Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6654427, 12 pages
https://doi.org/10.1155/2021/6654427

Research Article

Energy Balanced Source Location Privacy Scheme Using Multibranch Path in WSNs for IoT

Huijiao Wang, Lin Wu, Qing Zhao, Yongzhuang Wei, and Hua Jiang

Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China

Correspondence should be addressed to Huijiao Wang; whj@guet.edu.cn

Received 23 December 2020; Revised 3 February 2021; Accepted 12 February 2021; Published 25 February 2021

Academic Editor: Chien-Ming Chen

Source location privacy, one of the core contents of Wireless Sensor Network (WSN) security, has a significant impact on extensive application of WSNs. In this paper, a novel location privacy protection routing scheme called Energy Balanced Branch Tree (EBBT) is proposed by using multibranch and fake sources. This scheme has three phases. In the first place, the data of the source are randomly sent to a certain intermediate node. Then, a minimum hop routing (MHR) from the intermediate node to the base station is formed. Then, branch paths with fake sources are generated dynamically from some nodes on the MHR path. Finally, a tree-shaped structure from real source nodes and fake source nodes to the base station is achieved. In difference to the previous schemes, the location of the real source in the EBBT scheme does not affect the location and the number of fake sources. During the formation of the tree-shaped multibranch paths, the residual energy of nodes is considered sufficiently, and the control of the direction of each branch path is also involved. The influence of the number and length of branches on the network lifetime and network security is also investigated. Experimental results show that the proposed algorithm has the advantages of long network security period and lifetime, as well as high path diversity. Our simulation further illustrates that the EBBT scheme has favorable privacy of the source location without changing the network lifetime.

1. Introduction

WSNs (Wireless Sensor Networks) as the main component of Internet of Things (IoT) applications have been used widely in monitoring field [1, 2], intelligent agriculture [3], medical health [4, 5], and so on. With the continuous expansion of WSN application scenarios, its securities, such as the authentication and encryption mechanism [6, 7], have attracted extensive attentions. The security and privacy of WSNs are usually affected by the wireless communication media, dynamic topology, and limited resources. Specifically, wireless communication media used in WSNs is easily monitored, intercepted, or modified by an attacker. An attacker may infer the location information by listening to the transmission of network data and analyzing the network traffic [8].

In the monitoring events or object tracking applications, networks consist of some monitoring sensor nodes. The node that senses the monitored object is called the source node or source location. In general, the source node is responsible for collecting and sending data to the sink node. Considering that physical locations of source nodes are closely related to the identity information of monitored objects, the location information of these nodes is important to the whole system. For instance, in the pander-hunter game model [9], the concept of the source location privacy (SLP) is presented. After sensing the panda, the source node sends the collected data to the base station periodically. The hunter acts as an attacker, and the hunter tries to catch the panda by backtracking. Keeping the location of the source node from being discovered by attackers is the goal of SLP.

The researches on SLP protection mechanism put forward different strategies to cope with different attacks. Random walk is one of the most classical methods to resist the hop-by-hop backtracking attacks. And the fake packet injection is another well-known method against backtracking attacks. This method embeds fake traffic into real data making it difficult for attackers to distinguish real traffic from fake data and thus guarantees the security of source location [10]. Due to the reason that the fake source is generated statically, attackers can easily find the location of the fake source and thus discard it to

find the real one [11]. The technology of multibranch path for forwarding fake packet would make the opponent deviate from the delivery path of the real packets and thus protect the source location when the attacker traces the source through hop-by-hop backtracking packets. This source location privacy protection can be implemented by building a tree structure route [12], where the backbone route and the transfer route are established. As the number of branches or false data traffic increases, the intensity of privacy protection will increase. Moreover, the energy consumption of network is also higher than before with the increase of branches.

The nodes that have balanced energy consumption in WSNs can greatly improve the lifetime of network. On the contrary, uneven energy consumption in network will incur hotspots, where the data traffic load is heavy. For instance, the nodes near the base station are responsible for more data packet forwarding. In addition, if a dead node appears in the network, the node energy of the nonhotspot area is about 90% of the remaining energy [13, 14]. Moreover, repeated usage of the same paths and nodes will lead to hotspots to reduce network lifetime. Therefore, how to improve the SLP with high energy efficiency appears to be a difficult task in WSN security.

Considering the limited energy of WSNs, balancing the energy consumption of nodes will help to improve the life of the network. Therefore, the residual energy will be an important factor in the selection of forwarding nodes. The diversity of paths can not only improve the security of the network but also balance the energy consumption of nodes. However, the number of dynamic false sources and the length of branches will be contradictory to the security and the network lifetime. The privacy protection mechanism in this paper will seek a dynamic balance between them. While strengthening the network security and fully improving the energy efficiency of nodes, we propose a novel energy balanced multibranch privacy protection scheme, called Energy Balanced Branch Tree (EBBT). Our contributions are summarized as follows:

(1) A novel three-phase SLP protection scheme is proposed. This scheme uses constrained random walk, multibranch paths, and fake sources to achieve the protection of SLP

(2) The selection of phantom source nodes is improved, and the constrained random walk is used. In the process of moving away from the real node, the phantom source moves towards the periphery of the network

(3) The fake sources and multibranch paths are generated dynamically. The direction range of the branch path is determined to make the path more dispersive. And, the compromise of the length and number of branches are tackled to balance between the energy consumption and the security

(4) Theoretical analysis and experiments are carried out to evaluate the security period, energy consumption, and the path diversity and data delay

The rest of this paper is organized as follows: Section 2 provides related work. The network structure and the adversary model are given in Section 3. The detailed EBBT scheme is described in Section 4. The properties of the EBBT scheme are discussed in Section 5. The simulations and performance evaluation are presented in Section 6. Section 7 concludes the results.

2. Related Work

In 2004, Ozturk et al. [9] firstly introduced the concept called privacy of source location and designed the panda-hunter model and the phantom routing (PR) protocol. The proposed protocol is a two-stage technique. In the first stage, data packets of the source node are randomly forwarded to the phantom source node. And the packets are forwarded to the sink node through flooding in the second stage. One year later, Kamat et al. [15] proposed an energy-efficient phantom single-path routing (PSPR) protocol in which the trade-off between energy consumption and privacy performance is considered. Later, Xi et al. [16] illustrated the drawbacks of the phantom flooding protocol, thus proposing greedy random walks (GROW) where both the source and the sink simultaneously performed random walk. This scheme enhances the strength of privacy protection but also increases the transmission delay. Wang et al. [17] introduced the concept of attacker listening range and developed the phantom routing with locational angle (PRLA). In their work, random walks are chosen according to the random angle. PRLA can improve the security period compared with the single path phantom protocol, but more computations are needed to recognize the geographic location information of nodes. To improve the phantom routing, Li and Ren proposed Routing to a Random Intermediate Node (RRIN) [18]. A minimum distance from the intermediate node, i.e., phantom node, to the source node holds so that it is difficult for the attacker to locate the real source. Similar to RRIN, the Sink Toroidal Region (STaR) [19] area is designed to ensure the proper location of the intermediate node.

Wang et al. [20] proposed a multipath routing named Weighted Random Stride (WRS) adopting two parameters, randomly forwarding angle and stride value, to build a multipath route. Although using multipath routing can confuse the adversary and make it take longer time to track the source node, the energy consumption also increases. To reduce the energy consumption in the multipath scheme, a constrained random routing (CRR) scheme was proposed [21]. The CRR technique uses the changed next-hop node to build the multiple path route. The selection domain is also designed for the selection of the forwarding node. The selection weight is calculated. The next-hop nodes are selected according to the weight.

Chen and Lou [22] proposed a forward random walk (FRW) scheme. According to hop counts of candidate nodes to sink, the next-hop node in the FRW scheme is selected. The nodes used to forward packets are randomly selected from the closer set. The FRW scheme forms a transmission path between the sink and the source, resulting in insufficient privacy protection. To hide the source by using seek strategy, Long et al. [12] proposed a tree-based (TB) routing, where a backbone route between the source and the sink was established to forward the real message. On the other hand, the

diversionary route can send fake messages in the nonhotspot region to minimize the energy of the hotspot area. The bidirectional tree (BT) privacy protection scheme was presented [22]. Some branches are established in the route between the real source and the sink node. In the BT scheme, the real data arrived at the sink node along the shortest path so that the data transmission delay is reduced. With the increase of branches, the security period is greatly enhanced while the tree-shape route leads the additional energy overhead. Moreover, the staircase and branching scheme was introduced to prevent backtracking along the routing paths [10].

The stochastic and diffusive routing using multiple virtual source (SDR-m) technology was proposed [23]. Similar to the TB method, the SDR-m scheme also exploits the excess energy in the nonhotspot area. It is a two-phase scheme. The packet of the source node first reaches the virtual source through dispersive routes. The concept of escape-angle, sequitur-angle is then introduced. In order to avoid a decrease in the security period due to the expansion path loop, a path extension method (PEM) was proposed, in which the fake source is dynamically generated and the principle of the extended path combination is specified [24]. A method of mixing the permanent and temporary fake sources was proposed [25]. The selection of fake sources is designed as a scheduling problem. By dispatching the temporary and permanent fake sources to send packets, the security of the source location is improved, but the energy consumption is still high. Zhou and Wen [26] came up with an energy-efficient scheme by applying ant colony optimization, in which the information of distance, pheromones, and remaining energy are used to figure out the next-hop node.

Different from the above literatures, a three-stage source privacy preservation using a random routing scheme (SLP-R) was proposed [27], which does not use any fake source or multiple paths. The reverse random walk is moving away from the source node. The same depth route increases the path angle between the sink node and the source node. The minimum hop route ensures fast convergence and low latency. This scheme provides better privacy protection but increases the path length and delay of packet transmission. Han et al. [28] proposed the DRBR algorithm. At the mixed link point, the probability p is used to generate branches, and the attacker is induced to arrive at the fake source to improve the privacy protection intensity. Meanwhile, the transmission of fake data packets consumes extra energy. Wang et al. [29] proposed the SPAC scheme. The original packet is mapped into a set of shares and transmitted with minimal energy consumption. The fake source sends fake packets to the sink node at the edge of the anonymous cloud to confuse the attacker. The sink node can recover the original message after receiving at least T shares. This scheme protects the privacy of source location, improves the confidentiality of data, but reduces the accuracy of data.

3. System Model

The network structure and the properties of application scenario are abstracted, and the corresponding adversary characteristics and attack methods are also given in this section.

3.1. Network Model. The network has N sensor nodes and a base station (BS). When the sensor node monitors an object, the collected data are forwarded to the base station in a hop-by-hop manner. In this scenario, the sensor node becomes the source. However, if a monitored object cannot be detected for some time, the source node would stop sending packets. Some settings about the network model are given.

(1) The sensor nodes are uniformly distributed. All the nodes have the same configuration including calculation ability, communication radius, storage, and initial energy

(2) The packet is encrypted. The attacker cannot directly decipher the data packet to distinguish the real data packet and fake data packet

(3) The deep configuration information is initiated by the BS. Deep configuration information is broadcast. After the deep configuration phase, all nodes finally get the hop value to the BS

(4) Nodes can communicate with each other if their distance is less than the communication radius. They can obtain the information of neighbor nodes, i.e., the hops, ID of neighbor node, or other interesting information

(5) The base station is secure enough that the attackers cannot get any information from it

3.2. Adversary Model. Getting the source location is of great value to the attacker. The attackers trace the source location in terms of temporal correlation of transmitted data and traffic analysis among nodes. The attacker has the following settings.

(1) The attacker has complicated listening equipment, enough energy, enough storage, and strong computational power

(2) The attacker monitors the local communication situation. All the packets in the listening range will be found. But the attacker cannot listen to the packets out of the eavesdropping range. Attackers use a combination of eavesdropping and hop-by-hop tracking

(3) The attacker performs passive attack. It means that the attacker will not modify routing paths and data packets or destroy sensor nodes, etc., assuming that the attacker cannot decrypt the encrypted information. Therefore, the attacker intercepts and analyzes the data traffic to obtain useful information

(4) The attacker initially listens to the network near the base station until it overheard a packet sent by some nodes. The attacker analyzes traffic to determine the last node sending information. If there is no wireless signal monitored in the specified time during the tracking, it is considered that the tracking has failed, and the attacker quickly go back to the base station to start listening again

(5) The movement speed of the attacker is lower than the speed of packet transmission. Therefore, the attacker moves a node distance each time

4. The Proposed EBBT

This section describes the details of the proposed EBBT scheme. As shown in Figure 1, the EBBT scheme includes three phases: the random walk (RW), the minimum hop routing (MHR), and the creation of branch paths. In the RW phase, the real data packet of the source node is firstly forwarded H hops to reach a certain intermediate node C_{node}. Then, the real packet is transmitted from C_{node} to BS along the MHR path. The node $B_{i-0}(i = 0, 1, \cdots, n)$ that generates the branch path $\text{Bran}_i(i = 0, 1, \cdots, n)$ on the MHR path is selected with a certain probability p. The node at the end of Bran_i (green node in Figure 1) will play the role of the fake source. The fake source generates fake packets. The transmission path of fake packets consists of Bran_i. The transmission paths of the real and fake data form a tree-shaped routing structure rooted at BS.

4.1. Random Walk. The node N_i knows the remaining energy and the hop value of its neighbor nodes. Neighbor nodes are divided into three sets ES, CS, and FS according to the hop value. ES, CS, and FS are represented, respectively, as the set of neighbors with equal hop, smaller hop, and larger hop than node N_i has. In the RW phase, ES and FS constitute the forwarding candidate set of N_i. Then, N_i selects randomly the nodes with the most remaining energy from the candidate set as the next hop. R represents the communication radius of the sensor node and defines the maximum hop M of the network as

$$M = \frac{\text{network size}}{R}. \tag{1}$$

The number of forwarding is controlled by the parameter H. The value of H is defined by

$$H = \left\lceil \frac{M}{5} \right\rceil. \tag{2}$$

For a packet is forwarded, H is reduced by 1. If H is reduced to zero, the RW phase terminates. The last node C_{node} is the intermediate node. The nodes that existed in the RW path are deleted from the candidate set to avoid the case that the forwarding path forms a loop. The construction of the RW path and building the candidate set are described in Algorithm 1.

4.2. Minimum Hop Routing. A minimum hop routing (MHR) will be built from C_{node} to the BS. In the construction process of MHR, the candidate set of the next-hop node is CS. There will be a situation that the RW path crosses the MHR path. As shown in Figure 1, N_4 selects C_{node} as the next-hop node from the set ES and FS in the RW phase. The CS set of C_{node} contains N_4 and N_5, and N_4 may appear in both RW and MHR paths. If N_4 is the next-hop node of C_{node} in MHR, the RW path intersects with the MHR path. Therefore, in order to avoid the intersection of the two paths, the range of the node is defined. As shown in Figure 2, the range $[\alpha_2, \alpha_4]$ represent the scope of the node in the RW path.

The visible area is introduced [17] to improve the security of the source location. A certain range around the source node is defined as the visible area. The circle around the source node represents this area in Figure 2. The tangential range of the nodes in the visible area and C_{node} is represented by the range $[\alpha_1, \alpha_3]$.

The geographic region of the nodes of the RW phase and within the visible range of the source is represented by the angle α, which satisfies the relation as follows:

$$\min(\alpha_1, \alpha_2) < \alpha < \max(\alpha_3, \alpha_4). \tag{3}$$

The range of candidate node is beyond α. In addition to the geographical distribution of nodes, we also need to consider the remaining energy of the node. Algorithm 2 gives the construction process of MHR and the selection method of the initial node of branches.

4.3. MHR with Multibranch. The node on the MHR path becomes the initial node of Bran_i, B_{i-0}, with the probability p. Here, the parameter p controls the number of branch paths. But a branch path generated at C_{node} is not controlled by probability p. As shown in Figure 1, the branch path Bran_0 is generated at C_{node}, i.e., C_{node} is the initial node in Bran_0. The initial node B_{0-0} in Bran_0 selects the next-hop node B_{0-1} among its candidate set FS with most energy. The construction process of the branch path Bran_0 is shown in Algorithm 3. The node B_{0-m} at the end of the branch path is the fake source, and consequently, the node set $B = \{B_{0-0}, B_{0-1}, \cdots, B_{0-m}\}$ forms the branch path Bran_0. H_0 represents the length of this branch path.

In the process of constructing $\text{Bran}_i(i > 0)$ paths, in order to obtain good path dispersion, the following mechanisms are specified for the selection of the next-hop node. As shown in Figure 3, the x-axis of the coordinate system is the connecting line from the intermediate node to the sink node. The forwarding candidate set of $B_{i-k}(i > 0)$ is composed of nodes in ES and FS with the angle θ in the range L. The direction range L of a generated branch path is defined as

$$L = [\gamma - \varepsilon, \gamma + \varepsilon], \tag{4}$$

where ε represents the angular deviation. γ is used to control the scattered path. If θ is in the range L, N_j is the candidate node. If there is no matching node, the range of angle L will increase by ε. The node with the most remaining energy in the candidate sets becomes a new branch node. This process will repeat until the value of H_i reduces to 0. The node at the end of the branch path is the fake source. The decentralized branch paths make the energy consumption more balanced and avoid to pull the attacker back to the real path. The construction of branch path $\text{Bran}_i(i > 0)$ is shown in Algorithm 4.

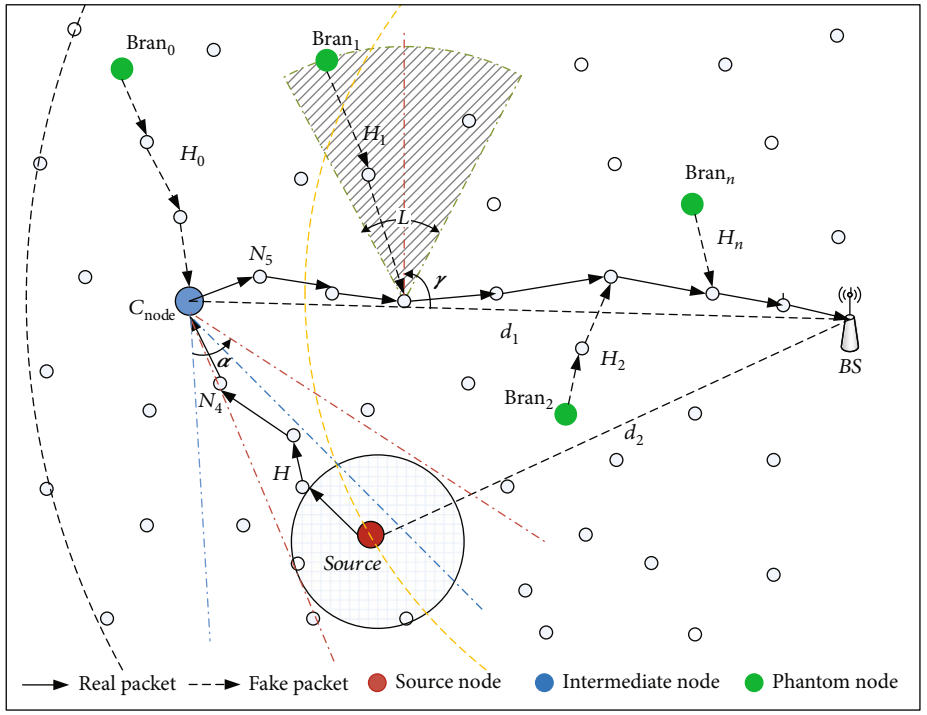

Figure 1: The proposed scheme of EBBT.

Input:
H: the hop number of RW phase;
N_i: the source node;
Output:
C_{node}: the intermedium node;
The routing path of RW phases;
1: Initiation: N_i_Next_hop =null;
2: N_i is put into RW set;
3: **while** N_i receive a configuration message **do**
4: Build the candidate set;
5: delete RW nodes from the candidate set;
6: Select the most energy node in candidate set as next hop N_{i+1};
7: $H = H - 1$;
8: **if** $H == 0$
9: return;
10: **else** RW scheme (N_{i+1});
11: **end while**

Algorithm 1: RW Scheme (Node N_i)

4.4. Analysis of the Number and Length of Branch Paths. In the EBBT scheme, the energy consumed by the branch path is the additional energy overhead. Therefore, the number and length of branch paths affect the security period and the network lifetime.

The nodes $B_{i-0}(i > 0)$ on the MHR path generates a branch path with probability p. Then, the length of the MHR path and probability p have influence on the number of branch paths. The length of the MHR path is defined as the hops between the C_{node} and the BS, i.e., $C_{node-hop}$. The probability p is dynamically changed and is controlled using a threshold value ω. Otherwise, the branch path should be far from the BS because the traffic load of the node near the BS is heavy. Therefore, the relationship among ω, $B_{i-0-hop}$ and p is defined as

$$\frac{C_{node-hop} - B_{i-0-hop}}{C_{node-hop}} + \frac{p}{\omega} = 1. \quad (5)$$

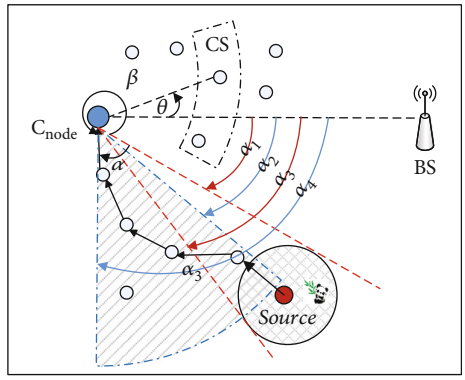

Figure 2: Angle range in the selection of next-hop node in the MHR path.

$B_{i-0-\text{hop}}$ is the hop value between the BS and B_{i-0}.The maximum value of $B_{i-0-\text{hop}}$ is $C_{\text{node-hop}}$. The maximum value of p is ω. Equation (5) is simplified as follows:

$$p = \left(1 - \frac{C_{\text{node-hop}} - B_{i-0-\text{hop}}}{C_{\text{node-hop}}}\right) * \omega. \tag{6}$$

It can be concluded from Equation (6) that p is related to $C_{\text{node-hop}}$, $B_{i-0-\text{hop}}$, and ω. The larger the $B_{i-0-\text{hop}}$ value, the larger the probability p is. The node B_{i-0} distribution exhibits a decreasing trend from C_{node} to the BS according to p. The relationship between the number of branches and the hops from C_{node} to the BS under different ω is shown in Figure 4.

Choosing an appropriate branch length is important because it affects energy consumption and security period. The length of the Bran_0 path is set to be the same as the length of the RW phase, that is, $H_0 = H$. $\text{Bran}_{i-\text{length}}$ represents the length of each $\text{Bran}_i(i > 0)$, where $\text{Bran}_{i-\text{length}}$ is related to $B_{i-0-\text{hop}}$ and H. $\text{Bran}_{i-\text{length}}$ is consistent with the change of $B_{i-0-\text{hop}}$. The maximum length of the branch path is H. The definition of $\text{Bran}_{i-\text{length}}$ is defined as

$$\text{Bran}_{i-\text{length}} = \frac{H}{\tau} + \left\lfloor \log\left(B_{i-0-\text{hop}}\right) \right\rfloor, \tag{7}$$

where τ is the adjustment parameter of the path length.

5. Performance Analysis

The average delay of data transmission, the energy consumption, and the security performance of this scheme are analyzed in this section.

5.1. The Average Hops. The forwarding path of real data in the EBBT scheme consists of two phases: RW and MHR. The average hops for MHR are shown as

$$\text{MHR}_{\text{avg}} = \frac{1}{M - d_2/R} \sum_{(d_2/R)+1}^{M} \frac{d_1}{R}. \tag{8}$$

As shown in Figure 1, d_1/R is the hops between C_{node} and the BS. d_2/R is the hops between the source and the BS. The maximum hop M is given in (1). According to the constraint of the random walk phase, the hop range of C_{node} is between d_2/R and M. The candidate region of C_{node} is in the area between the orange line and the edge of the network shown in Figure 1. The hops in the random walk phase are controlled by H. Therefore, the total delay for real data forwarding is described as

$$\text{EBBT}_{\text{delay}} = H + \text{MHR}_{\text{avg}}. \tag{9}$$

5.2. Energy Consumption. The main energy consumption of the EBBT scheme is to receive and forward real and fake data. Next, we analyze the energy overhead. Assuming that the frequency of source reporting event is equal to the frequency that each of the fake sources generates a fake data packet. When there is no real data, the fake sources stop generating the fake data packet. The path for transmitting a real packet includes RW and MHR. The RW path has H hops, the energy consumed by the single-hop propagation packet is set to be E_u, and the energy consumption is $H * E_u$. d_1 is the length of the MHR path. Then, the energy consumption of the MHR path is $d_1/R * E_u$. Therefore, the energy consumption for transmitting a real data packet is expressed as

$$\text{Cost}_T = H * E_u + \frac{d_1}{R} * E_u. \tag{10}$$

When the source node sends real data packets to the base station, fake packets are generated together. The fake packets are forwarded along Bran_i. The length of $\text{Bran}_i(i > 0)$ is calculated by Equation (7). Therefore, the energy consumption of the fake packets is represented as

$$\text{Cost}_F = H * E_u + E_u * \sum_{i=1}^{n} \left(\frac{H}{\tau} + \log\left(B_{i-0-\text{hop}}\right)\right). \tag{11}$$

The total energy consumption is calculated by

$$\text{Cost} = \text{Cost}_T + \text{Cost}_F. \tag{12}$$

The energy consumption of the node B_{i-0} is mainly divided into two parts: forwarding data packets coming from other nodes and forwarding data packets from B_{i-0} to BS. The energy consumption while B_{i-0} forwards the data packet in the random walk phase is $H * E_u$. The hop count between the intermediate node and B_{i-0} is $d_1/R - B_{i-0-\text{hop}}$, and the energy consumption is $\left(d_1/R - B_{i-0-\text{hop}}\right) * E_u$. The energy consumption while B_{i-0} forwards the packet to the BS is $B_{i-0-\text{hop}} * E_u$. B_{i-0} only receives the fake packet coming from B_{i-1} but not forwards it. The energy consumption of B_{i-0} is measured by the packet transmission on the average hop

Input:
The sink node;
N_i: the intermedium node;
Output:
The routing path of MHR phases;
The branch path;
1: Initiation: N_i_Next_hop = null;
2: Get visible area angle (α_1,α_3);
3: Get RW angle (α_2, α_4);
4: range of α = (min (α_1, α_2), max (α_3, α_4));
5: **while** N_i!=sink node **do**
6: Built the candidate set; the candidate set delete RW node;
7: Select the node with the most energy in candidate set and node not in angle α as the next hop;
8: **if** N_i== C_{node}
9: Bran0 (C_{node}, hop);
10: **else**
11: Generates a random number (Ran) that is uniformly distributed between 0 and 1;
12: **if** Ran< p
13: Bran(N_i, hop);
14: **end if**
15: **end if**
16: the next-hop node as current node N_i;
17: **end while**

Algorithm 2: MHR Scheme (Node N_i)

Input:
N_i: the intermedium node;
H_0: the length of Bran_0;
Output:
The routing path of Bran_0;
1: Initiation: N_i_Next_hop = null;
2: **while** receive Bran_0 message do
3: Build the candidate set;
4: The candidate set delete RW and Bran_0 node;
5: Select the node with the most energy from the candidate set as the Next_hop;
6: $H_0 = H_0 - 1$;
7: **if** $H_0 == 0$
8: return;
9: **end if**
10: **end while**

Algorithm 3: Bran_0 Scheme (Node N_i, Hop H_0)

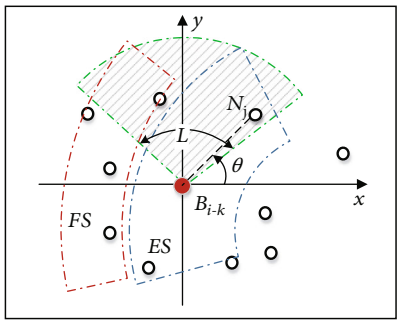

Figure 3: Selection of candidate node in the branch path.

count. Therefore, the total energy consumption of B_{i-0} is represented as

$$\text{Cost}_{B_{i-0}} = \frac{H * E_u + \left(d_1/R - B_{i-0-\text{hop}}\right) * E_u + B_{i-0-\text{hop}} * E_u}{B_{i-0-\text{hop}}} = \frac{(H + d_1/R) * E_u}{B_{i-0-\text{hop}}}. \tag{13}$$

The probability of repeated usage is reduced because the residual energy of the node is considered when selecting the next hop. And the energy of the node is reduced after a round of data transmission. In the next round, the node has no

Input:
N_i: the node in the MHR path;
H_i: the length of Bran$_i$;
Output:
The routing path of Bran$_i$;
1: Initiation: N_i_Next_hop = null;
2: **while** receive Bran-i message do
3: Build the candidate set; the candidate set delete FW, Bran, and CW node;
4: Get branch angle θ;
5: Select the most energy node in the candidate set and θ in range of γ as Next_hop;
6: $H_i = H_i - 1$;
7: **if** $H_i == 0$;
8: return;
9: **end if**
10: **end while**

Algorithm 4: Bran Scheme (Node N_i, Hop H_i)

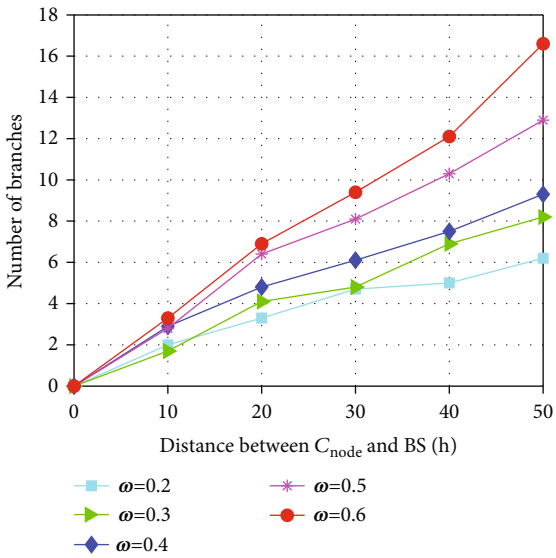

Figure 4: Number of branches changed with the hop from C_{node} to BS.

advantage of becoming the next hop. For example, in the candidate set of the node B_{0-0}, B_{0-1} is selected due to its most remaining energy. So, the remaining energy of B_{0-1} is reduced. Compared with other neighbor nodes, it would not be the next-hop node in the next round. At the same time, the next-hop node is also far away from the base station. Therefore, the branch path will extend to the edge of the network. In summary, the dispersed paths and the factor of the node remaining energy make the energy consumption to be uniform and prolong the lifetime of the network.

5.3. Security Period. In this section, we give the security analysis of the EBBT scheme. Assume that the monitored objects are randomly strolling in the protected area. The monitored object moves to another location after a period. Define the dwell time T. The source node sends the information to the BS if it senses the object. The information is sent at a rate of ψ messages per second. The security period Δ refers to the number of data packet sent by the source node before being captured by the attacker. If $T\psi < \Delta$, it means that the monitored object has left before the attacker arrives. Otherwise, the attacker may track back to the source node if the routing path is unique. A routing strategy with multibranch paths could cause an attacker to deviate from the real source and increase the network security period.

The total probability that the attacker chooses the branch path and the attacker chooses the correct path is 1. Since the attacker goes back to the base station, the probability that the attacker selects the true path or selects the branch path is all 1/2. The probability that an attacker always chooses the correct path is

$$P_{\text{correct}} = \left(\frac{1}{2}\right)^{n+1}, \tag{14}$$

where n represents the number of branches. C_{node} generates a branch path. The security period becomes long because the attacker has a higher probability of selecting the branch path and is trapped near the fake source node.

The number and length of branch paths in the EBBT scheme have a significant impact on the security period. As mentioned in Section 4, the length of the Bran$_i$$(i > 0)$ is described as Bran$_{i-\text{length}}$. The length of Bran$_0$ path is H. Therefore, the average of Bran$_{i-\text{length}}$ is defined as

$$\text{Bran}_{i-\text{ave-len}} = \frac{H + (H/\tau) + \lfloor \log\left(B_{i-0-\text{hop}}\right) \rfloor}{2}. \tag{15}$$

When the attacker is trapped in the branch path, it takes the attacker 2Bran$_{i-\text{length}}$ in the branch path and $2B_{i-0-\text{hop}}$ in the MHR path. According to formula (14), the real path is one of $n+1$ paths. Therefore, the safety period is defined theoretically as

$$S_H = H + \sum_{i=1}^{n} 2\text{Bran}_{i-\text{ave-len}} + \sum_{i=1}^{n} 2B_{i-0-\text{hop}}. \tag{16}$$

The safety period S_H is represented by the backtracking distance of the attacker. If the movement rate of the attacker is Δ_s, then the security period is defined as

$$S = S_H * \frac{\Delta_s}{\psi}. \quad (17)$$

When the source node is close to the BS, the network security period will reduce. The hop value between C_{node} and the BS can be increased by increasing the value of H in order to increase the security period. In addition, the threshold ω can be increased. In the EBBT scheme, the branch path, the fake sources, and the intermediate nodes are dynamically changed. The shortest path that is from the intermediate node to the base station is also generated dynamically. Therefore, there are enough paths to confuse the attacker. So, the source location is fully protected.

6. Simulation Experiment and Performance Evaluation

The proposed EBBT scheme in this paper is based on random walk and the routing of multibranch tree structure. Therefore, in the simulation, we use the schemes SLP-R [27], BT [22], and FRW [22] for comparison. The simulation experiment is performed on the MatlabR2014a simulation platform. Experimental environment settings are as follows: 1300 m × 1300 m network, 11000 sensor nodes are randomly and uniformly deployed. The purpose of this setting is to make the average number of neighbors per node up to about 11. The communication radius R of the sensor node and attacker is all 25 m. The node power and the communication radius could be adjusted when existing no suitable neighbor node, assuming that the parameters of EBBT are $\omega = 40$, $\gamma = 90$, $H = 10$, and $\Delta_s = \psi = 1$. The settings of related parameters are verified by the following simulation experiments. Under these settings, the system obtains a better network lifetime and security period. The initial energy of the node is 3 J. The data transmission energy consumption model used in the simulation experiment is referenced to the model in [30]. The length of the packet is 4000 bits. The simulation results given here are the average results of 25 trials, and the source node sends 500 packets in each round.

6.1. EBBT Algorithm Parameter Evaluation. The influence of parameter ω on the network lifetime is shown in Figure 5, and the influence of parameter ω on the security period is shown in Figure 6. The coordinate of the source node is (700, 650). As the value ω increases, the lifetime of network decreases slightly while the security period increases. Since ω is the threshold value of the control probability p, we can infer that p increases with the value of ω increases. Probability p controls the number of branches. The larger the value of p is, the more branches are created. In the analysis of Section 5, the more the number of branches is, the longer the security period has gained, which is consistent with the experimental results.

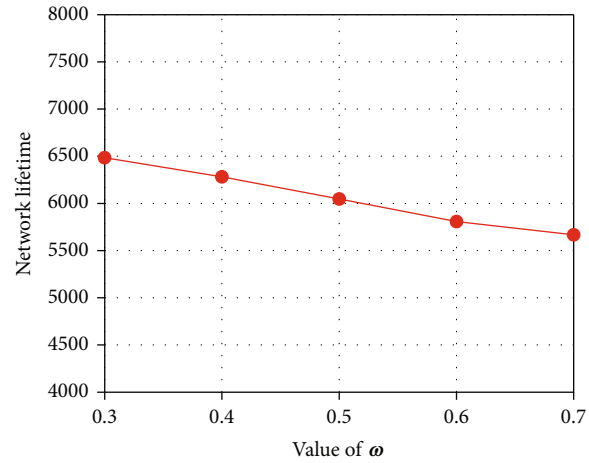

Figure 5: Effect of parameter ω on network lifetime.

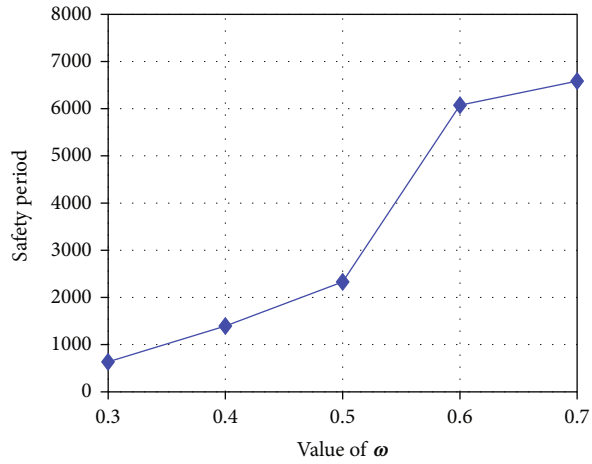

Figure 6: Effect of parameter ω on security period.

Figure 7 shows the effect of parameter H on the total number of nodes in branches. Since H_0 is equal to H, the length of H_i and H is positively correlated according to Equation (7). It indicates that there are more nodes in branches with the increment of H. It is consistent with the overall trend shown in Figure 7. The number of nodes is about twice times when the distance of the source to BS changes from 500 m to 1000 m. At the same time, with the increase of the distance from the source node to the base station, the number of nodes involved in the branch paths also increases. Therefore, selecting the appropriate number and length of branch path must be compromised between the network lifetime and security period. Figure 8 shows the effect of the lengths of the RW phase on the network security period. The larger the H value is, the longer the security period can be obtained.

6.2. Security Period. The simulation comparison of the source location privacy protection security period of four schemes FRW, SLP-R, BT, and EBBT are shown in Figure 9. The source node location in this scenario is between 500 m and 1000 m. The ordinate of the source node is 650 in y-axis. The security period tends to increase as the distance between the source

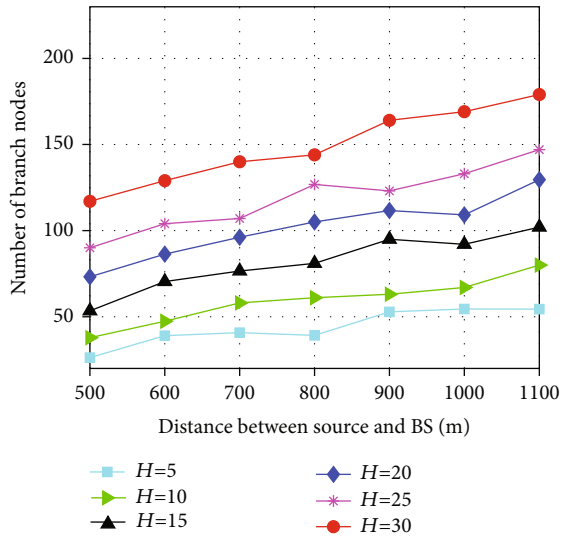

Figure 7: Effect of parameter H on the number of nodes in branches.

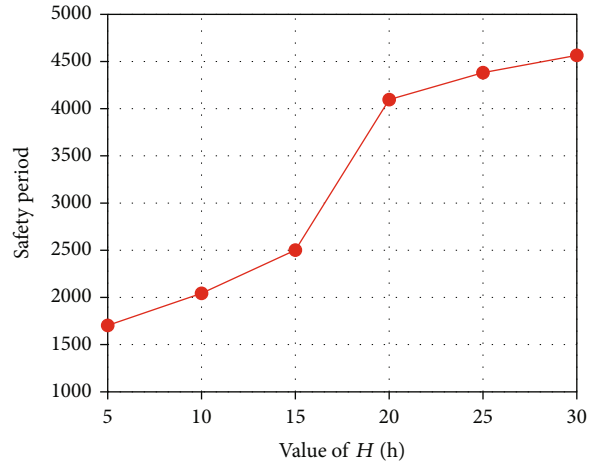

Figure 8: Effect of parameter H on security period.

node and the BS increases. In the FRW and SLP-R, there is only one path between the BS and the source. Therefore, the network has a shorter security period. An attacker could trace back to the source by a hop-by-hop manner. The security period of the BT method is higher than FRW and SLP-R due to the increasing branch path, but lower than the EBBT method. Because the node in the branch path may have a looped routing in the BT scheme, the attacker can directly skip the loop to reach the upstream node. The attacker's backtracking time is reduced, so the security period is lower than EBBT.

6.3. Network Lifetime. Figure 10 shows the effect of different source-to-BS distances on network lifetime. In the four schemes, the network lifetime of EBBT is the longest. The parameter γ in the EBBT controls the direction of the branch path such that the branch paths are relatively dispersed. When selecting the nodes in the path, the node with the most remaining energy becomes the next-hop node. Therefore, since the remaining energy of the node is reduced in the next

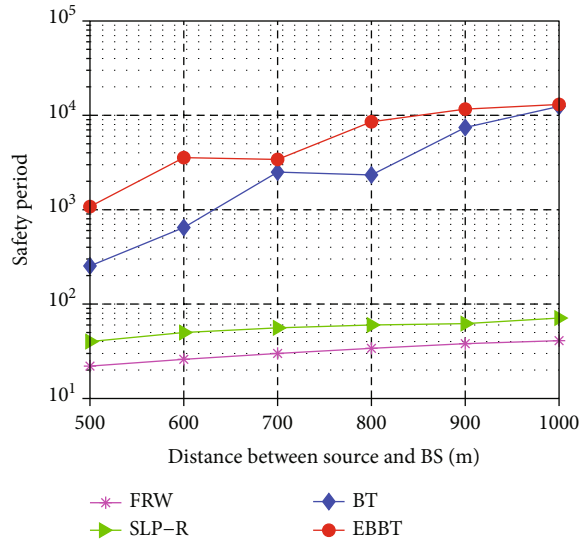

Figure 9: Network security period versus different distance to BS.

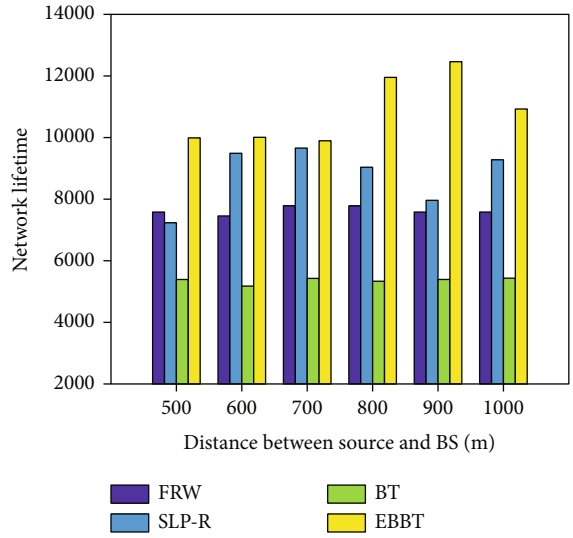

Figure 10: Network lifetime versus different distance to BS.

selection, there will be no advantage of becoming the most residual energy of the node. This guarantees the participation of all nodes. The lifetime of the network is increased by dispersing branch path load traffic to more sensor nodes.

BT and FRW have no obvious change of the network lifetime. It is because the routing path of the BT and FRW is limited in a narrow scope. Therefore, nodes may be reused. Heavy-duty nodes consume more energy, and they will become the bottleneck of the network lifetime. The number of nodes participating in the routing in the SLP-R scheme is lower than in the BT. And the SLP-R scheme selects the next-hop node in the overall network randomly. So, the lifetime is longer than FRW's and BT's. Although the number of nodes involved in transmission is increased in the EBBT scheme compared with the SLP-R scheme, the balanced consumption of communication energy in the network plays a positive role in prolonging the network lifetime.

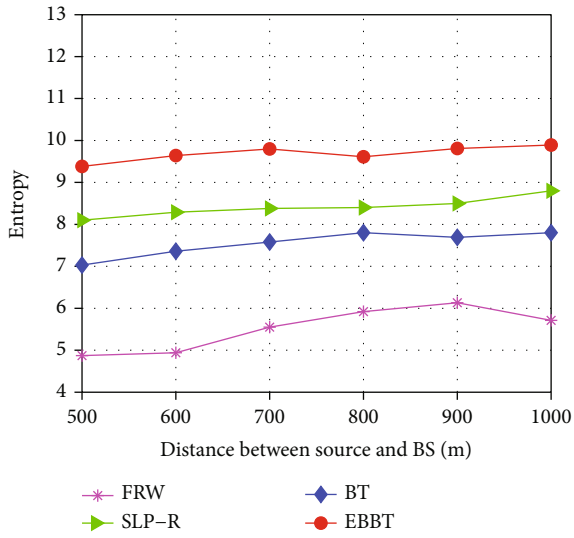

Figure 11: Path diversity versus different distance to BS.

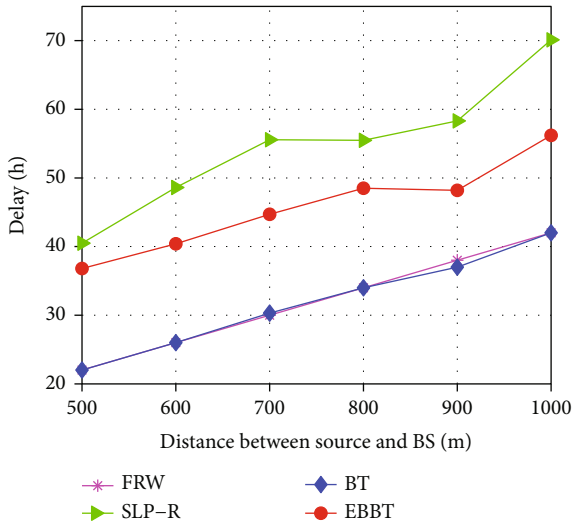

Figure 12: The average delay versus different distance to BS.

6.4. Path Diversity. The path diversity is evaluated by entropy. Entropy, a mathematical tool, is widely used as an indicator to measure the uncertainty of information. The entropy is defined as

$$H(N) = -\sum_{i=1}^{N} \frac{p(n_i)}{M} \log_2 \frac{p(n_i)}{M}. \quad (18)$$

$H(N)$ is referenced to the model in [27]. The higher value of $H(N)$ implies that the degree of the path diversity is better.

Figure 11 compares the path diversity of the four schemes using entropy. The source node sends 100 packets to the BS and computes $H(N)$ to measure the degree of the path diversity. As shown in Figure 11, the path diversity of the EBBT scheme is the highest. It indicates that there are more nodes involved in the path from the source to the BS in the EBBT scheme. The average load on each sensor node is low. Thereby, the network lifetime is improved.

6.5. Average Delay. The average delay refers to the time that packets are transmitted from the source node to BS. In the EBBT scheme, this time is described by the number of hops that data packets are forwarding on the route. Figure 12 shows the average delay of the four schemes. The next-hop node in the FRW is close to the base station. The data packets in BT are sent to the BS through the shortest path, so the average delay in BT and FRW is the least. The SLP-R technology has the largest delay, and the SLP-R algorithm used a three-phase routing path; its routing path is long, and the delay increases. The routing path of the real data in SLP-R is longer than BT and EBBT. The delay of the EBBT scheme is mainly originated by random walk. The packets are sent to a certain intermediate node firstly, which increases the delay. The EBBT solution improves the degree of privacy protection at the cost of relatively less delay.

7. Conclusions

In this paper, an energy-efficient source location privacy protection scheme (EBBT) was proposed which is mainly based on random walk and fake sources. In EBBT, multiple branches are used to transmit fake data streams. The real data flow and the fake traffic flow are mixed, together with the sources forming a tree-like routing structure. With the dynamic change of intermediate nodes, the whole tree structure also changes dynamically. The branch path control angle and the remaining energy of nodes are fully considered which greatly improve the utilization of energy. Theoretical analysis and the results of simulation experiment show that the proposed scheme has a higher network security period and network lifetime than the related scheme, and it has high path diversity. Our future work is to design the data privacy protection method while protecting the location of the source node.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The work is supported by the National Natural Science Foundation of China (61862011), the Natural Science Foundation of Guangxi (2019GXNSFGA245004 and 2019GXNSFAA245053), the Science and Technology Major Project of Guangxi (AA19254016 and AA18118025), and the Project of Guangxi Key Laboratory of Trusted Software (KX202056).

References

[1] L. Lombardo, S. Corbellini, M. Parvis, A. Elsayed, E. Angelini, and S. Grassini, "Wireless sensor network for distributed environmental monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 5, pp. 1214–1222, 2018.

[2] H. Mshali, T. Lemlouma, M. Moloney, and D. Magoni, "A survey on health monitoring systems for health smart homes," *International Journal of Industrial Ergonomics*, vol. 66, pp. 26–56, 2018.

[3] X. J. Kong, X. T. Liu, B. Jedari, M. Li, L. Wan, and F. Xia, "Mobile crowdsourcing in smart cities: technologies, applications, and future challenges," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8095–8113, 2019.

[4] F. Wu, C. K. Qiu, T. Y. Wu, and M. R. Yuce, "Edge-based hybrid system implementation for long-range safety and healthcare IoT applications," *IEEE Internet of Things Journal*, vol. 8, 2021.

[5] C. M. Chen, B. Xiang, T. Y. Wu, and K. H. Wang, "An anonymous mutual authenticated key agreement scheme for wearable sensors in wireless body area networks," *Applied Sciences*, vol. 8, no. 7, p. 1074, 2018.

[6] T. Y. Wu, L. Yang, Z. Lee, C. M. Chen, J. S. Pan, and S. K. Islam, "Improved ECC-based three-factor multiserver authentication scheme," *Security and Communication Networks*, vol. 2021, Article ID 6627956, 14 pages, 2021.

[7] S. Kumari, P. Chaudhary, C. M. Chen, and M. K. Khan, "Questioning key compromise attack on Ostad-Sharif et al.'s authentication and session key generation scheme for healthcare applications," *IEEE Access*, vol. 7, pp. 39717–39720, 2019.

[8] P. M. Chanal and M. S. Kakkasageri, "Security and privacy in IOT: a survey," *Wireless Personal Communications*, vol. 115, no. 2, pp. 1667–1693, 2020.

[9] C. Ozturk, Y. Y. Zhang, and W. Trappe, "Source-location privacy in energy-constrained sensor network routing," in *SASN '04: Proceedings of the 2nd ACM workshop on Security of ad hoc and sensor network*, pp. 88–93, Washington. DC. USA, October 2004.

[10] B. Chakraborty, S. Verma, and K. P. Singh, "Staircase based differential privacy with branching mechanism for location privacy preservation in wireless sensor networks," *Computers & Security*, vol. 77, pp. 36–48, 2018.

[11] A. Jhumka, M. Leeke, and S. Shrestha, "On the use of fake sources for source location privacy: trade-offs between energy and privacy," *The Computer Journal*, vol. 54, no. 6, pp. 860–874, 2011.

[12] Jun Long, Mianxiong Dong, K. Ota, and Anfeng Liu, "Achieving source location privacy and network lifetime maximization through tree-based diversionary routing in wireless sensor networks," *IEEE Access*, vol. 2, pp. 633–651, 2014.

[13] A. F. Liu, P. H. Zhang, and Z. G. Chen, "Theoretical analysis of the lifetime and energy hole in cluster based wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 71, no. 10, pp. 1327–1355, 2011.

[14] Y. Ouyang, Z. Y. Le, G. L. Chen, J. Ford, and F. Makedon, "Entrapping adversaries for source protection in sensor networks," in *2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06)*, Buffalo-Niagara Falls, NY, USA, June 2006.

[15] P. Kamat, Y. Zhang, W. Trappe, and C. Ozturk, "Enhancing source-location privacy in sensor network routing," in *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, Columbus, OH, USA, June 2005.

[16] Y. Xi, L. Schwiebert, and W. S. Shi, "Preserving source location privacy in monitoring-based wireless sensor networks," in *Proceedings 20th IEEE International Parallel & Distributed Processing Symposium*, pp. 25–29, Rhodes Island. Greece, April 2006.

[17] W. P. Wang, L. Chen, and J. X. Wang, "A source-location privacy protocol in WSN based on locational angle," in *2008 IEEE International Conference on Communications*, pp. 1630–1634, Beijing, China, May 2008.

[18] Y. Li and J. Ren, "Source-location privacy through dynamic routing in wireless sensor networks," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, San Diego, CA, USA, May 2010.

[19] L. Lightfoot, Y. Li, and J. Ren, "STaR: design and quantitative measurement of source-location privacy for wireless sensor networks," *Security and Communication Networks*, vol. 9, no. 3, pp. 220–228, 2016.

[20] H. D. Wang, B. Sheng, and Q. Li, "Privacy-aware routing in sensor networks," *Computer Networks*, vol. 53, no. 9, pp. 1512–1529, 2009.

[21] W. L. Chen, M. S. Zhang, G. W. Hu, X. L. Tang, and A. K. Sangaiah, "Constrained random routing mechanism for source privacy protection in WSNs," *IEEE Access*, vol. 5, pp. 23171–23181, 2017.

[22] H. Chen and W. Lou, "On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks," *Pervasive and Mobile Computing*, vol. 16, pp. 36–50, 2015.

[23] R. Manjula and R. Datta, "A novel source location privacy preservation technique to achieve enhanced privacy and network lifetime in WSNs," *Pervasive and Mobile Computing*, vol. 44, pp. 58–73, 2018.

[24] W. Tan, K. Xu, and D. Wang, "An anti-tracking source-location privacy protection protocol in WSNs based on path extension," *IEEE Internet of Things Journal*, vol. 1, no. 5, pp. 461–471, 2014.

[25] A. Jhumka, M. Bradbury, and M. Leeke, "Fake source-based source location privacy in wireless sensor networks," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 12, pp. 2999–3020, 2015.

[26] L. M. Zhou and Q. Y. Wen, "Energy efficient source location privacy protecting scheme in wireless sensor networks using ant colony optimization," *International Journal of Distributed Sensor Networks*, vol. 11, pp. 1–14, 2014.

[27] M. Raja and R. Datta, "An enhanced source location privacy protection technique for wireless sensor networks using randomized routes," *IETE Journal of Research*, vol. 64, no. 6, pp. 764–776, 2017.

[28] G. J. Han, M. T. Xu, Y. He, J. F. Jiang, J. A. Ansere, and W. Zhang, "A dynamic ring-based routing scheme for source location privacy in wireless sensor networks," *Information Sciences*, vol. 504, pp. 308–323, 2019.

[29] N. Wang, J. Fu, J. Li, and B. K. Bhargava, "Source-location privacy protection based on anonymity cloud in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 100–114, 2020.

[30] G. J. Han, H. Wang, J. F. Jiang, W. B. Zhang, and S. Chan, "CASLP: a confused arc-based source location privacy protection scheme in WSNs for IoT," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 42–47, 2018.