York: IEEE.
  28. Meghanathan N, Thomas D and Addison ES. Multicast extensions to the flow-oriented routing protocol and node velocity-based stable path routing protocol for mobile ad hoc networks. In: *Proceedings of the international conference on ultra modern telecommunications and workshops*, Saint Petersburg, Russia, 12–14 October 2009, pp.1–6. New York: IEEE.
  29. Zhou G, He T, Krishnamurthy S, et al. Impact of radio irregularity on wireless sensor networks. In: *Proceedings of the international conference on mobile systems, applications and services*, Boston, MA, 6–9 June 2004, pp.125–138. New York: ACM.
  30. Sharma A and Tewolde GS. Considerations in low power wireless sensor networks. In: *Proceedings of the IEEE international conference on electro/information technology (EIT)*, Dekalb, IL, 21–23 May 2015, pp.626–631. New York: IEEE.