An Improved Approach for Probabilistic Dual Source Location Privacy Protection Scheme in WSN

Kotturi Sai Charmila[1], Gunupusala Satyanarayana[2]

[1]M.Tech Scholar, Computer Science & Technology, Sir C.R.R College of Engineering, Andhra Pradesh, India

[2]Assistant Professor, Sir C.R.R College of Engineering, Andhra Pradesh, India

Abstract- Wireless Sensor Networks (WSNs) have been widely deployed to monitor valuable objects. In these applications, the sensor node senses the existence of objects and transmitting data packets to the sink node (SN) in a multi hop fashion. The SN is a powerful node with high performance and is used to collect all the information sensed by the sensor nodes. Due to the open nature of the wireless medium, it is easy for an adversary to trace back along the routing path of the packets and get the location of the source node. Once adversaries have got the source node location, they can capture the monitored targets. In this project, we focus on the source location privacy problem in WSNs, a hot research topic in security, and propose a novel approach of two new cluster-based source location privacy protection schemes in WSNs called cluster-based dual phantom node source location privacy protection scheme (DPS) and probabilistic source location privacy protection scheme (PSLP) for WSNs. A more powerful adversary, which can use Hidden Markov Model (HMM) to estimate the state of the source, is considered in this study. To cope with this type of adversary dual phantom nodes and fake sources, which are responsible to mimic the behaviour of the source, are utilized to diversify the routing path. Then, the weight of each node is calculated as criteria to select the next-hop candidate. In addition, two transmission modes are designed to transmit real packets. We evaluate our schemes through theoretical analysis and experiments. Experimental results show that compared with other schemes, our proposed schemes are more efficient and achieves higher Security, as well as keeping lower total energy consumption. Our proposed schemes can protect the location privacy of the source node even in resource constrained wireless network environments.

Index Terms—Wireless sensor networks, source location privacy, phantom node, fake source.

I. Introduction

In recent years, WSNs have played an important role in a number of security applications, like remotely monitoring objects etc. In such applications, the location of the monitored object is tightly coupled with the sensor that detects it, called the data source. Therefore, preserving the location of data source is important for protecting the object from being traced. Such preservation cannot be simply accomplished by encrypting the data packets as the location of the data source can be disclosed by analyzing the traffic flow in WSN. There have been extensive techniques proposed to preserve source location privacy against different attack models: Local-eavesdropping model - Local-eavesdropping assumes the attacker's ability to monitor the wireless communication is limited to a very small region, up to very few hops Global-eavesdropping model - The attacker is assumed to be capable of monitoring the traffic over the entire network. Both being unrealistic, because the former stringently restricts the attacker's ability, while the latter exaggerates it, considering resources and cost required for launching such an attack. Semi-Global eavesdropping model - A more practical attack model, in this semi-global eavesdropping model, the attacker is able to eavesdrop on wireless communications in a substantial area that is much smaller than the entire monitoring network. This attack model allows the attacker to gather substantially more information than a local eavesdropper. Under the semi-global eavesdropping model, we explore a novel protocol for preserving source-location privacy by using data mules. Traditionally, data mules are used in WSNs for reducing energy consumption due to the data transmission between sensors and facilitating Communication in disconnected networks. A data mule picks up data from the data source and then delivers them directly to the base station. We adapt the functionality of data mules so that they not only maintain their traditional functionality, but also facilitate the preservation of the location privacy of data sources. Wireless Sensor Networks (WSNs) are WSN networks comprised of a large number of small and costless devices (sensor nodes) which provide traditional computers with the ability to feel and reason about their surroundings, thus providing intelligence to the environment and enabling the Ambient Intelligence (AmI) paradigm.

The reduced cost and size of sensor nodes is one of the main advantages of WSNs but it is also one of its main limitations, since it greatly constrains the capabilities of sensor nodes. These devices must cope with a processor or memory equivalent to that of computers thirty years ago. Moreover, they are mainly battery powered and in most cases these are irreplaceable. Due to the lack of resources, sensor nodes are extremely vulnerable to different types of attacks, from the hardware to the application layer In general, privacy in AmI environments has traditionally been related to what is known as social privacy, that is, the need to prevent individuals from being tracked without their explicit consent. However, there are also network privacy considerations that must be taken into consideration. An attacker might analyze the network operation in order to retrieve information about the network itself and the data being collected.

II. Literature Survey

Meanwhile, wireless sensor networks vehicles also have to be prevented from the misuse of the private information and the attacks on their privacy. There is a number of research works focusing on providing the anonymous authentication with preserved privacy in VANETs [01]. They specifically provide a survey on the privacy-preserving authentication (PPA) schemes proposed for VANETs. We investigate and categorize the existing PPA schemes by their key cryptographies for authentication and the mechanisms for privacy preservation. In wireless sensor networks, it is important to provide confidentiality to the sensor's location. In this section, we describe previous proposed technologies that were designed to preserve the source location in wireless sensor networks. For a more comprehensive taxonomy of techniques of preserving privacy in WSNs, readers may refer to the state-of-the-art survey Fan et al. [02] preserve location privacy by using homomorphic encryption operations to prevent traffic analysis in network coding. In [03], each cluster header can filter the dummy packets received from the sensor nodes of its cluster to reduce the number of dummy packets. However, the scheme requires much computation overhead due to using asymmetric-key cryptography, and the packet delivery delay is long because the cluster header sends packets with a fixed rate regardless of the number of events it collects. Mehta et al. [04] formalize the location privacy problem using a global adversary model and compute a lower bound for the overhead required for achieving a given level of privacy protection. The proposed scheme by Alomair et al. [05] can guarantee event indistinguish ability by achieving interval indistinguish ability, where the adversary cannot distinguish between the first, the middle, or the end of the interval. In [06], dummy packets can be filtered at proxy nodes, and the lifetime of the WSN is analyzed at different proxy assignment methodologies. Hong et al. [07] propose a scheme that can thwart time correlation attack. In this attack, the adversary exploits the time correlation of transmissions in successive links to learn the end-to-end route. Zhou and Yow [08] propose an anonymous geographic routing algorithm which includes three components to avoid the explicit exposure of identity and location in communication. For local-eavesdropping based attack, flooding based approach was first introduced in [10], where each sensor broadcasts data that it receives to all its neighbors. However, this technique suffers from high communication overhead for sensors. In [09], each data packet is first relayed to a randomly selected intermediate sensor in the network and then is forwarded towards base station along the shortest path. In [01], FitProb Rate is proposed to maintain source anonymity, which is an exponentially distributed dummy traffic generation scheme. The Fitprob parameter decides the dummy traffic generated at a dynamic rate, which differs from other similar works. It is a great improvement over source simulation and fake sources but still has the drawback of having overhead due to dummy packet generation.

III. System Model

In this section, the system model contains the network model and the adversary model, and assumptions are interspersed in both two parts. The background application is the protection of wild rare animals. In the wild environment, sensor nodes are randomly deployed. After being deployed, the locations of these sensor nodes keep unchanged. Then, sensor nodes monitor the acts of animals. the design features of the proposed network model and adversary model are familiarized, and assumptions are presented. The terrain of our underlying network is a finite two-dimensional grid, which is further divided into cells of equal size. The network is composed of one base station, static sensors, and mobile agents, called data mules. Static sensors - All static sensors are homogeneous with the same lifetime and capabilities of storage, processing as well as communication. They are deployed uniformly at random in the cells, and assumed to guarantee the connectivity of the network. Data mules - Data mules are the mobile agents which can be artificially introduced in the network [10]. We assume they move independently and do not communicate with each other. Also, they are assumed to know their own locations when they are moving all the time. Their mobility pattern can be modeled as a random walk on the grid, whereby in each transition it moves with equal probability to one of the horizontally or vertically adjacent cells.

After a data mule moves into a cell, it stays there for tpause time period before its next transition. At the beginning of the pause interval, the data mule announces its arrival by broadcasting Hello Message. Only data source will respond and relay buffered data to the data mule. We assume the data mule does not communicate with sensors when moving. The data mule's communication range is larger than that of a sensor, thus a data source which cannot directly transmit data to the data mule will use multi-hop routing. A. Network model : The network model in this study is based on the typical Panda-Hunter model. 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, 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.

IV. The Probabilistic Source Location Privacy Protection Scheme

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, increase 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. An overview of PSLP is shown in Fig. As mentioned in the adversary model, the adversary can use HMM to estimate the state of the source and then perform targeted search. What we need to do is to increase more possible states of the source. Phantom nodes and fake sources perfectly match our needs. 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.

A. The determination of phantom nodes

As mentioned before, phantom nodes are nodes deployed around the source to simulate the function of the source. Considering the function of phantom nodes, we can see that the longer the distance between a phantom node and the source, the stronger the privacy protection is. The main purpose of this setup is to direct the adversary away from the real source. For more details, when the source appears, it sends packets to one of its neighbors within H hops via directed random walk. Then, the neighbor sends packets to a node in its far neighbor list and decreases H by one. When H becomes zero, the current node changes into a phantom node and forwards packets sent by the source. The phantom node changes during each data transmission. In addition, the phantom node must stay outside the visible area (circle around the source). Because when the adversary backtracks to the visible area, it recognizes the source immediately. Moreover, the source sends packets to the phantom node once during the initialization. So, the transmission between the source and the phantom node is assumed to be safe. Noted that the determination of phantom nodes relates to the distance between the source and the sink, which will be presented later.

B. The determination of fake sources

As described in previous definition, fake sources are gener- ated around the sink to increase directions from where packets come. The deployment range of a fake source is specified by angle θ2 in Fig. 3. First of all, the sink divides the network into several rings. Then, these rings are divided into n sectors. For the sake of separating fake sources and the source, fake sources are only selected in the right part of the line which is perpendicular to the line linking the source and the sink. The number of fake sources is determined by the actual application. At the initialization, the fake source sequence is generated. Each fake source is preferably to stay in different sectors, which guarantees that the direction of each fake packet is different. Since the adversary knows the source state in a specific time, it needs to analyze the packet flow to find the source. Therefore, by adopting fake sources to diversify the source location, source location privacy is protected. A node acts as a fake source for a fixed period. When the time period exhausts, another fake source appears. In order to alleviate the energy consumption of fake sources, we assume that there only exists one fake source for a certain period of time.

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of PSLP. All the results provided in this section are the average values of the experimental data. A. Overview In this area, four metrics are evaluated in the simulation, namely, the safety time, the energy consumption, the network lifetime, and the transmission delay. First of all, we give the definition of each metric. The safety time is the difference between the time when the source sends the first packet and when the adversary finds the source's location. To be more specific, we use the hop count of backtracking taken by the adversary to represent the safety time. The energy consumption represents the average energy costed per simulation run. As control packets only take up very little energy, so we ignore this part and mainly focus on the energy consumption during packets transmission. The network lifetime refers to the time difference between the network establishment and the death of the first node. The transmission delay means the average packet transmission and the data processing time per simulation run. PSLP is compared with two other schemes, which are the dynamic single path routing algorithm (Dynamic SPR) and the enhanced protocol for source location protection (SLPE) [9]. Dynamic SPR uses fake sources to protect the source location, while the SLP-E adopts phantom nodes to implement this. These two methods are integrated in PSLP. Therefore, we choose Dynamic SPR and SLP-E for the comparison in NS2 tool. The Total Overhead, Transmission range and Transmission delay are shown in Fig 1, Fig 2 and Fig 3 respectively.

Fig 1. Total Overhead

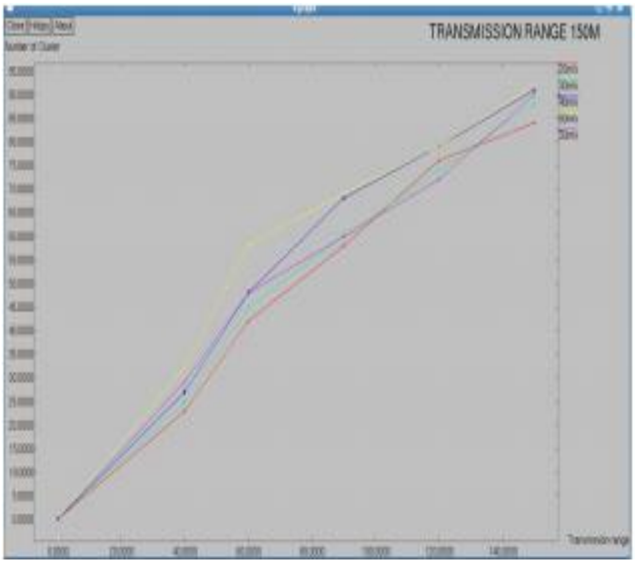

Fig 2. Transmission range

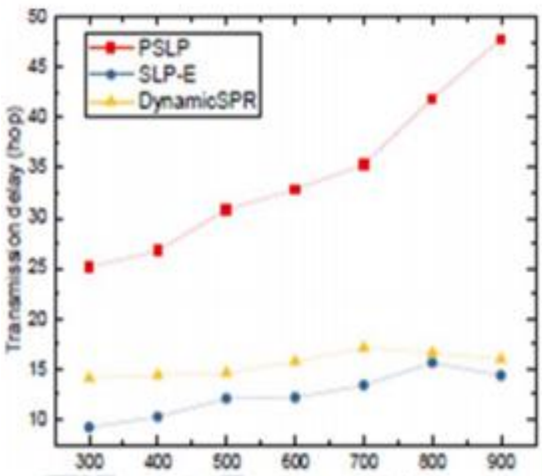

Fig 3. Transmission delay

VI. Conclusion

Studying security in WSNs became increasingly important during the last decade. In this project, we focused on the source location privacy, a research hotspot in security, and proposed a probabilistic source location privacy protection scheme (PSLP) based on WSNs. A powerful adversary which utilizes Hidden Markov Model (HMM) is considered in this study. To cope with it, phantom nodes, fake sources, and weight are adopted to change the packets' transmission directions. Considering the distance between the source and the sink, two types of routing modes are designed. Compared with Dynamic SPR 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.

References

[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] Ozturk, Y. Zhang, and W. Trappe, "SourceLocation 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] K Koteswara Chari and M Krishna, An Efficient Scalable Data Sharing in Cloud Storage Using Key Aggregate Encryption, International Journal of Science Engineering and Advance Technology, 3(11), pp: 945-946, Nov-2015.

[15] D Poul Joseph, M Krishna and K Arun, Cognitive Analytics and Comparison of Symmetric and Asymmetric Cryptography Algorithms, International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 2(3), pp: 63-68., Mar-2015.