Evaluating the Privacy Measure of the Source Location Privacy Scheme in a Wireless Sensor Network

Aparna Gurjar[1], Prof. A R Bhagat Patil[2]
Dept. of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India

ABSTRACT:
Implementing Source Location Privacy Makes It Possible To Hide The Location Information Of The Transmitting Node. Classified As A Contextual Privacy Protection Technique, The Source Location Privacy Is An Essential Feature Of Those Real Life Sensor Networks Which Have Been Deployed For Monitoring Events Happening At Particular Locations. This Paper Designs A Source Location Privacy Scheme Using Cluster Based Anonymization And Random Routing. The Privacy Measure Index Is Then Evaluated In Order To Estimate The Overall Privacy Achieved By The SLP Scheme. The Effect Of The Privacy Scheme On End To End Message Delay Is Observed, For Estimating The Network Performance Degradation And Establishing The Efficacy Of The SLP Scheme.

KEYWORDS: *Anonymization, Message Delay, Privacy Measure, Routing, Source Location Privacy*

I. INTRODUCTION

The privacy threats of wireless sensor networks (WSN) can be classified into two broad categories: Content privacy and context privacy. Source location privacy comes under the purview of contextual privacy, in which the context information, like the location of the source node, its identity; or temporal information like the time and duration of transmission, is kept hidden from unintended users. Content privacy generally relies on data encryption for providing data security in WSN. This approach of encryption can become counter-productive because cryptographic algorithms are computationally intensive and can deplete the scarce energy resources of the WSN. This paper designs a source location privacy scheme which does not require encryption and uses the concepts of anonymization of node identities along with random routing of source messages in order to enable privacy. The model of the SLP scheme is first developed and then analyzed using information theoretic approach of entropy and probability. Simulation is performed to test the process of cluster formation, anonymization of node identities and random routing, and obtain relevant observations to evaluate the privacy measure index of the SLP scheme.

The section wise organization of the rest of the paper is as follows. Section 2 discusses the various context oriented privacy protection techniques. Section 3 describes the model and analyzes the proposed SLP scheme. Section 4 presents the experimental observations achieved through simulation. Section 5 concludes the paper.

II. CONTEXTUAL PRIVACY TECHNIQUES

From the study of related works in privacy protection techniques it is clear that context privacy protection can be implemented in variety of ways. In [1] Celal Ozturk et discuss the strategy of baseline flooding for implementing source location privacy. In this technique every new message is transmitted only once by each node that receives it; and every node broadcasts the message to its neighbours provided it is receiving the message for the first time. But this process consumes energy and it is also possible to backtrack to the source node. They also describe the technique of fake message flooding where fake messages are fed into the actual message traffic so that the source of the message remains unknown. In [2] Guillermo Suarez-Tangil et al. conclude that fake messaging gives excellent source location privacy, but it becomes difficult to efficiently create fake sources and define the optimal message generation rate for the fake sources. In [3] Hillol Kargupta et al. focus on the data perturbation approach to mask the data. In [4] Xiaoxin Wu uses an approach of

pseudonym, where instead of the real identity of a node, a false identity is used. The author uses positions of destinations as pseudonyms to be used in place of actual node identities. . In [5] Jing Deng et.al show that the nodes near the base station handle more number of packets than those situated away from the base station and hence can be identified as important nodes by an adversary. This helps him to move closer to the base station. Various methods are described viz. "Multi-parent routing scheme", "random walk" and "fractal propagation" which randomize the packet routes, prevent traffic analysis attack and also prevent the adversary from locating the base station or the source node.

III. THE SOURCE LOCATION PRIVACY SCHEME

The proposed SLP scheme is designed for the sensor network which is similar to the generic sensor-network application, called the "Panda-Hunter Game" in [1]. Here the sensor nodes continuously monitor the area of interest and report the presence of specific entities within its sensing range. There is an adversary trying to get the location information of those entities. The sensor nodes remain static after their random deployment. The sensing capabilities of the adversary are similar to that of the WSN nodes.

3.1. The model of the SLP scheme

The privacy scheme is implemented in two phases: The cluster based anonymization phase and the random routing phase. In the first phase the nodes are randomly deployed initially. The area to be sensed is divided into equal partitions with the number of partitions being fixed at five. Each partition represents a cluster. The process of clustering uses distance as a clustering criterion. All nodes lying within the boundary of a particular partition form a cluster and choose their cluster head (CH) randomly. The CH implements the anonymization mechanism by assigning random number to a node which needs to transmit event related information. The node then replaces its real identity with this number and then transmits its message. The mapping of, which node has been assigned which random number as the node identity is available with the cluster head. In case an adversary is able to read the message header he only gets to know the fake Identity number and not the real one. Thus this scheme prevents the Correlation-based source node identification attacks described in [6].

The primary aim of the second phase i.e. random routing is to hide the source location of the transmitting node so that the adversary is not able to detect which node started the transmission after generation of the desired event. Therefore the cluster heads after receiving the data from their respective source nodes forward the data to a randomly chosen node. Based on a random number of hops, each node forwards the data to another random node, thus forming a logical link across the network. Then finally one of the link nodes transmits it to the base station. The blending of source location information of the transmitting node with the network traffic is thus achieved and the location of the transmitting node gets diffused.

3.2. The analysis of the SLP scheme

In this section we introduce and define the various parameters associated with the SLP Scheme.

The degree of privacy: The degree of privacy $(DoP)_A$ contributed by the anonymization phase is defined as a percentage of the maximum Privacy.

The maximum privacy occurs when the adversary is not able to pin point the location of the source node and any one of the total nodes present in the WSN can be the probable source nodes. It is measured by using entropy based method described in [7], [8].

The $(DoP)_A$ is calculated as follows:

$$(DoP)_A = E_S / E_{MAX} \quad (1)$$

Where,

$$E_{MAX} = -\sum_{i=1}^{N} \frac{1}{N} * \log_2 (\frac{1}{N}) = \log_2 N$$

$$E_S = -\sum_{i=1}^{Ns} \frac{1}{Ns} * \log_2 (\frac{1}{Ns}) = \log_2 Ns$$

N= Total no. of sensor nodes in WSN

N_S = No. of nodes in the sensed area

Degree of Disclosure: Degree of Disclosure is defined as the amount of location information that one message is able to disclose to the adversary.

If the routing path is fixed then it is possible for an adversary to backtrack to the source node as the location information of each node on the routing path becomes known to the adversary due to fixed correlation between

node location and its identity. For a fixed path of length L hops, the Degree of Disclosure (DoD) will be given by (2)

$$DoD = (1/L) \quad (2)$$

If there is n number of fixed paths each having length L_1, L_2, L_3.....Ln then DoD will be computed as in [8] and given by (3)

$$DoD = (1/\{ L_1 + L_2 + L_3 + ...Ln \}) \quad (3)$$

The degree of privacy due to routing, $(DoP)_R$ is given by (4)

$$(DoP)_R = 1\text{-}DoD \quad (4)$$

Here 1 is taken as the value for maximum privacy when no message leaks any kind of location information to the adversary.

The Privacy Measure Index (PMI) describes the overall privacy achieved by the SLP scheme and is defined as the average privacy achieved as a result of anonymization and random routing. Thus PMI is given by (5).

$$PMI = [(DoP)_A + (DoP)_R] / 2 \quad (5)$$

IV. SIMULATION RESULTS

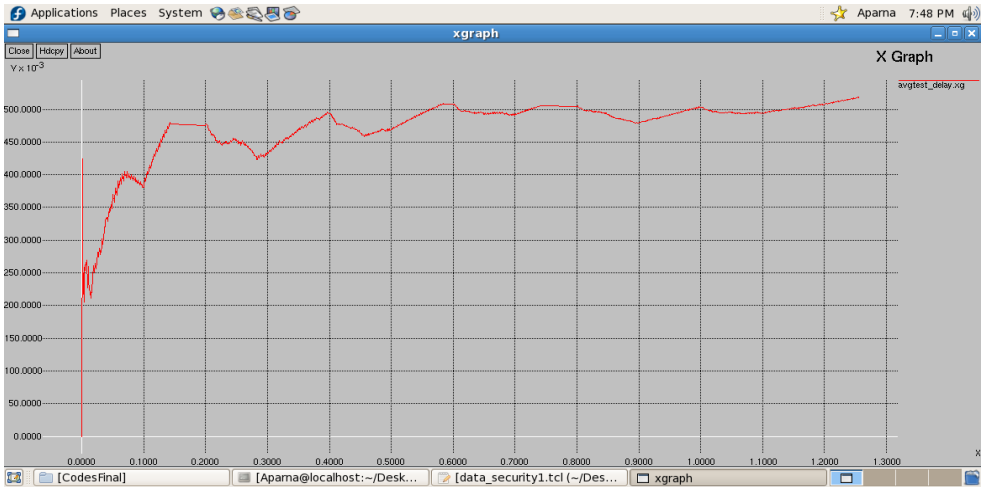

Figure1. Average end to end message delay

TABLE1. Calculating privacy measure index (PMI)

No. of Nodes	Affected nodes	E_S	E_{MAX}	Degree of Privacy $(DoP)_A = (E_S/E_{MAX})$	% Average $(DoP)_A$	No. of Hops from Source to Sink	Degree of Disclosure (DoD) for fixed path	Degree of Disclosure (DoD) for n disjoint fixed path n=4	Degree of Privacy $(DoP)_R$ - 1-DoD	Privacy Measure Index (PMI)
50	3	1.585	5.644	0.281	18.5%	7	0.142	1/27=.037 or 3.7%	96.3%	57.4%
	1	0		0		5	0.2			
	3	1.585		0.281		5	0.2			
	2	1		0.177		10	0.1			
100	4	2	6.644	0.301	20.9%	6	0.167	1/27=0.037 or 3.7%	96.3%	58.6%
	2	1		0.150		9	0.111			
	2	1		0.150		10	0.1			
	3	1.585		0.238		2	0.5			
200	6	2.585	7.644	0.338	31.4%	4	0.25	1/30 =.033 or 3.33%	96.7%	64.05 %
	5	2.322		0.304		8	0.125			
	2	1		0.131		9	0.111			
	13	3.701		0.484		9	0.111			
400	15	3.907	8.644	0.452	44.3%	6	0.167	1/25=0.04 Or 4.0%	96%	70.15 %
	13	3.701		0.428		8	0.125			
	14	3.807		0.440		4	0.25			
	15	3.907		0.452		7	0.143			

The following analysis can be done from values of TABLE 1

- The simulation was done for 4 random node configurations of 50,100, 200 and 400 nodes
- The value of maximum entropy E_{MAX} increased from 5.644 to 8.644
- The % average degree of privacy due to anonymization $(DoP)_A$ increased from 18.5% to 44.3%
  - This happens because the anonymity set for the adversary increases with increase in node density. Due to anonymization the real ID of the node is replaced by a randomly generated pseudo-identity. So, even if the adversary finds out the source ID by reading the header information of the nodes in his sensing area, he cannot find source correctly because he gets to know only the pseudo-identity number which is not linked with source location.
- Four random paths were considered for message transmission from source to sink for each configuration
- The degree of disclosure in case of fixed path routing was more than values obtained for 4 disjoint fixed paths
  - The value of (DoP) routing increases due to the increase in number of routing paths (fixed at four, in this case) and their path lengths.
  - For random routing, as the number of possible independent disjoint paths increases infinitely the (DoP) routing can effectively be considered as 1(In this case the messages leak negligible source location information)
- The degree of privacy due to routing $(DoP)_R$ remained more or less constant at a high value of around 96%
- The Privacy Measure Index increased from 57.4% to 70.15%

Fig 1 shows the plot of average end to end delay in seconds for all received messages on the Y-axis along with the event time stamp on the X-axis. Initially the delay is less, after that it increases within a short period of time and finally stabilizes to almost constant value for the rest of the transmission period. The average message delay was measured as 0.5 msec. There is an increase in average end to end message delay by a factor of 10, compared to a scheme which does not implement SLP. This is expected as the concept of shortest path, which is implemented in most routing algorithms, is not adhered to in this scheme, in order to diffuse the source location information.

V. CONCLUSION

This SLP scheme is effectively able to maintain the location privacy of the transmitting nodes using anonymization and routing technique and gives high degree of privacy values as seen by the simulation results. There is degradation in network performance in terms of end to end message delay, which is acceptable if the requirement of privacy protection is of paramount importance for the sensor network application. This scheme also prolongs network life time as it does not use the computationally intensive process of encryption.

REFERENCES

[1] Source-Location Privacy in Energy-Constrained Sensor Network Routing Celal Ozturk, Yanyong Zhang, Wade Trappe *SASN'04, October 25, 2004, Washington, DC, USA. Copyright 2004 ACM 1-58113-972-1/04/0010*

[2] An Experimental Comparison of Source Location Privacy Methods for Power Optimization in WSNs Guillermo Suarez-Tangil, Esther Palomar, Benjamin Ramos, Arturo Ribagorda *in Advances in sensors, signals and materials, ISSN: 1792-6211 / ISSN: 1792-6238, ISBN: 978-960-474-248-6*

[3] Random data Perturbation Techniques and Privacy Preserving Data Mining Hillol Kargupta, Souptik Datta *Extended version of paper presented in 2003 IEEE International Conference on Data Mining*

[4] Xiaoxin Wu,"Applying Pseudonimity for Anonymous Data Delivery in Location-Aware Mobile *Ad Hoc* Networks" *IEEE Transactions on Vehicular Technology, Vol. 55 NO. 3, May 2006*

[5] DecorrelatingWireless Sensor Network Traffic To Inhibit Traffic Analysis Attacks Jing Deng Richard Han Shivakant Mishra *Elsevier Pervasive and Mobile Computing Journal, Special Issue on Security in Wireless Mobile Computing Systems, vol 2, issue 2, April 2006, pp. 159-186*

[6] Yun Li, Jian Ren, and Jie Wu, "Quantitative Measurement and Design of Source-Location Privacy Schemes for Wireless Sensor Networks" *IEEE Transactions on Parallel and Distributed Systems, Vol. 23, No. 7, July 2012*

[7] Towards an information theoretic metric for anonymity Andrei Serjantov, George Danezis

[8] Source-Location Privacy in Wireless Sensor Networks Song-Woo Lee, Young-Hun Park, Ju-Hyung Son, Seung-Woo Seo, U Kang, Ho-Kun Moon, and Myoung-Soo Lee *Korea Institute of Information Security and Cryptology Journal 2007*

[9] Wireless Sensor Network Security: A Survey John Paul Walters, Zhengqiang Liang, Weisong Shi, and Vipin Chaudhary in Security in Distributed, Grid, and Pervasive Computing Yang Xiao,(Eds.) pp. 2006 Auerbach Publications, CRC Press

[10] Protecting Location Privacy in Large-Scale Wireless Sensor Networks Lei Kang ICC 2009 proceedings