Replica Node Detection in Static Wireless Sensor Networks using Neighborhood Information

Ali Reza Khodadadi
Department of Computer Engineering, Baft Branch, Islamic Azad University
Baft, Iran

Fereshteh Khodadadi
Department of Computer Engineering, Malayer Branch, Islamic Azad University
Malayer, Iran

Mozhgan Khodadadi
Department of Computer Engineering, Malayer Branch, Islamic Azad University
Malayer, Iran

ABSTRACT
This study proposes a novel, simple, and efficient algorithm using explorer nodes to overcome the replica node attack in static wireless sensor networks. In the replica node attack, the adversary captures a legitimate node in the network and extracts its important information, including its ID, to generate and inject several replica nodes in the network. These nodes are controlled by the adversary and have the ability to link legitimate nodes. Therefore, these replica nodes can easily have their corruptive impacts on network integrity. The main notion of the proposed algorithm is to collect spatial and neighborhood information by mobile explorer nodes in the network environment to detect replica nodes. The proposed algorithm consists of two parts: 1- recording information in the buffer of explorer nodes, and 2- verifying buffer content to detect potential replica nodes. The proposed algorithm is implemented and its efficiency is evaluated by a set of experiments in terms of replica node detection rate. Furthermore, evaluation results are compared with existing algorithms and indicate that the proposed algorithm outperforms other methods.

Keywords
Static sensor networks; replica node attack; neighborhood information; explorer node

1. INTRODUCTION
Wireless sensor networks are ad hoc networks, which consist of hundreds to thousands of small and cheap sensor nodes. These networks have various applications in the military, medicine, health, and other sciences and are mostly appropriate to study environments, where human presence is dangerous and costly. Sensor nodes have many limitations in terms of memory, computational power, radio range, and energy. According to these limitations and considering the unattended deployment of sensor nodes, the wireless communication nature of these networks, and their increasing application in military domains, security of these networks has become an important and challenging issue and the focus of many researchers.

Node replication or node replica attack is one of the dangerous attacks in wireless sensor networks. Considering the unattended deployment of nodes in the operational environment, the adversary can capture one (or more) legitimate nodes in the network and extract its important information, e.g. keying materials to generate replicated (or replica) nodes. Since replica nodes contain the exact specifications and information (e.g. ID, keying materials, etc.) of the captured legitimate node, they can establish keys with other legitimate nodes in the network. After deploying in the network, replica nodes can also use this inner network position to initiate different attacks [4] [5] [3] [6].

So far, there have been many algorithms (e.g. [7-12]) to overcome the aforementioned replica node attack. Most of these algorithms are based on sending spatial claim messages to nodes or witness locations in the network. Thee algorithms have three fundamental problems, which are very high memory overhead, high communication overhead, and low replica node detection rate.

This study proposes an algorithm based on neighborhood information and explorer nodes to detect replica nodes in static wireless sensor networks to eliminate the weaknesses of previous algorithms. The rest of this paper is organized as follows. Section 2 discusses previous works, section 3 presents system assumptions, and section 4 introduces the proposed algorithm. Section 5 evaluates the performance and provides simulation results and finally, section 5 presents the conclusions.

2. RELATED WORK
Four algorithms, called NNB, DM, RM, and LSM are proposed [7], which employ public key encryption. Another centralized algorithm, called RED, is proposed to overcome the replica node attack [8]. RED is executed in fixed time periods. This algorithm consists of two stages. In the first stage, a random value r is shared among all nodes. This can be centralized (e.g. by a satellite or base station) or distributive (e.g. by leader nodes, which are selected in a distributive manner). In the second stage (i.e. the detection phase), each node digitally signs (by a private key) and propagates its claim, which includes its ID and geographical location. Moreover, RED has been reviewed and its feasibility has been assessed further [9]. In fact, using analysis and another set of simulations, this review shows that RED can be implemented in real wireless sensor networks.

Another protocol, called SET, is proposed to detect replica nodes [10]. The main notion of this algorithm is inspired by the observation that a sensor network can be modeled by non-overlapping areas. Four replica node detection protocols, B-MEM, BC-MEM, C-MEM, and CC-MEM, are also proposed [11], which are based on sending location claim messages. Furthermore, two distributive solutions, called UTLSE and MTLSE, are proposed for replica node detection in mobile sensor networks [12]. The main notion of these two algorithms is using the mobility characteristic of nodes, as well as their temporal-spatial claims. Moreover, two algorithms, called RAWL and TRAWL, are proposed in [13]. In RAWL, for each node u, several random hops are navigated in the network and the passed nodes are selected as witnesses of node u. TRAWL is based on RAWL, but adds a trace table to each node to reduce memory costs.

Another protocol is proposed in [14] to overcome the replica node attack. This protocol utilizes a symmetric polynomial to establish pairwise keys [17] and a group-based deployment model. In this protocol, sensor nodes are deployed in separate groups or generation in the environment. The main notion of this protocol is that using a symmetric polynomial, each deployed sensor node can be linked to a unique generation or group, to which it belongs. More specifically, even if the adversary captures a node and generates replicas, these replica nodes will belong to the same group to which the captures node belongs. Two other algorithms, called SDC and P-MPC, are proposed based "local multicasting" or LM to detect replica nodes [16]. These algorithms work in sensor networks with the grid topology.

3. SYSTEM ASSUMPTIONS

A sensor network consists of two sets of normal sensor nodes (SN) and explorer nodes (EN). The total number of nodes in the network is n=EN+SN. Nodes are randomly distributed in a two-dimensional area. Normal sensor nodes are responsible for the network mission. Whereas, explorer nodes are responsible for detecting replica nodes. After being deployed in the environment, normal sensor nodes remain static and are not aware of their location. However, explorer nodes are equipped with GPS and after deployment, they can move in the network environment based on mobility models, including the random waypoint model. Each node has a unique ID and a constant radio range of r. Moreover, it is assumed that the sensor network is deployed in a hostile environment. Therefore, the network is unsafe and the adversary can capture nodes, generate replicas from them, and inject them in the network.

4. THE PROPOSED ALGORITHM

Considering the limitations of sensor nodes in terms of processing power, memory size, energy, etc., a desirable and applicable algorithm for such networks should be simple enough to impose little memory, processing, and communication overheads on these limited resource sensor nodes. The main notion of the proposed algorithm is collecting spatial and neighborhood information by mobile explorer nodes in the network environment to detect replica nodes. In sum, in the proposed algorithm, there are a number of explorer nodes, which explore in the network environment and collect necessary information from different points of it to detect replica nodes.

In contrast to most existing methods, in the proposed algorithm, only explorer nodes are equipped with GPS. Each explorer node has a buffer with size S, which can store items as figure 1.

NodeID	X-Pos	Y-Pos

Figure 1. The template of the data stored in explorer node buffers

After deploying nodes in the network, explorer nodes start exploring, collecting information, and detecting replica nodes. The proposed algorithm consists of two phases: 1- recording information in the buffer of explorer nodes and 2- verifying the buffer content to detect potential replica nodes. In what follows, the details of these two stages are presented.

4.1 Information Recoding Phase

After deploying nodes in the network, each explorer node u selects a random destination P (Xp, Yp) and starts to move towards it. After reaching location P, u broadcasts a request message and asks each node in its neighborhood, i.e. P, to return an ACK message. Each sensor node v, which receives this message, responds with an ACK message. Explorer node u then inserts a data item corresponding to v, including <<v, Xp, Yp>>, in its buffer. After time period t, explorer u runs the verification procedure and then selects another random destination and moves towards it.

This process is periodically executed (for R times). At this stage, malicious replica nodes cannot remain hidden and refuse to send an ACK message, since all legitimate neighbors will realize it and remove them from their neighborhood table. Usually, in all scenarios of wireless sensor networks, each node has a table, called the neighborhood table, which stores the IDs of its neighbors. Moreover, the buffer of explorer nodes is considered cyclic. It means that if the buffer is full, new data replaces the older ones.

4.2 Verification Phase

The verification phase is also performed periodically each time the explorer nodes move to a new destination and record neighborhood information in their buffers. Each explorer node navigates its buffer and if it finds two data items, e.g. <<v, Xp, Yp>> and <<v, Xq, Yq>> for a particular node, e.g. v, and equation (1) is satisfied, it considers this as the presence of a replica node in the network.

$$\sqrt{(Xp-Xq)^2+(Yp-Yq)^2} > \alpha \quad (1)$$

Where, α is a threshold adjusted in proportion to radio range (r) ($\alpha \geq r$). This equation shows that a particular node v is simultaneously present at two different areas of the network, which indicates that it has replica nodes in the network.

5. PERFORMANC EVALATION AND SIMULATION RESLTS

This section evaluates the performance of the proposed algorithm, present simulation results, and compares the results with those of existing algorithms.

5.1 Memory Overhead

Considering the memory limitations of sensor nodes, a desirable and applicable algorithm is one with low memory consumption. Therefore, this section evaluates the proposed algorithm in terms of memory consumption and compares it with other approaches. The proposed algorithm imposes no memory overhead to normal sensor nodes and only poses a memory overhead of S for the buffers of explorer nodes. Table 1 compares the memory overhead of the proposed algorithm with that of existing approaches.

5.2 Communication Overhead

At each iteration of the proposed algorithm, each explorer node broadcasts a request message for its neighbors and normal sensor nodes, which are located at its neighborhood, respond with an ACK message. Therefore, the communication overhead of the proposed algorithm for an iteration of the proposed algorithm is at most O(n) and for R iterations is O($n \times R$). Table (1) compares the communication overhead of the proposed algorithm with that of previous methods.

Table 1. Comparison of the proposed algorithm and other methods in ter5ms of memory and communication overheads

Algorithm	Memory overhead	Communication overhead
LSM	$O(\sqrt{n})$	$O(n\sqrt{n})$
SET	$O(\frac{n}{T})$	$O(n)$
P-MPC, SDC	$O(w)$	$O(r\sqrt{n})+O(s)$
RED	$O(d)$	$O(n\sqrt{n})$
RAWL	$O(\log n\times\sqrt{n})$	$O(\log n\times\sqrt{n})$
Proposed Algorithm	$O(S)$	$O(n\times R)$

5.3 Simulation Model

C++ programming language is used to implement the proposed algorithm. The simulation model is as follows.

- The simulations assume that the network has n sensor nodes, which are randomly deployed in a two-dimensional area.
- The adversary generates and deploys M replica nodes in the network.
- The transmission range of node is r=10 at the first 3 experiments. However, it is changed for the rest of them.
- The threshold is considered $\alpha = r = 10$.
- The mobility model of explorer nodes is random and their maximum speed is 20m/s.
- The buffer size of explorer nodes for the first 3 experiments is S=100 and for the rest of them is S=300.
- The number of explorer nodes in the network is EN.
- In order to insure the credibility of results, each simulation is repeated 500 times and the final result is obtained by averaging these 500 repetitions.

5.4 Experimental Results

First experiment: this experiment aims to investigate the effect of parameters M and R on the performance of the proposed algorithm. In this experiment, parameters are set to n=300 and EN=5. The number of replica node generated from each captured node is M=2~5 and the number of algorithm executions is R=25~200. Figure 2 presents the results. Results of this experiment shows that increasing the number of algorithm executions and the number of replica nodes also increases the detection rate. It is clear that replica node detection rate is increased by increasing the number of algorithm executions and thus, the number of explorer node moves. The reason is that explorer nodes navigate a larger number of areas in the network environment and thus, it is more likely to detect all replica nodes. Moreover, a larger number of replica nodes in the network mean that a node with a specific ID is simultaneously in different areas of the network. Therefore, explorer nodes can detect replica nodes more rapidly.

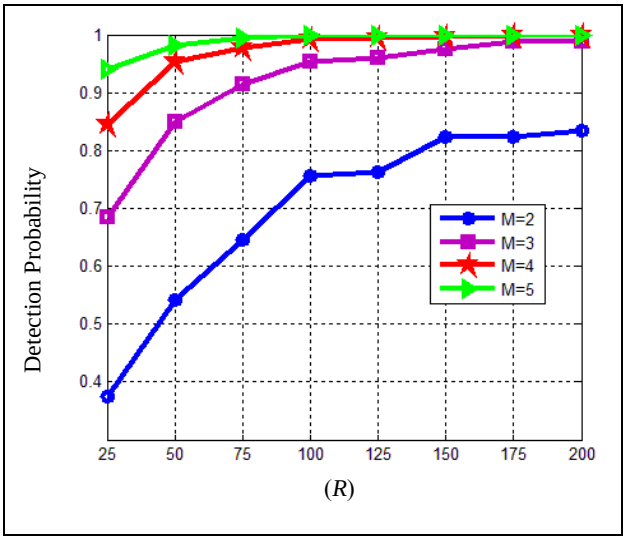

Figure 2. The effect of parameters *R* and *M* on the performance of the proposed algorithm in terms of detection rate

Second experiment: this experiment aims to investigate the effect of EN (the number of explorer nodes) on the performance of the proposed algorithm. In this experiment, parameters are set to n=300 and M=4. The number of explorer nodes is EN=3~6 and the number of algorithm executions is R=25~200. Figure 3 presents the results. Experimental results show that increasing the number of explorer nodes in the network also increases the detection rate of replica nodes. The reason is very clear, since, increasing the number of explorer nodes allows navigating a larger portion of the network, which increases the detection rate.

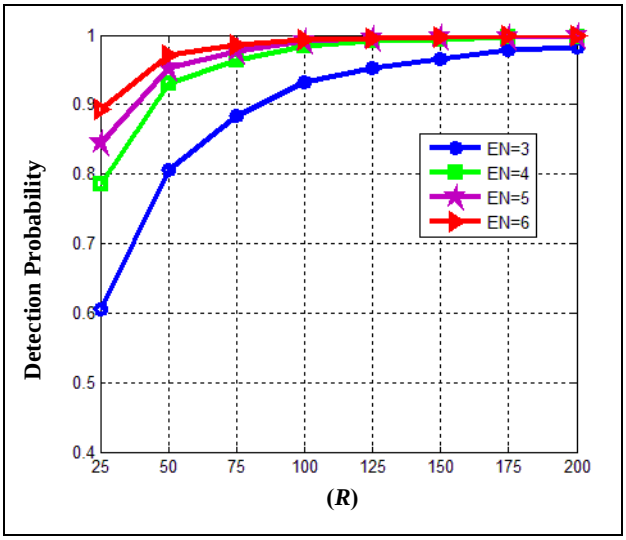

Figure 3. The effect of parameter *EN* on the performance of the proposed algorithm in terms of replica node detection rate

Third experiment: this experiment aims to investigate the effect of n (the number nodes) on the performance of the proposed algorithm. In this experiment, parameters are set

M=4, *EN*=5, and *R*=100. Moreover, the number of nodes is *n*=100~800. Figure 4 presents the results.

Experimental results show that increasing the total number of nodes in the network slightly decreases the detection rate of replica nodes. The reason is that increasing the number of nodes also increases the density of the network and thus, the average number of neighbors of each node. This frequently replaces the data in the buffer of explorer nodes with new ones and thus, reduces their lifetime. Therefore, the likelihood is reduced that there are two different data items about a particular node, e.g. *v*, is in the buffer of an explorer one. That reduces the detection rate of replica nodes.

Fourth experiment: this experiment aims to compare the proposed algorithm and several other methods in terms of detection rate. In this experiment, total number of nodes is *n*=1000~5000. Moreover, the number of explorer nodes is *EN*=20, *S*=400, *R*=500, and the most difficult establishment of replica nodes is considered, i.e. *M*=2. Furthermore, the radio range of nodes is adjusted for each node to have about *d*=20 neighbors. Table 2 presents a list of the evaluated algorithms, as well as the corresponding experimental results. As it is shown, in most cases, the proposed algorithm provides more desirable and superior results. Only when the number of nodes is too high, i.e. more than 4000 or 5000, the detection rate of the proposed algorithm is reduced. Of course, under such circumstances, increasing the number of explorer nodes, buffer size (*S*), or *R* can maintain a desirable performance for the proposed algorithm.

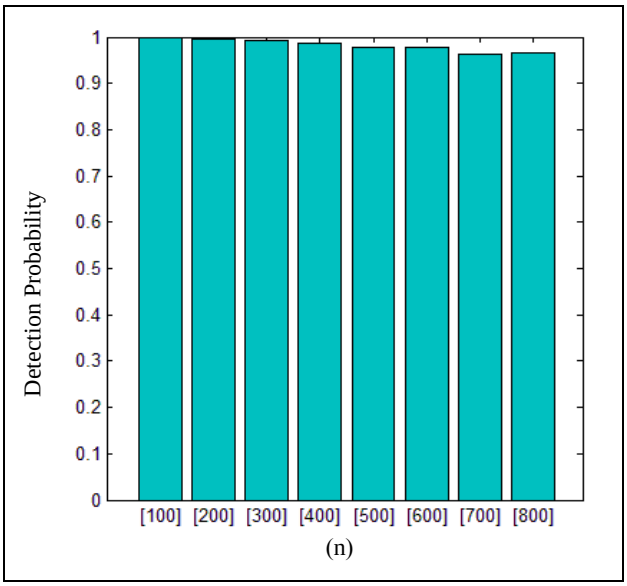

Figure 4. The effect of the number nodes on the performance of the proposed algorithm in terms of replica node detection rate

Table 2. performance Comparison of the proposed algorithm, SDC, P-MPC, and LSM in terms of the replica node detection rate

Algorithm	Detection Probability				
	n=1000	**n=2000**	**n=3000**	**n=4000**	**n=5000**
LSM	0.89	0.89	0.81	0.86	0.86
B-MEM	0.86	0.86	0.86	0.81	0.83
BC-MEM	0.93	0.93	0.97	0.96	0.93
C-MEM	0.95	0.93	0.93	0.89	0.95
CC-MEM	0.99	0.98	1	0.99	1
Proposed Algorithm	1	1	0.9	0.88	0.86

6. CONCLUSIONS

This study proposed an algorithm based on neighborhood information and explorer nodes to detect replica node in wireless sensor networks. The main notion of this algorithm is that explorer nodes roam the network and collect neighborhood information to detect replica nodes. The proposed algorithm was simulated and its performance was evaluated by several experiments. Experimental results were compared with several other algorithms, which indicate the superiority of the proposed algorithm. Future works aim to utilize explorer nodes to overcome Sybil attack, as well as the replica node attack in mobile wireless sensor networks.

7. REFERENCES

[1] Akyildiz Ian F. and Kasimoglu Ismail H., "Wireless sensor and actor networks: research challenges", in: Proceedings of the Ad Hoc Networks 2, pp. 351–367, 2004.

[2] Yick J., Mukherjee B. and Ghosal D., "Wireless sensor network survey", in: Proceedings of the Computer Networks 52, pp. 2292–2330, 2008.

[3] Karlof C. And Wagner D., "Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures", in: Proceedings of the Ad Hoc Networks, pp. 299-302, 2003.

[4] Walters J.P., Liang Z., Shi W. and Chaudhary V., "Wireless Sensor Network Security: A Survey", in: proceedings of the Distributed, Grid, and Pervasive Computing, Vol. 1, Issue 2, CRC Press, pp. 1-50, 2007.

[5] Sharma K. and et al., "A Comparative Study of Various Security Approaches Used in Wireless Sensor Networks", in: Proceedings of the International Journal of Advanced Science and Technology, Vol. 17, April, 2010.

[6] Manjula V. and Chellappan C., "REPLICATION ATTACK MITIGATIONS FOR STATIC AND MOBILE WSN", in: Proceedings of the International Journal of Network Security & Its Applications (IJNSA), Vol. 3, No. 2, March 2011.

[7] B. Parno, A. Perrig, and V. D. Gligor. "Distributed Detection of Node Replication Attacks in Sensor Networks", IEEE Symposium on Security and Privacy, 2005.

[8] M. Conti, R. D. Pietro, and L. V. Mancini, "A Randomized, Efficient, and Distributed Protocol for the Detection of Node Replication Attacks in Wireless Sensor Networks", In Proc. of ACM MobiHoc, September 2007.

[9] M. Conti, R. D. Pietro, L. V. Mancini, and Alessandro Mei "Distributed Detection of Clone Attacks in Wireless Sensor Networks", IEEE TRANSACTIONS ON DEPEDABLE AND SECURE COMPUTING, 2010

[10] H. Choi, S. Zhu, and T. F. La Porta. "SET: Detecting Node Clones in Sensor Networks", In SecureComm '07, pages 341–350, 2007.

[11] Zhang M., Khanapure V., Chen S., Xiao X., "Memory Efficient Protocols for Detecting Node Replication Attacks in Wireless Sensor Networks", in: Proceedings of the IEEE International Conference on Network Protocols (ICNP), 2009.

[12] Deng X., Xiong Y., and Chen D., "Mobility-assisted Detection of the Replication Attacks in Mobile Wireless Sensor Networks", in: Proceedings of the IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications, 2010.

[13] Yingpei Zeng, Jiannong Cao, Senior Member, IEEE, Shigeng Zhang, Shanqing Guo and Li Xie,"Random-Walk Based Approach to Detect Clone Attacks in Wireless Sensor Networks", IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 28, NO. 5, JUNE 2010.

[14] C. Bekara and M. Laurent-Maknavicius, "A new protocol for securing wireless sensor networks against nodes replication attacks", in WIMOB '07: Proceedings of the Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications. Washington, DC, USA: IEEE Computer Society, 2007.

[15] Thakur G., "CINORA:Cell Based Ientification of Node Replication Attack in Wireless Sensor Networks", www.cise.ufl.edu/~gsthakur/cinora.pdf

[16] B. Zhu, V. G. K. Addada, S. Setia, S. Jajodia, and S. Roy, "Efficient Distributed Detection of Node Replication Attacks in Sensor Networks", Annual Computer Security Applications Conference (ACSAC), December 2007.

[17] R. Blundo, A. D. Suntis, A. Herzbeg, S. Kutten, U. Vaccaro and M. Yung, "Perfectly secure key distribution for dynamic conferences", In Proc. of the 12th Annual International Cryptology Conference on Advances in Cryptology, Lecture Notes in Computer Science, vol. 17, Springer-verlag, pp. 471-486, 1992.