GREEDY RANDOM BASED LOCATION PRIVACY IN WIRELESS SENSOR NETWORK

A.Deepika and R.Nirmalan

PG Scholar, Department Of Computer Science and Engineering , Sri Vidya College Of Engineering and Technology, Tamilnadu, India
Assistant Professor, Department Of Computer Science and Engineering ,Sri Vidya College Of Engineering and Technology, Tamilnadu, India

ABSTRACT

A wireless sensor network is a class of transducer which is used to monitor and record the condition of a wireless environment. The three key challenges in REAL, namely, self-organization, high accuracy, we design a state transition process, a locking mechanism and time delay mechanism, respectively. In this paper, we propose GROW, a two-way random walk, i.e., from both source and sink, to reduce the chance an eavesdropper can collect the location information. We compare the performance of GROW with current protocols through simulated experiments. The results show that GROW protects location privacy, provides more accurate query answers, and reduces communication and computational costs.

Keywords:- *Wireless Sensor Network, REAL, Greedy Random Walk, Backtracking.*

I. INTRODUCTION

Wireless sensor networks (WSNs) are new large-scale wireless networks that consist of distributed, autonomous, low-cost, low-power, small-size devices using sensors to cooperatively collect information through infrastructure less ad-hoc wireless network. Basically, location monitoring applications use sensors to gather personal locations and provide location-based services [2],[5]. However, with an reliable server, an adversary may abuse its received location information to personal sensitive information. As a result, monitoring personal locations pose privacy threats to the monitored individuals information [4], [3]. Such privacy threats, an effective way is to use k-anonymity techniques [5],[6],[7],[3]. Individuals information [4], [3].Such privacy threats, an effective way is to use k-anonymity techniques [5], [4], [7], [3]. Basically, a k-anonymity technique enlarge a person's location to a cloaked area that covers this person and at least $k-1$ other persons, so that this person is indistinguishable from the persons residing in the cloaked area. In wireless sensor networks (WSNs), this kind of cloaked areas is defined as k-anonymized aggregate locations. A k-anonymized aggregate location R is represented in a form of a cloaked area A along with the number of persons (moving objects) N residing in A, where $N \geq k$, written as R= (A,N).written as R = ⟨A, N⟩ hereafter. Previous work has defined the reciprocity property as a sufficient condition for spatial k-anonymity, i.e., the persons in a cloaked area share the same k-anonymized cloaked area [6], [7]. In the context of WSNs, the reciprocity property need that sensor nodes in the same aggregate location area share the same aggregate location, that is, each sensor node is included into one and only one aggregate location. Four steps implemented by location monitoring applications. (1)Wireless sensor nodes are deployed in the large to communicate with a small wireless transmitter worn by persons in order to determine their exact locations and identities [2], [5]. (2) Each sensor node counts the number of persons in its sensing area; note that a certain person is counted by only one sensor node, since the transmitter worn by the person maintains only one connection to a certain sensor node,

although it may be sensed by more than one sensor node. (3) Each sensor node sensing a k-anonymized aggregate location and only reports the aggregate location to the server. (4) The spatial histogram technique is employed to answer aggregate queries based on k-anonymized aggregate locations reported from sensor nodes.

In this paper, we are motivated to propose a GROW algorithm for generating k-anonymized Aggregate Locations in WSNs. The objectives of GROW are to (a) partition the whole system area into a set of aggregate locations such that each aggregate location covers at least k persons and does not overlap with any other aggregate locations, and (b) minimize the areas of aggregate locations in order to maximize their accuracy and thus provide location-based services with better quality. (C) The communication and computational costs of GROW are significantly lower than the existing spatial cloaking methods, so GROW is energy-efficient, which is essential for prolonging the lifetime of WSNs [4].

1.1 Bloom filter

A Bloom filter is a simple space-efficient randomized data structure for representing a set in order to support membership queries. Burton Bloom introduced Bloom filters in the 1970s, and ever since they have been very popular in database applications. Recently they started receiving more widespread attention in the networking literature. We then consider four types of network-related applications of Bloom filters:

1. **Collaborating in overlay and peer-to-peer networks:** Bloom filters can be used for summarizing content to assist collaborations in overlay and peer-to peer networks.
2. **Resource routing:** Bloom filters allow probabilistic algorithms for locating Resources.
3. **Packet routing:** Bloom filters provide a means to speed up or simplify packet routing protocols.
4. **Measurement:** Bloom filters provide a useful tool for measurement infrastructures used to create data summaries in routers or other network devices.

We emphasize that this simple categorization is very loose; some applications fit into more than one of these categories, and these categories are not meant to be perfect. Indeed, we suspect that new applications of Bloom filters and their variants will continue to "bloom" in the network literature. Also, we emphasize that we are providing only brief summaries of the work of many others. The theme unifying these diverse applications is that a Bloom filter offers a succinct way to represent a set or a list of items. There are many places in a network where one might like to keep or send a list, but a complete list requires too much space. A Bloom filter offers a representation that can dramatically reduce space, at the cost of introducing false positives. If false positives do not cause significant problems, the Bloom filter may provide improved performance. We call this the Bloom filter principle, and we repeat it for emphasis below.

The Bloom filter principle: Wherever a list or set is used, and space is at a premium, consider using a Bloom filter if the effect of false positives can be mitigated.

II.IMPLEMENTATION

In this section, we describe the implementation of our Greedy Random Walk Algorithm. The Grow algorithm where the sensor nodes execute the GROW model for every reporting period to generate their k-anonymized aggregate locations and send them to the server.

1.Source Node S_{ix}

2. Set of Nodes around the source SN_{ix}

3. Destination Node D_{ix}

4. Set of Neighbor Nodes λN_{ix}

5. Node Time λN_{ixt}

6. Packets Transmission ηp_{ckt}

7. Set of Path αP_{th}

8. Back Tracking B_t

9. Bloom Filter B_f

10. For ($\lambda N_{ix} = SN_{ix}$; $\lambda N_{ix} \leq D_{ix}$; λN_{ix} ++) // Random walk

11. if (λN_{ix} =1)

12. A neighbor node that has already participated in the forwarding

13. Else

14. (λN=0)

15. A neighbor node that hasn't participated in the forwarding // Greedy Condition Satisfied

16. $N\lambda_{ix} \leftarrow \alpha P_{th} + \lambda N_{ixt} + P\eta_{ckt}$ // Each Time neighbor Nodes Select Path by time

17. $B_t \leftarrow \alpha P_{th}$ //A random path might backtrack to itself after some time

18. Eavesdropper may infer

19. $B_t \leftarrow \lambda N_{ix}$

20. $B_t \leftarrow$ Store Current Node and its neighbor Information

21.End if

22. END

23.Packets reach Destination D_{ix}

III. SYSTEM MODEL

This section presents our problem, GROW's key entities (Fig. 1), communication and privacy models

3.1 Problem definition. Given an anonymity level k, Va set of moving objects o1, o2, . . . , on and a set of sensor nodes s1, s2,. . . , Sn with sensing areas a1, a2, . . . , an, respectively, the goal of GROW is to cloak the whole system area into a set of k-anonymized aggregate locations G = {Gi = ⟨Ai , Ni⟩} such that: (1) G satisfies the reciprocity property, i.e., each Gi = ⟨Ai , Ni⟩ ∈ G covers at least k objects and does not overlap with any other Gj = ⟨Aj , Nj ⟩ ∈ G, formally ∀i, Ni ≥ k and ∀i≠ j, Ai∩Aj = ∅. (2) G minimizes the average size of all areas Ai in G to maximize their accuracy and thus provide location-based services with better quality.

3.2 Sensor nodes. A sensor node, also known is a node in a sensor network that is capable of performing some processing, gathering information and communicating with other connected nodes in the network. In each reporting period, every sensor node is aware of its location and sensing area and responsible for determining the number of persons in its sensing area. All sensor nodes autonomously organize their sensing areas into a set of non-overlapping k-anonymized aggregate locations and report them to the server.

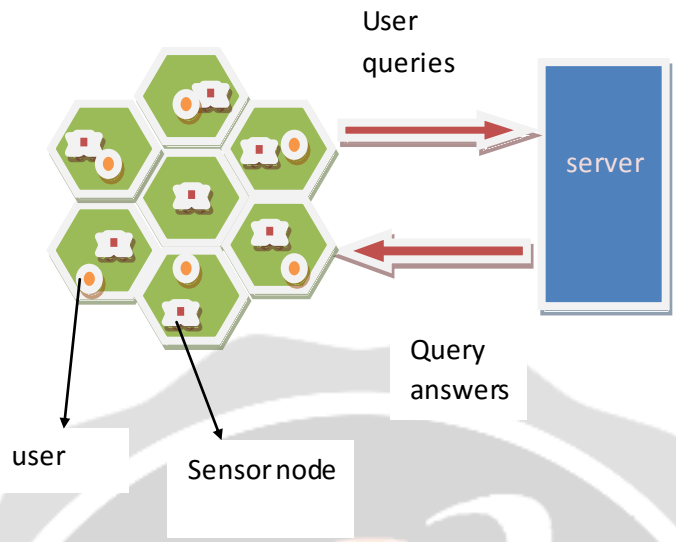


Fig. 1: The system architecture of GROW

3.3 Server: The server collects k-anonymized aggregate locations from sensor nodes, estimates distribution of monitored persons using the spatial histogram method, and provides location-based services through answering aggregate queries from users, for instance, "what is the number of persons in a certain area?" The spatial histogram divides the whole monitored area into disjointed equal-sized grid cells and maintains an estimator of the number of objects within each grid cell. Further, only the system administrator can change the anonymity level k of the system by disseminating a message with a new value of k to all the sensor nodes.

3.4 Users: Users are the persons monitored by the system. They can also issue aggregate queries to the system via the sensor nodes. The server answers the queries based on the estimated object distribution.

3.5 Communication models: By maintaining a routing table and bloom filter, a sensor node knows how to communicate with others even if the network topology is changing due to node failure. Once a sensor node receives a message of any type, it immediately confirms the receipt by sending an acknowledgement message. Thus, if a message gets lost, the source sensor node will send it again until it receives the acknowledgement message.

3.6 Privacy model: The GROW algorithm partitions the whole system area into a set of areas such that each area covers at least k persons and does not overlap any other areas. Second, through the anonymous communication techniques for communication between sensor nodes and a server, the server only knows that the sender of a k-anonymized aggregate location.

IV. A GREEDY ALGORITHM APPROACH

The greedy algorithm is expensive for solving the influence maximization problem on a large scale network. So we propose a community based greedy algorithm which mine the Influential nodes in each community rather than the whole network. Greedy random walk 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. GROW, a source and sink-based random walk as the alternative against this kind of attack. Improve the basic random walk by using local broadcasting and bloom filter. Simulation result show that it is practical to use our approach in a large scale wireless sensor network to protect source location privacy.

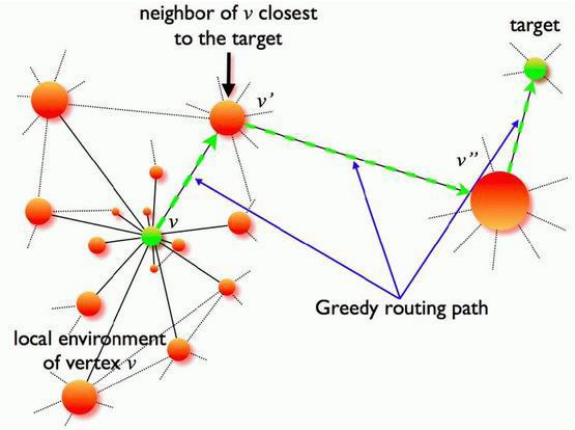
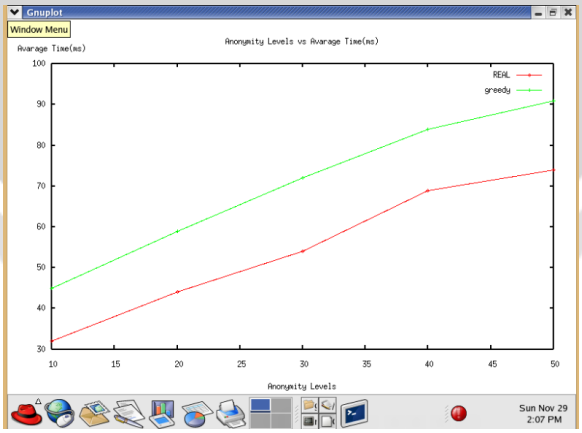


Fig. 2: GROW Model Communication

V. PERFORMANCE METRICS

We apply four popular performance metrics. (1) **Attack success ratio**. This metric measures the resilience of the k-anonymity protocols to the attacker model. Let M be the total number of sensing areas in an experiment, and m be the number of sensing areas that are derived with less than k objects by the attacker model. The ratio of successful attacks is measured as m/M. (2) **Query answer error**. This metric measures the quality of query answers, i.e., the accuracy of aggregate locations. Let N^ be the estimated number of objects within a query region using the spatial histogram, and N be the actual number of objects within the query region. If $N > 0$, the query answer error is $|N^ - N|/N$; otherwise, the error is N^. (3) **Average number of messages.** This metric measures the communication cost by calculating the average number of messages sent by each sensor node per reporting period. (4) **Average execution time.** These metric measures the computational cost (i.e., average execution time) needed for each sensor node to be part of an aggregate location per reporting period.



VI. RELATED WORK

Privacy has attracted a lot of attention in data protection [9], information retrieval [15], and especially location-based services. Approaches for preserving location privacy include enforcing privacy policies [3], anonymizing identities of data [15], location obfuscation [5], [4], space transformation [15], differential privacy techniques [7], and spatial cloaking. Among these approaches, only the spatial cloaking technique can provide aggregate location information to the server and strive for a balance between privacy protection and the quality of services by tuning privacy requirements, e.g., k-anonymity. There are mainly four reasons. (1) Both privacy policy enforcement and

identity anonymization approaches cannot prevent internal data thefts or inadvertent disclosure. (2) The location obfuscation approach reports n different locations with only one exact location and thus cannot provide high-quality monitoring services due to a large amount of false location information. (3) The space transformation approach maps the location information into another space but still reveals the monitored object's exact location information in the mapped space; hence, this technique fails to provide privacy-preserving monitoring services. (4) The large noise added by the differential privacy techniques usually dominates the true object counts in a sensing area, which will severely deteriorates the quality of monitoring services. Therefore, we apply the spatial cloaking technique to preserve the monitored object's location privacy in our location monitoring system.

Based on system architecture, current spatial cloaking techniques for k-anonymity can be classified into centralized [5], [7], distributed [8], and peer-to-peer [7], [10], [11]. (1) Centralized. The centralized approaches employ a central server that performs the aggregation on the behalf of individuals, e.g., users or sensor nodes. However, they suffer from two severe drawbacks: (a) All sensor nodes must trust the central aggregation server which is a single point of attack and poses a serious security threat. For example, if the server is compromised by an attacker, the history of all user movements may be revealed. (b) The server may become bottleneck since it must handle frequent location updates as users move. (2) Distributed. The distributed methods assume that there exist local infrastructures instead of using a global server. For instance, the techniques [8], [9] utilize base stations for users to communicate with each other and the study [8] collaboratively uses multiple servers and a third party to learn whether there are at least k persons in a certain area. None of these distributed methods is applicable to WSNs, in which there do not exist any local servers to assist sensor nodes to cloak their sensing areas into aggregate locations. (3) Peerto-peer (P2P). The P2P protocols do not depend on any global or local servers and require sensor nodes to collaborate with one another to organize their sensing areas into aggregate locations. The P2P protocols can be grouped into greedy, greedy-enhanced and random. (i) Greedy. Most current protocols [7], [10] use a greedy approach to find a cloaked area; in each step the greedy approach searches for a neighbor with the largest score calculated by a predefined function. (ii) Greedy-enhanced. In Tiny Casper [11], [12] the cloaked area obtained by the greedy approach is iteratively refined based on extra communication among the sensor nodes until it reaches the minimal possible size. (iii) Random. The study [14] employs a random method to search neighbors rather than using the greedy approach. Although these P2P protocols have been used for WSNs, they cannot satisfy the reciprocity property, since a sensor node may be involved in more than one aggregate location due to the lack of a locking mechanism.

In this paper, that we partition a whole system area into a set of non-overlapping k-anonymized aggregate locations. In WSNs, there are a few distributed clustering algorithms [1]. They are proposed for various objectives including energy saving [6], connectivity [4], management [3], fault tolerance [13], and load balancing [8], [9]. However, to the best of our knowledge, there is no research on distributed clustering algorithms for location privacy. These distributed clustering algorithms determine a set of cluster heads that are selected randomly from sensor nodes or are assumed as the special nodes with richer resources than the ordinary sensor nodes, allow the overlapping among clusters, and do not restrict the size of each cluster including the upper bound and the lower bound. Thus, none of the clustering algorithms can be applied to our problem for the following reasons: (1) In the context of privacy protection, our problem considers the k-anonymity requirement (the lower bound constraint) that is a much stricter constraint than the constraints of existing distributed clustering algorithms. (2) Our problem does not allow any overlapping between clusters (i.e., aggregation locations) to avoid location privacy breaches. (3) Our problem also aims at minimizing the spatial area size of a cluster (the upper bound constraint) to provide accurate aggregate locations. (4) Our problem chooses the sensor node with more objects as a leader (i.e., cluster head) with a higher probability, which is essentially different from the random selection method or the pre-assignment to the special sensor nodes with richer resources in the distributed clustering algorithms. Our location privacy-preserving problem is also different other privacy related problems include: (a) source location privacy that hides the sender's location and identity, (b) aggregate data privacy that preserves the privacy of the sensor node's aggregate readings during transmission, (c) data storage privacy that hides the data storage location, and (d) query privacy that avoids disclosing the personal interests.

VII.CONCLUSION

In this paper, the GROW model is proposed for privacy preserving location monitoring services in WSNs. An attack model is defined that leads to a privacy breach in existing protocols, because they generate overlapping aggregate locations. By generating non-overlapping k-anonymized aggregate locations GROW satisfies the reciprocity property to avoid this privacy breach. A state transition process is designed in GROW to accomplish self-organization among sensor nodes. A locking mechanism to guarantee the reciprocity property, and the delay mechanism to improve the accuracy of aggregate locations. By comparing with the REAL solutions, the experimental results show that GROW protects location privacy, provides more accurate query answers and saves communication and computational costs.

REFERENCES

[1] A. A. Abbasi and M. Younis. A survey on clustering algorithms for wireless sensor networks. Computer Communications, 30(14-15):2826–2841, 2007.

[2] G. Aggarwal, T. Feder, K. Kenthapadi, R. Motwani, R. Panigrahy, D. Thomas, and A. Zhu. Anonymizing tables. In ICDT, 2005.

[3] J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: A survey. IEEE Wireless Communications, 11(6):6–28, 2004.

[4] D. Baker and A. Ephremides. The architectural organization of a mobile radio network via a distributed algorithm. IEEE TCOM, 29(11):1694–1701, 1981.

[5] B. Bamba, L. Liu, P. Pesti, and T. Wang. Supporting anonymous location queries in mobile environments with privacygrid. In WWW, 2008.

[6] S. Bandyopadhyay and E. J. Coyle. An energy efficient hierarchical clustering algorithm for wireless sensor networks. In IEEE INFOCOM, 2003.

[7] J. Bao, H. Chen, and W.-S. Ku. PROS: A peer-to-peer system for location privacy protection on road networks (Demo). In ACM SIGSPATIAL, 2009.

[8] P. Belsis and G. Pantziou. A k-anonymity privacy-preserving approach in wireless medical monitoring environments. Personal and Ubiquitous Computing, accepted to appear, 2012.

[9] H. Chan and A. Perrig. Ace: An emergent algorithm for highly uniform cluster formation. In EWSN, 2004.

[10] J. Chen, H. Xu, and L. Zhu. Query-aware location privacy model based on p-sensitive and k-anonymity for road networks. In Internet of Things. 2012.