Security Preserving Source Location Identification in Wireless Sensor Networks

Satya babu Sonti (M.Tech(SE))
Aurora’s Technological & Research Institute
satyababusonti1@yahoo.com

V. Satish M.Tech
Assistant Professor
Aurora’s Technological & Research Institute
Satish.vuyyala@gmail.com

Abstract
In wireless sensor network is deployed to monitor certain events and pinpoint their locations, the location information is intended only for legitimate users. However, an eavesdropper can monitor the traffic and deduce the approximate location of monitored objects in certain situations. We first describe a successful attack against the flooding-based phantom routing, proposed in the seminal work by Celal Ozturk, Yanyong Zhang, and Wade Trappe. Then, we propose Random Walk, a two-way random walk, i.e., from both source and sink, to reduce the chance an eavesdropper can collect the location information. We improve the delivery rate by using local broadcasting and greedy forwarding. Privacy protection is verified under a backtracking attack model. The message delivery time is a little longer than that of the broadcasting-based approach, but it is still acceptable if we consider the enhanced privacy preserving capability of this new approach. At the same time, the energy consumption is less than half the energy consumption of flooding-base phantom routing, which is preferred in a low duty cycle, environmental monitoring sensor network. The results show that our system provides high quality location monitoring services for system users and guarantees the location privacy of the monitored persons.

Keywords: Location privacy, wireless sensor networks, location monitoring system

I. Introduction

The advance in wireless sensor technologies has resulted in many new applications for military and/or civilian purposes. Many cases of these applications rely on the information of personal locations, for example, surveillance and location systems. Wireless communication had gained more popularity in recent years. The application driven force behind the popularity is easy deployment and mobility. Besides the wide applications of wireless local network today, emerging applications of wireless communication include wireless sensor networks and Mesh Networks [4]. It can be easily seen that wireless networking will gain more popularity and vast information will be carried on wireless networks in the near future. However, wireless communication media is a broadcast media, which poses a big challenge of how to protect information running on the network. Despite strong encryption of the data, wireless communication media still exposes some information about the traffic carried on the network. This is an inherent side effect of wireless communication. Mobility means that the communication is expected everywhere in the deployment area, which subsequently exposes the communication to possible attackers. Easy deployment means that there is certain openness in the protocol, which subsequently exposes some protocol information to possible attackers. Location privacy is an important security issue. Loss of location privacy can enable subsequent exposure of identity information because location information enables binding between cyberspace information and physical world entities. For example, web surfing packets coming out of a
home in a Mesh network enable an eavesdropper to analyze the surfing habits of one family if the source location of those packets can be determined. In a wireless sensor network, location information often means the physical location of the event, which is crucial given some applications of wireless sensor networks. For example, in a battlefield, the location of a soldier should not be exposed if he initiates a broadcast query. In the panda-hunter problem, the location of the panda should not be exposed to hunters [8]. A wireless sensor network can be a low duty cycle network. Often, traffic has a strong correlation with a certain event at certain time. This gives big advantages to an eavesdropper since he does not need sophisticated techniques to discriminate traffic among different events.

In this paper, we study the source location privacy problem under the assumption of one single source during a specific period. However, we need to point out that such a scenario can happen in a real wireless sensor network. To preserve location privacy, we propose to use source and sink-based random walk for packet delivery. The sink first sets up a path through random walk which serves as a receptor. Each packet from a source is then randomly forwarded until it reaches the receptor. At that point, the packet is forwarded to the sink through the pre-established path. A random walk greatly reduces the chance of packets being detected. Even if an eavesdropper happens to detect one packet, the next packet is unlikely to follow the same path, thus rendering the previous observation useless.

II. Related Work

Our work is inspired by [8, 6]. An application scenario of a wireless sensor network for monitoring a panda is presented. Enabling outside monitoring of a panda without exposing the location of the panda to hunters is proposed as the Panda-Hunter problem. Phantom routing is used for message delivery from the location of the panda to the sink for preserving its location privacy. The phantom routing algorithm is composed of two phases. In the first phase, the source initiates a random walk. In the second phase, the packet is being delivered through flooding or single path routing. In this paper, we specifically address a possible attack against the flooding-based delivery method. This paper proposes a privacy-preserving location monitoring system for wireless sensor networks to provide monitoring services. Our system relies on the well established k-anonymity privacy concept, which requires each person is indistinguishable among k persons. In our system, each sensor node blurs its sensing area into a cloaked area, in which at least k persons are residing. Each sensor node reports only aggregate location information, which is in a form of a cloaked area, A, along with the number of persons, N, located in A, where N ≥ k, to the server. It is important to note that the value of k achieves a trade-off between the strictness of privacy protection and the quality of monitoring services. A smaller k indicates less privacy protection, because a smaller cloaked area will be reported from the sensor node; hence better monitoring services. However, a larger k results in a larger cloaked area, which will reduce the quality of monitoring services, but it provides better privacy protection. Our system can avoid the privacy leakage in the example given in Figure 1 by providing low quality location monitoring services for small areas that the adversary could use to track users, while providing high quality services for larger areas. The definition of a small area is relative to the required anonymity level, because our system provides better quality services for the same area if we relax the required anonymity.
level. Thus the adversary cannot infer the number of persons currently residing in a small area from our system output with any _delity; therefore the adversary cannot know that Alice is in room R3. To preserve personal location privacy, we propose two in-network aggregate location anonymization algorithms, namely, resource- and quality-aware algorithms. Both algorithms require the sensor nodes to collaborate with each other to blur their sensing areas into cloaked areas, such that each cloaked area contains at least k persons to constitute a k-anonymous cloaked area. The idea of using intersecting paths to deliver packets has been proposed in rumor routing [1]. In rumor routing, an event is known by some sensors in the small neighborhood of the event location. A query is sent through random walk. A usable delivery ratio is achieved by a large number of query random walks intersecting with each other. This is different from our approach. In our approach, both event and query source use random walk to advertise themselves. Also, our concern is to provide privacy protection; thus a more dynamic structure than rumor routing is needed. In [10], asymptotic of three query strategies over a sensor network are discussed. Proofs are given that the probability of unsuccessful delivery using source and receiver driven ‘sticky’ Brownian motion decays much faster than using a single Brownian motion with increasing random walk length. \((t−5/8) vs (log(t)−1)\) where \(t\) is how long the Brownian motion has lasted) This result gives us a lower bound on the performance for our approach. In a real sensor network, the performance can be improved due to a limited size network. Also, in our approach, pure Brownian motion is not required for providing enough privacy protection. In [3], the problem of hiding the location of the base station in sensor networks is discussed. An attack model of determining the base station location through traffic analysis is used. To hide the traffic pattern, randomly delaying the sending time is proposed to hide the parent-child relationship given a traffic rate model. Our work instead addresses the spatial pattern of the traffic. In [5], the problem of sharing the location information without revealing the identity privacy in the mobile data collection applications, such as a cell phone periodically reporting its location, is discussed. Multi-target tracking algorithms can be used to identify each trajectory even when there is no identity information. A perturbation algorithm over multiple user paths is proposed to confuse the attacker. The algorithm takes advantage of the possible intersections of different paths and modifies location samples according to a nonlinear optimization solution.

Data Privacy

Data privacy protections target privacy of data collected by a network and queries posted to a network. There are two types of adversaries threatening the data privacy { external adversary and internal adversary. The external adversary only eavesdrops communication in a network. This kind of adversary can be easily defeated by encryption techniques such as SPINS [PST+02] or pDCS [SZZC07]. On the other hand, the internal adversary controls one or more nodes and usually has an access to encryption keys of these nodes. In such a case, the easiest way to protect privacy of data sent from nodes to the base station is to use end-to-end encryption based on keys shared between the sending node and the base station. However, such encryption makes data aggregation within the network impossible. Therefore, one of the challenges is to provide secure and privacy preserving data aggregation in the presence of an internal adversary. Multiple schemes were proposed to solve this problem [ZWF08, HLN+07].

Context privacy

Even though data privacy might be sufficiently protected, a sensor network may still leak a valuable context-oriented information. Typical context-oriented information is information on source location, sink location and timing of events. This kind of information can be usually obtained by an external adversary using traffic analysis techniques [Ray01]. We summarize state-of-the-art protections in the following subsections.

Location Privacy

Location privacy is extremely important in WSNs. Information on location of events or on location of base stations can be of a primary concern of an adversary. Suppose the Panda-
Hunter Game [KZTO05] where a WSN is employed to monitor end angered pandas in their habitat. It is sufficient for the adversary to and out location of sensors currently monitoring the panda to successfully localize and capture the panda. Similarly, the adversary only needs to and out location of the base station to be able to mount a physical or other DoS attack on the base station and thus inactivate the whole network. Problem of source location privacy was introduced and first studied by Chaum [Cha81] who proposed a mix net to hide information on a data source. Onion routing was proposed by Reed et al. [SGR97] to provide source location privacy in public computer networks. Gruteser and Grunwald proposed an approach for anonymous access to the location-based services [GG03]. This approach removes _ne details that could compromise user location privacy. Anonymous on-demand routing [KH03] was proposed to provide route anonymity and location privacy for mobile ad-hoc networks that are closely related to WSNs. However, many existing protocols are inappropriate for resource-constrained WSNs, because of requirements on memory, energy and computational power. Several protocols were designed especially for WSNs and we summarize them in this section. They are divided into four classes based on the nature of their main idea.

Temporal Privacy
In addition to location, another sensitive contextual information that can be inferred by an external adversary is timing of monitored events or a message rate. A clever adversary may abuse such information for example for victim tracking. Knowing the time and place of the message creation, she can estimate the victim movements. The problem of temporal information protection is referred to as the temporal privacy. The concept of the temporal privacy in WSNs was first defined by Kamat et al. [KXTZ07]. They have formalized the problem and proposed the Rate-Controlled Adaptive Delaying (RCAD) to protect the temporal privacy. In the RCAD, every node buffers an incoming message and randomly delays its retransmission according to the exponential distribution. Buffer preemption strategy is included to cope with the problem of overloaded buffers. When the node buffer is full, this strategy chooses a message to be transmitted immediately without further delay. Several such strategies are proposed and evaluated in [KXTZ09]. The RCAD is suitable for WSN applications where a reasonable delay can be tolerated.

GROW Algorithm
Previous analysis of random walk is based on a planar graph. However, this is not the actual communication graph in a wireless sensor network. If we treat the communication graph as a non planar graph during the implementation of the random walk, the probability of the source path and the sink path intersecting is much less than the previous asymptotic result. We use local broadcasting to solve this problem. Whenever a sensor forwards a packet, all its neighbors overhear this packet and create a route entry for the source pointing to the forwarding sensor. This does not require additional transmissions. Essentially the random walk is sticky not only for the sensors on the forwarding path but also for the neighboring sensors of this path. In effect, we build a pipe along the forwarding path. The scenario not only exists between two paths, but also exists on a single random path itself. A random path might backtrack to itself after some time. However, we would like the path to extend as far as possible and as quickly as possible. The sensor might forward the packet to one of its previous hop’s neighbors. Such a forwarding decision is not good since the random walk does not make much progress. To prevent this case, we use a Bloom filter [2] to store all current neighbors in the forwarding packet. When the next hop randomly picks up one of its neighbors, it checks whether that neighbor is already in the filter. Given a limited number of neighbors, the probability of false positives can be made very small by using a reasonable size filter within a packet. In other words, the packet will be forwarded to a sensor that has not seen the packet before with high probability. However, the potential for backtracking still exists. The only possible way to prevent backtracking is to remember all the sensors which have already seen this packet. This is not realistic for a large scale network. Currently, we did not address this issue in this paper. Instead, we rely on increasing the random walk length to increase the coverage of the path. We are working on an improved method to address this issue. To decrease the chance of backtracking, each sensor keeps a Bloom filter to store those neighbors that have already participated in the forwarding. Each time a sensor is forwarding a packet, it will store the last hop from which the packet came and the next hop which it forwards the packet to. When the random walk backtracks to a sensor, it will choose one neighbor that has never forwarded the packet before. In this-way, we hope to maximize the coverage given a fixed path length. If
the source and the sink are close to each other, the two random paths have a greater chance to intersect, thus the intersection points are closer to the source and the sink. This enables the eavesdropper to possibly trace the path. To prevent this from happening, we require a minimum path length of the source random walk. Note that we do not assume any routing infrastructure in GROW for generality. If extra information is available, we can certainly use the information to improve the performance. For example, if geographical locations of sensors are known, it is easy to identify which part of the network has not been visited. Thus a more effective greedy forwarding based on this information can be used.

III. Random Walk with Privacy

The use of random walk is desired for protecting source location privacy. A random walk does not disclose any information about the source since the forwarding decision is made locally and independent of the source location. In fact, an eavesdropper can not distinguish two random walks from two different sources. Using random walk also forces the adversary to use backtracking strategies, rendering the attack described in section 3 impossible. Since a Brownian motion path eventually hits a, it is important to see at what speed it converges to a. A recent work from Shakkottai [10] investigated this problem. The convergence is quantified as how fast the non-delivery probability decreases. It is shown that the probability decays as \((\log(t))^{-1}\), where \(t\) is how long the Brownian motion path has lasted. A more interesting result shows that if there is also a random walk from the sink at the same time, the probability of those two random walks not intersecting with each other decays as \(t^{-5/8}\), which means that it is exponentially better than using only one random walk. However, directly applying this approach is still not appropriate for practical applications because using random walk within the radio range of the eavesdropper is not useful to protect the source location privacy. For example, in , the eavesdropper can move to the sensor from which it first hears the packet. Thus, the local random walk within the eavesdropper’s radio range consumes extra energy and causes longer delivery time.

Algorithm:

1) Random Search(Desired number of Nodes N)
2) Begin
3) Probe_table=Probe()
4) If Returned_result>= N then Return Result
   Else
5) \(K= \text{Choosing}_K(\text{Probe_table}, \text{MAX}_TTL)\)
6) While Returned_Results<=N
7) Neigh=Random Choosing Unvisited Neighbr
8) D= degree(Neighbor)
9) TTL= Calculating_TTL(k,d)
10) Forwarding Query to neighbour with K and TTL
11) End While
   End if
12) Return Returned_result
   End Begin

To provide source location privacy, it is necessary to relax the requirement for the delivery time. This is because if packets are always forwarded through the shortest path, it is easy for the eavesdropper to backtrack the path. There is certainly a trade-off between privacy and delivery time.

IV. Conclusion

In this paper, we describe a possible attack against the flooding-based phantom routing. We propose random walk, a source and sink-based random walk as the alternative against this kind of attack. We improve the basic random walk by using local broadcasting and a Bloom filter. Simulation results show that it is practical to use our approach in a large scale wireless sensor network to protect source location privacy. Energy consumption is greatly reduced compared to the flooding-based phantom routing while there is only slight additional delay for message delivery. However, the delay is still acceptable. We believe that random walk is a basic approach for protecting source location privacy. However, there is still room for us to optimize the performance of this approach. Our future work is to find more efficient ways to build random paths.

References

and Privacy for Emerging Areas in Communication Networks (SecureComm), 2005.