JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# A Probabilistic Source Location Privacy Protection Scheme in Wireless Sensor Networks

**O.V.SOWMYA**
**ASSISTANT PROFESSOR**

**Abstract**-Wireless Sensor Networks (WSN) is one of the main zone of research and it has been more well known in the real life difficulties by giving minimal effort arrangements. The system comprises of little sensor nodes capable for detecting, handling, computation and communication. The system comprises of various sorts of assault, the most harming assault is sinkhole assault. In this kind of assault, the sinkhole node tries to draw in information to itself by transmitting counterfeit data to neighbor nodes and henceforth it interferes with the usefulness of such systems. Thusly with a specific end goal to overcome from this sort of assault giving security is critical. In this paper we are proposing sink and also source area protection systems. With a specific end goal to give more protection these procedures like privacy protection scheme and Hidden Markov Model (Base station location anonymity and security technique) is utilized. On account of forward random walk conspire requires every node to acquire its hop count to the sink, which can be accomplished by utilizing a sink-based flooding. Toward the starting, the sink will start a flooding, after which every node can get the both its neighbors hop count to the sink. On account of PPS conspire the center thought is to change the transmission scope of an arrangement of some chose sensors around the base to befuddle the assailant. Through this procedure, we can make an arrangement of fake base stations which can't be recognized by a solid assailant. The simulation results demonstrate that the proposed PSLP scheme improves the safety time without compromising the energy consumption.

**Keywords: Wireless Sensor Networks, Privacy Protection Scheme, Hidden Marko Model.**

## 1. INTRODUCTION

Wireless Sensor network arrange comprises of minute sensor nodes, low-power, light weight, minimal effort. Because of the ease of these nodes, the situating can be in the request of result of thousands to million nodes. The sensor nodes perform desire estimations, process the information and send it to a base station, which is usually referred to as sink node. The base station gathers the information from every one of the nodes and assesses the information from every node. Source Location Privacy (SLP) in sensor arrange implies the status that the data about the locations of events identified by sensors nodes is appropriately secured from unauthorized users, for instance, the sink of the network, can acquire the data. Observing and identifying events is an ordinary utilization of sensor networks. Preserving source location in sensor networks is challenging primarily because of the accompanying reason, the restricted resources accessible in sensor networks requires exceptionally effective security protection components

Wireless Sensor Networks (WSNs) are materializing as a encouraging technology because of their large range of uses and applications in industrial, environmental watching, military and civilian fields. Because of economic deliberation, the sensor nodes are commonly simple and of less cost.

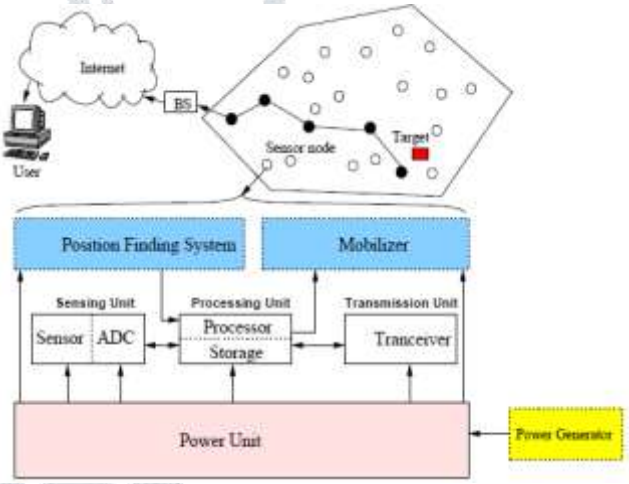


**Figure 1:** Components Of sensor nodes

Figure 1 shows the structural diagram of sensor node components. Generally, sensor node consists of sensing, processing, transportation, mobilizer, position detecting system, and power units. The same diagram represents the communication structure of a WSN. Sensor nodes are basically dispersed in a sensor field, which is the region where the sensor nodes are placed. Sensor nodes cooperate between themselves to generate good-quality information about the physical surroundings. Every sensor node takes its decisions on its goal, the data it presently has, and its familiarity of its computing, communication, and power resources. Each of these dispersed sensor nodes has the ability to gather and forward information either to other sensor nodes or back to an outside base station. A base station may be a static node or a dynamic node which is capable of connecting the wireless sensor network to an already presented communications structure or to the network where a user can have entry to the reported information. In the working of Wireless Sensor network, the sensor nodes are often neglected, nevertheless, and are thus likely to go through from various types of novel attacks such as compromised node

attack, Denial of service attack, Black hole attack, etc. A black hole attack (BLA) is one of the most basic and common type of attacks, in which the attacker captures a sensor node and drops all data packets that are forwarded through this sensor node, concluding in important and sensitive data being rejected or not able to be forwarded to the BS target. Because the network comes on the conclusion depending on the sensor nodes' captured data, the aftermath is that the network will totally break and, more seriously, take wrong decisions. And hence, how to find and prevent various attacks is of great implication for security in WSNs

.

## II LITERATURE SURVEY

Many researchers have paid attention to the location privacy since Ozturk first proposed his concept [12]. Recently, location privacy has been widely researched in industrial wireless sensor networks [13], vehicular ad-hoc networks [14], cloud computing [15], social network [16] and so on. Location privacy covers the source location privacy and the sink location privacy. In this paper, we focus on the source location privacy protection. Manjula et al. used virtual sources to protect the source location privacy [17]. In their scheme, a routing technique was proposed to maximize the safety time. By adding random walk into the routing process, nodes in non-hotspot areas participated in the establishment of multiple routing paths. Hence, the safety time increased without influencing the network lifetime. Matthew et al. proposed two algorithms using fake sources to protect the source location privacy [8]. In the first algorithm, fake sources were dynamically deployed around the sink. Then, the sink used flooding to select fake sources. This algorithm can provide a good source location privacy at the expense of the huge energy consumption. To cope with this, another algorithm called dynamic single path routing algorithm (DynamicSPR) was proposed. By using directed random walk, nodes away from the source were selected as fake sources, which significantly reduced the energy consumption. However, fake sources were related to the relative location of the source and the sink, sensor nodes in a specific area might exhaust energy. Jing et al. considered a more powerful adversary and proposed a privacy enhancing routing algorithm to protect location privacy [18]. In their research, a global adversary using Bayesian maximum-a-posteriori (MAP) estimation strategy tried to monitor the communication between nodes. Then, a decision-making framework was put forward to reduce the adversary's detection probability. Finally, the problem was converted into the adjustment of parameters. Huang et al. focused on the energy utilization rate in WSNs while maintaining the source location privacy [19]. They proposed a redundancy branch-based source location privacy scheme. In their scheme, many redundancy branches were generated from the source to the sink. The number of branches was determined by the energy collected by nodes. In addition, these branches were converged into several routing paths later. However, the number of converged routing paths was not clearly defined and the energy collected by nodes around the sink might be less than the energy costed by transmitting packets. Chen et al. in [20] proposed a constrained random walk mechanism. In their mechanism, a next-hop candidate selection domain was generated based on the offset angle of current node's neighbors and the danger distance, which made the selection domain look like an ellipse. Then, the weight of each node in the domain was calculated by the ratio between a current node's offset angle and the sum of total offset angle. The smaller the ratio, the higher the probability that this node became the next-hop candidate. However, the offset angle of a node was fixed, and thereby the weight might not change. Thus, nodes which acted as the next-hop candidate might consume too much energy. Chen et al. utilized phantom nodes and proposed a limited flooding algorithm to protect the source location privacy [9]. The limited flooding was performed by the source to get the information of nodes in the limited flooding area. Then, nodes on the edge of the limited flooding area were chosen as phantom nodes to simulate the function of the source. If a phantom node stayed behind the source, packets sent by this phantom node first bypassed the visible area and were then transmitted to the sink using the shortest path. However, the limited flooding was repeatedly performed, which might not suitable for a large scale network. Li et al. in [21] proposed a scheme using random intermediate nodes and ring to protect the source location privacy. First, the authors introduced the criteria to quantitatively measure the source location information leakage. Then, to reduce the leakage probability, random intermediate nodes were added to make the routing path disperse. Packets were first transmitted to an intermediate node and then forwarded to a node in ring around the sink. Packets were routed on the ring for a random hop and then sent to the sink. Mutalemwa et al. divided the whole network into regions and proposed a scheme based on region transmission [22]. In this scheme, the sink was located in the center of the network and regions were generated around the sink. The transmission between regions was implemented by a set of relay nodes which were selected strategically. These strategic relay nodes took up two regions and were responsible for forwarding packets to the sink. However, the distribution of these nodes were close to the sink. Relaying too many packets would consume a lot of energy. Thereby, the average energy efficiency was not high. Wang et al. considered the source location privacy against a new type of adversary in [23]. The adversary model had two properties, global and local. Under normal circumstances, the adversary was a local adversary. When a potential area where the source may stay was located, the adversary became a global adversary in this area. To cope with it, a message mapping sharing method was presented and a cloud containing many dummy packets was created around the source to hide the location. Each message copy was transmitted by random routing, which provided sufficient source location privacy. Considering the time correlation during the transmission between sensor nodes, Mayank et al. used the data mule to protect the source location privacy [24]. Data mule worked as the mobile data collection unit and collected data when the source was in its communication radius. In this condition, the source location privacy was changed into the protection of the mule's moving track. Then, the authors proposed three extended versions of angle-based scheme to protect the source location. However, since the mule moved grid by grid, the protection of the mule was not given enough attention. There was still a lot of research space in reducing the time correlation. To further reduce the time correlation during the transmission, Proa~no et al. proposed a traffic decorrelation technique to reduce the threat of a global adversary [25]. The proposed traffic normalization scheme reduced the communication overhead and the transmission delay. In addition, the whole network was partitioned into a set of minimum connected areas with a circular queue, which could reduce the active nodes during transmission and the adversary's eavesdrop probability. The privacy in their work was quantified to the distance between location estimated by the adversary and the location of the source. From the above works, it can be seen that source location privacy protection has experienced a great improvement, techniques like fake sources, phantom nodes, random walk, and the weight are developed. However, these techniques are only used in a simple way, which gives us an inspiration.

## III. PROPOSED SYSTEM

Security of WSNs involves many aspects, such as data privacy [5] and location privacy [6]. Data privacy can be protected by encryption algorithms while location privacy cannot be protected to the extreme. Due to the time correlation in data transmission between two nodes, the adversary can infer location information through analysis. From a time correlation perspective, location privacy consists of the source location privacy and the sink location privacy. Given the importance of the source, in this paper, we focus on the source location
privacy, which is an emerging research topic in the field of security. There are many techniques, like secure routing [7], fake sources [8], phantom nodes [9], fake cloud [10], and cluster [11], that can be applied to protect the source location privacy. We propose a probabilistic source location privacy protection scheme (PSLP), which adopts phantom nodes and fake sources for the reason that these two techniques can diversify the routing path. In the above steps, the visible area is a special area. When the adversary backtracks to this area, the source can be recognized immediately. Two types of packets exist in the transmission, which are the real packets and the fake packets. Real packets are generated by the source while fake packets are generated by fake sources. In order to hide the source location, real packets sent by the source are first transmitted to a phantom node through directed random walk. Here, considering the distance between the source and the sink, two transmission modes are taken into consideration and details will be given later. During the transmission of real packets, fake packets are also transmitted to the sink with a fixed period. The proposed PSLP has exhibited a better performance than two other recent schemes in our simulations with regard to increasing the safety time while balancing the energy consumption. The main contributions of this paper are: 1) Both phantom nodes and fake sources are integrated into the proposed PSLP, which enhance the source location privacy. 2) A more powerful local adversary, which can use Hidden Markov Model to estimate the state of the source, is taken into consideration. 3) Two data transmission modes are designed based on the distance between the source and the sink, which further enhance the source location privacy.

The network model in this study is based on the typical Panda-Hunter model [12]. As shown in Fig. 1, a WSN which is composed of many sensor nodes is deployed to monitor the activities of pandas. Once a sensor node detects a panda, it becomes the source and sends packets to the sink through multiple hops. The essence of privacy protection is reducing the probability that the adversary finds the source location. Therefore, we make the following assumptions: 1) Sensor nodes are randomly deployed. After being deployed , the location of each sensor node remains unchanged. What's more, all sensor nodes are homogenous, which means that they have the same initial energy, the same computing ability, and the same cache memory. 2) Routing is based on the weight. Each sensor node is assigned a weight that is updated regularly. The weight here represents the probability that this node is selected as the next hop, or it can be understood as the preference in selecting the next hop node, which is related to the residual energy, the communication quality, and the hop count to the sink. Details of this weight will be given later. 3) Only one sink exists in the network. As in other schemes or protocols [12], [15], [22], the sink remains in the network center. 4) Each sensor node has knowledge of its own adjacent neighbors. Packets sent by each sensor node are encrypted with an encryption algorithm. However, this part is beyond the scope of this study.

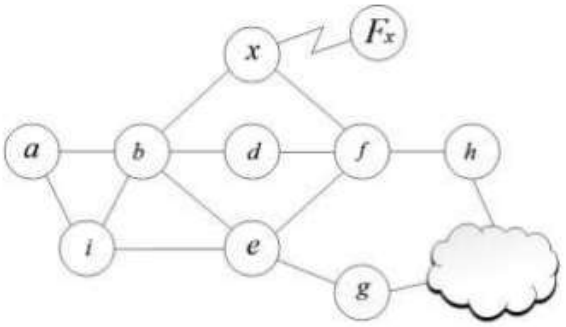


Fig. 1. Example of a node isolation attack: node x claims to know every two-hop neighbor of b, as well as Fx, a non-existent node.

**Fig 1. The Hunter Model**

In this section, a detailed description of PSLP is given. In the initialization process, the beacon message is periodically broadcasted by the sink to sensor nodes. When a node receives the message, it records the hop count stored in it, increases the hop count by one, repackages the packet, and sends to its neighbors. Each node only records the minimum hop count. Subsequently, all nodes know their hop count to the sink and their neighbors. Since the adversary may know the state of the source at a given time while the location of the source is still unknown, we intend to increase the possible locations of the source. PSLP contains three steps: the first step is the determination of phantom nodes; the second step is the determination of fake sources; the third step is the routing from the source to the sink. of the two is different. The phantom node refers to nodes around or nearby the source, which simulate the function of the source. The fake source also refers to nodes which simulate the function of the source. But the location of fake source is around the sink, which is far from the source. The motivation of combining the phantom node and the fake source together is to create the diversification of the transmission directions. Both phantom nodes and fake sources are selected in non-hotspot area, which has little influence on the network lifetime.

## IV METHODOLOGY

**Modules**

1) Source In this module, Source browses the file, choose the destination and sends to the router. In Source at the same time as uploading the document, encrypt after which uploads the file. File content material will be initialized to all the nodes

2) Router Routers are small electronic gadgets that join multiple computer networks collectively through either stressed out or wireless connections. In this module, router includes four Networks, each Network consists of specific nodes. When Source sends the file to begin with it comes to the Network1 and passes via the Network1 nodes, if any congestion located within the Network1 node, it robotically selects every other node and actions to Network2 and Network 3 and Network4 and reaches the destination. The energy length additionally be changed, view the Network information. In router the routing course and time postpone may be regarded

three) Router Manger In this module, Router Manger perspectives the attacker information by using checking the energy info and find attackers.

Four) Destination In this module, Receiver request for file call and mystery key and receives the content material from the router. Time postpone might be calculated by sending the file from supply to destination and time taken to reach the destination.

5) Attacker In this module, attacker selects the Network and node, receives the unique strength size and modifies the strength size for the node.

The first and major approach for development of a assignment starts from the thought of designing a mail enabled platform for a small firm wherein it is easy and handy of sending and receiving messages, there may be a seek engine ,cope with book and also inclusive of some wonderful video games. When it is accepted by using the business enterprise and our assignment manual the primary hobby, ie. Preliminary research starts.

Algorithm 1:
Detection of Attack
/* After the question-issuing node receives all reply messages */

1: INPUT: Top-ok Result, Relpay_Message
2: OUTPUT: Route Information
3: Route Information ← ∅
4: for each Relpay_Message do
five: for each Top-ok Result do
6: if Replay_path_Vale includes the node ID of a node processing a facts object in Top-okay Result and Score_Value_List does no longer include the records object then
7: Insert a route from the node with the missingdata item
8. The query-issuing node into SendRoute
nine: cease if
10: cease for
eleven: quit for
12: if Route Information= ∅ then
thirteen: Detect Attack
14: give up if

False Notification Attack

The nodes are grouped with some similar residences. Each group can have a set in-charge that is elected with the aid of Nodes maximum ID. If a few node inside any organization identifies an assault primarily based on the algorithm 1 it's going to record the malicious node id to is institution in-fee and this records is shared with all other group in- rate within the community. Each organization in-price will try to conform climate the node is malicious node or lire node (LN) .Where LN are regular node with a view to contribute a fake fee , No cost to top-K question. The LN nodes will update the rating price in Score_Value_List so it is not a malicious node, to affirm this the query issuing node will send a request to this LN node to send its score value. Then the values are compared with the values collected from replay messages if the values are of in extra variant the LN nodes are categorised.

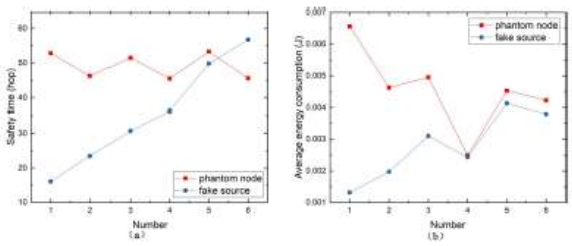


Fig 2: Influence of Safety and Average energy Consumption

Figure.1 suggests the accuracy of question end result received with the aid of query-issuing node. The X-axis denotes the quantity of requested information gadgets and Y-axis denotes the accuracy. The proposed pinnacle-k question technique will increase the accuracy even when the range of asked records gadgets is big. Figure.2 shows the traffic took place when question results are forwarded in more than one routes. The X-axis denotes the wide variety of requested data items and Y-axis denotes the visitors. Figure.Three shows the malicious node identity ratio that represents maximum wide variety of identified malicious node with the aid of issuing less range of queries. The X-axis denotes the query issuing time and misidentification.

The suggest approach suggests the question end result improves when the malicious nodes are recognized and removed and additionally while the malicious nodes are gift the query end result accuracy is low as shown in determine 2. The parent 3 indicates the site visitors float whilst the queries are issued inside the community and it's far in comparison with the attack and without assault. The site visitors is excessive while there may be a malicious nodes inside the community on account that they make a contribution false facts inside the community this lead the regular node to send extra query to relax on correct result.

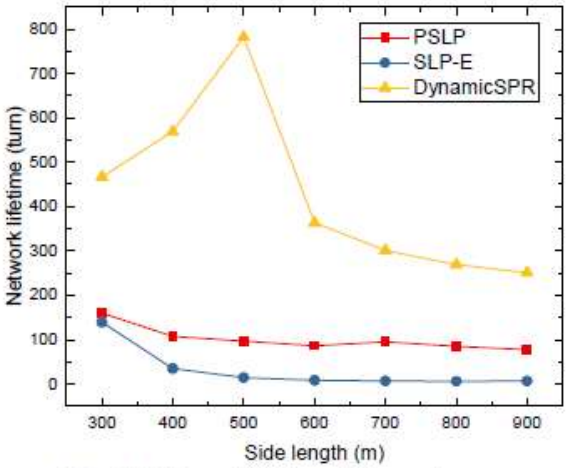


Fig. 12: Network lifetime of three schemes.

**Fig 3: Network Lifetime of Three Schemes**

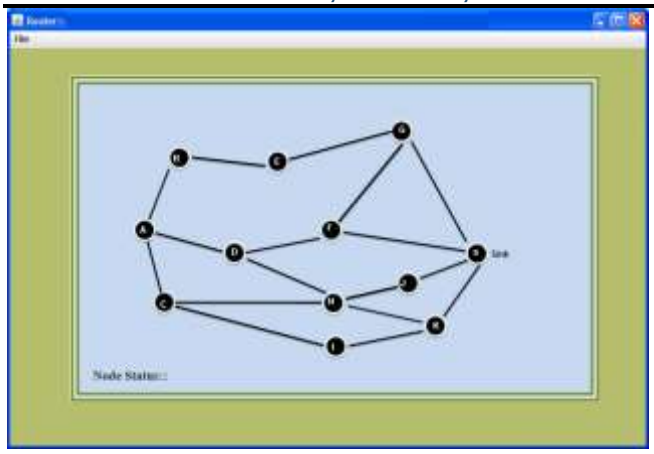

**Fig 4: Router with Various Routes**

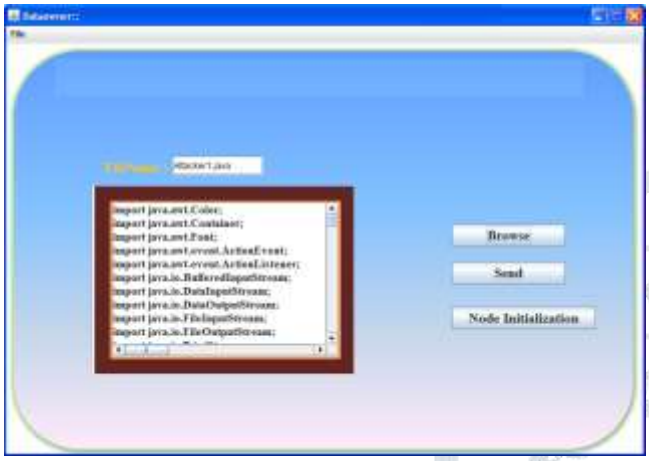


**Fig 5: Source Window to Transmit Data**

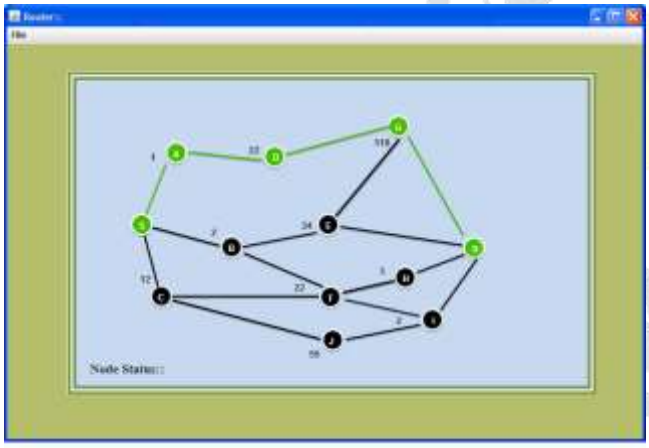

**Fig 6: Router without Attacks**

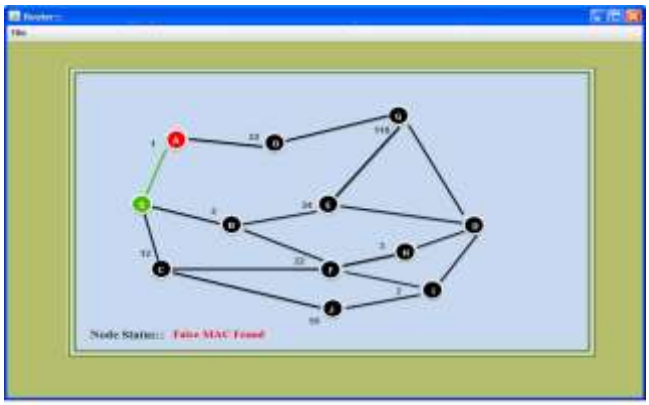

**Fig 6: Router with Attacks and Rerouting**

## V. CONCLUSION

WSNs have hugely extended in playing a key job for the data decisive selection and shipment. As the sensor nodes are deployed in neglected environment, they are more vulnerable to the attacks. If the attack is not detected in WSN, there is a chance of taking wrong decision and it is a serious issue in the sensitive fields such as in military, defense, etc. So there is a need to consider the issue of detecting and preventing such types of attacks and make the safe and reliable routing of data packets. In this paper spotlight is on the studying various types of attacks that can be performed on the sensor node such as denial of service, Black hole attack, etc. Also the techniques to detect the sensor node under the control of adversary and the type of attack happened on the node are presented in the paper. It also includes a solution for secure and reliable routing of the data in WSN. Considering the distance between the source and the sink, two types of routing modes are designed. Compared with DynamicSPR and SLPE, the simulation results demonstrate that the proposed PSLP achieves a high safety time and balances the energy consumption of each node. Future studies will concentrate on protecting the source location by reducing the adversary's monitoring probability and secure communication among nodes

## REFERENCE

[1] H. Lu and J. Li, "Privacy-preserving authentication schemes for vehicular ad hoc networks: a survey," Wireless Communications and Mobile Computing, vol. 16, no. 6, pp. 643-655, Apr. 2016.

[2] G. Han, X. Yang, L. Liu, S. Chan, and W. Zhang, "A Coverage- Aware Hierarchical Charging Algorithm in Wireless Rechargeable Sensor Networks," IEEE Network Magazine, pp. 1-7, Nov. 2018, DOI: 10.1109/MNET.2018.1800197

[3] G. Han, H. Guan, J. Wu, S. Chan, L. Shu, and W. Zhang, "An Uneven Cluster-Based Mobile Charging Algorithm for Wireless Rechargeable Sensor Networks," IEEE Systems Journal, pp. 1-12, Nov. 2018, DOI: 10.1109/JSYST.2018.2879084

[4] G. Han, H. Wang, J. Jiang, W. 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, Sept. 2018.

[5] H. Lu, J. Li, and M. Guizani, "Secure and Efficient Data Transmission for Cluster-based Wireless Sensor Networks," IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 3, pp. 750-761, Mar. 2014.

[6] G. Han, L. Zhou, H. Wang, W. Zhang, and S. Chan, "A Source Location Protection Protocol Based on Dynamic Routing in WSNs for Social Internet of Things," Future Generation Computer Systems, vol. 82, no. 5, pp. 689-697, Aug. 2018.

[7] H. Lu, J. Li, and H. Kameda, "A Secure Routing Protocol for Clusterbased Wireless Sensor Networks Using ID-based Digital Signature," Proceedings of IEEE Global Communications Conference, Dec. 2010.

[8] M. Bradbury, A. Jhumka, and M. Leeke, "Hybrid online protocols for source location privacy in wireless sensor networks," Journal of Parallel and Distributed Computing, vol. 115, pp. 67-81, May 2018.

[9] J. Chen, Z. Lin, Y. Hu, and B. Wang, "Hiding the Source Based on Limited Flooding for Sensor Networks," Sensors, vol. 15, no. 11, pp. 29129-29148, Nov. 2015.

[10] G. Han, X. Miao, H. Wang, M. Guizani, and W. Zhang, "CPSLP: A Cloud-Based Scheme for Protecting Source-Location Privacy in Wireless Sensor Networks Using Multi-Sinks," IEEE Transactions on Vehicular Technology, vol. 68, no. 3, pp. 2739-2750, Jan. 2019.

[11] G. Han, H. Wang, M. Guizani, S. Chan, and W. Zhang, "KCLP: A kmeans Cluster-Based Location Privacy Protection Scheme in WSNs for IoT," IEEE Wireless Communications Magazine, vol. 25, no. 6, pp. 84-90, Dec. 2018.

[12] C. Ozturk, Y. Zhang, and W. Trappe, "Source-Location privacy in energy-constrained sensor network routing," ACM Workshop on Security of Ad Hoc and Sensor Networks, pp. 88-93, Jan. 2004.

[13] J. Wang, R. Zhu, S. Liu, and Z. Cai, "Node Location Privacy Protection Based on Differentially Private Grids in Industrial Wireless Sensor Networks," Sensors, vol. 18, no. 2, pp. 410-425, Jan. 2018.

[14] A. Boualouache, S. Senouci, and S. Moussaoui, "A Survey on Pseudonym Changing Strategies for Vehicular Ad-Hoc Networks," IEEE Communications Surveys & Tutorials, vol. 20, no. 1, pp. 770-790, First quarter. 2018.

[15] Y. Gong, C. Zhang, Y. Fang, and J. Sun, "Protecting Location Privacy for Task Allocation in Ad Hoc Mobile Cloud Computing," IEEE Transactions on Emerging Topics in Computing, vol. 6, no. 1, pp. 110-121, Mar. 2018.