An Overview of Replica Node Detection in Wireless Sensor Networks

V.Ram Prabha
Ph.D Scholar, PET Engg. College,

P.Latha
Asso.Prof.,Govt.College of Engg.,Tirunelveli

ABSTRACT
Wireless Sensor **Networks** deployed in hostile environment are vulnerable to many attacks due to the nature of deployment and poor physical security. An attacker can take control over the network by compromising few nodes in the network either by extracting the secret keys from the sensor node or reprogram the sensor nodes. After compromising the sensor nodes, making the replicas of them and placing them back to the network , an adversary may perform various types of attacks on the network .Replica node attack is a very hazardous one and must be detected as early as possible to protect the network from those attacks. Many schemes have been proposed for detecting the replica nodes in centralized and distributed environment. In this paper we review these schemes and comparing their performance.

Keywords
Wireless sensor network, replica node, SET, RED, EDD

1. INTRODUCTION
Wireless sensor network is a collection of sensor nodes with limited energy, memory and processing capabilities. Due to the nature of deployment they are unattended. So an adversary can capture a sensor node easily and compromise it to get the keying materials and the program inside the node. Providing security to an unattended node is a critical one. The most hazardous attack in this scenario is a node replication attack. By capturing a single sensor node, the adversary can create as many replicas as he has the hardware. Time taken for placing the replicas should be less than the time and effort taken for capturing and compromising the nodes.

These replica nodes are placed again in the network for more malicious activities. Finding a replica and recover from the replica node attack in the network is an essential one for providing security to the sensor nodes. Detection of a replica node is not so easy since they have the legitimate keys which make them to consider as legitimate member of the network.After compromising a sensor node, an adversary can perform various attacks on the network in many ways. He can simply listening the traffic flow to gather information that passes through the nodes. He can perform jamming attack so that the legitimate signals cannot be transmitted. Alternatively he can inject some false information to corrupt the sensor node's operation or he can change the various network protocols for formation of clusters and then disabling the functions of the network. Several schemes have been proposed by different people for detecting replica node in static and mobile sensor networks. In this paper we combine those schemes and analyze them.The remainder of this paper is organized as follows. Section II describes about replica node attack. Section III presents an overview of existing schemes for detecting replicas in static sensor networks in both centralized and distributed environment. Section IV presents an overview of existing schemes in mobile sensor networks. Section V discusses the performance and overhead of these schemes. Section VI describes the discussion on the results. At last section VII describes the conclusion and future research extension to replica node detection.

2. REPLICA NODE ATTACK
In replica node attack, an adversary may capture the node and take the data into his own sensor. Then he deploys those sensors in to the network for various malicious activities. Replica node attack is a dangerous one since all the replica are having legitimate keys which makes the replica to be an benign node. Since there is no difference between the benign node and replica in terms of their authentication it is difficult to detect replica. Several researchers [10][11][12] showed a number of attacks that can be made on the network using replica nodes. Once a node is compromised the information get leaked, the adversary may inject false data on the nodes or modifying the data which is passed between the nodes. So finding a replica node is an important one for protecting the network from various attacks. Protection of sensor networks can be done in two ways. Both centralized and distributed approaches are needed. And also approaches needed for static sensor networks and wireless sensor networks.

3. DETECTION OF REPLICA IN STATIC SENSOR NETWORKS

3.1. Centralized schemes for detection of replica nodes

3.1.1. Simple Approach
In a simple Centralized approach, the Base Station (BS) acts as centralized entity, each node sends a list of its neighbour nodes and their claimed locations to a base station. If the base station finds that there are two far distant locations for one node ID, then the node clone must have occurred. The BS simply broadcasts through the whole network to expel the cloned nodes. Then, the BS will revoke the replicated nodes. This solution has several drawbacks, for instance: Single point of failure (BS) or any compromise to BS, and high

communication cost due to the relevant number of exchanged messages.

3.1.2. SET: Detecting node clones

A simple way of finding the replica [7] is that each node sends its authenticated report and its neighbour's to the base station. But this scheme has high communication overhead due the repeated information send to the base station. Witness based schemes [8][9] are based on public key cryptography and that is not suitable in case of wireless sensor networks.

In SET [1] a sensor network is spited into several non overlapping regions. All nodes are having distinct identifiers. Since each node has a different identifier the intersection of those regions will give an empty set. When the node has a replica the intersection will not be empty and the replica node can be detected.

3.1.3. RED: A Randomized, Efficient and Distributed protocol for the detection of node replication attacks

Each and every node in the network known its location and the nodes are static nodes and they use public key crypto systems [2]. In this a random value r is shared between the nodes with centralized broadcasting. Each node signs its claim and sends it to neighbours.

Set of witness nodes is formed using pseudo rand function. This takes the arguments ID, current r value, and the number of locations that have to be generated. The ambiguity is verified for the witnesses for the claim. Each claim is signed with the private key of that node. For each claim received by the witness node it verifies the signature first. To check for the validity of the message the coherence between the time inserted in the message and the current time is verified.

3.1.4. Real Time detection of replica node attacks

Detection of replica node is done with computation of finger print for each sensor node based on the characteristics of the neighbourhood [4]. Then verification of this finger print is done at base station and the neighbouring sensors.

Before deployment a superimposed s-disjunct code C is pre computed .A binary matrix C defines an s-disjunctive code if and only if the Boolean sum of any s-subset of columns of C does not cover any other column of C that are not in the s-subset. After deployment each sensor sends a codeword to the neighbourhood. Then it calculates the fingerprint for each node that sends their codeword and monitoring the messages sent in the neighbourhood. After getting the data the sensor calculates the fingerprint of the neighbour from the code word collected.

After calculation the sensor verifies that with the already stored one. If there is a mismatch then it sends an indication to base station. BS sends a query to the neighbourhoods to get the details of their fingerprint. Then BS decides which sensor should be revoked.

Alternatively BS itself finds the replica without any help from the sensors. BS has details of fingerprint with sensor ID's. It collects the message from the sensor and verifies its signature. If it does not match, then replica is detected.

3.1.5. Detection of Clones using Random Key Distribution

Keys that are present on the cloned nodes are detected by looking at how often they are used to authenticate nodes in the network[13]. Detection is by analyzing node authentication statistics. Each node randomly selects k keys from pool. Key usage refers to the number of times a key is used to set up connections between neighbouring nodes.

When clones are inserted into the network, the key usage distribution is skewed. Cloned keys are present on a greater number of nodes than normal and are therefore used more frequently than keys that have not been cloned. By collecting key usage statistics, we can determine which keys have been cloned.

3.2. Distributed schemes for replica node detection

3.2.1. Node-to network Broadcasting

This detection approach utilizes a simple broadcast protocol. Basically, each and every node in the network
uses an authenticated broadcast message to flood the network with its location information.[1] Each node stores the location information for its neighbours and if it receives a conflicting claim, revokes the offending node. A big problem of Node-to network Broadcasting is high energy consuming.

3.2.2. Random Multicast

In the Random Multicast (RM),[1] when a node broadcasts its location, each of its neighbours sends (with probability p) a digitally signed copy of the location claim to a set of randomly selected nodes. Assuming there is a replicated node, if every neighbor randomly selects O(_n) destinations, then exploiting the birthday paradox, there is a non negligible probability at least one node will receive a pair of non coherent location claims. The node that detects the existence of another node in two different locations within the same time-frame will be called witness. The RM protocol implies high communication costs: Each neighbor has to send O(_n)messages.

3.2.3. Line Selected Multicast

In the Line Selected Multicast (LSM) [1] protocol, uses the routing topology of the network to detect replication, each node which forwards claims also saves the claim. That is, the forwarding nodes are also witness nodes of a node which has the node ID in a claim. Therefore, LSM gives a higher detection rate than that of RM.

3.3.4. Resilient against node replication attack

Extremely efficient detection protocol XED is based on the strategy *remember and challenge* [3]. There is no need for a sensor node to know their location. When a sensor node *p*

meets another node it sends a random number to that node. Each and every node will have a number associated with the other node they met. When a node *p* meets the other node *q* it ask for the number. If the node does not replies or sending a wrong number indicates a replica.

When *p* and *q* are in the same range, they produce the random numbers $r_{p \rightarrow q}$ and $r_{q \rightarrow p}$ of *n* bits where they belongs to 0 to 2^{n-1} and then transfer them to each other. They have entries for ID of the node, the original random number and the received one. When *p* and *q* does not met before they exchange random numbers. If not *p* asks for the random number and *q* sends it. Now *p* checks for whether it is correct or not. Replica is detected if it is not matched with the stored one or *q* does not respond.

3.2.5. *Active detection of node replication attacks*

Each node is actively test if m other random nodes are replicated or not [14]. They are called as scrutinized nodes. In order to test whether the scrutinized node A is replicated or not. n nodes in the network are randomly chosen and asked to forward a request for a signed location claim to A. If two replicas exist , each will probably receive that and if both answers for that request, then two conflicting claims will be obtained. So that replica can be detected.

4. DETECTION OF REPLICA IN MOBILE SENSOR NETWORKS

4.1. Efficient and distributed detection of node replication attack

EDD [5] is based on the assumption that for the network with only benign nodes, the number of times n_1 that the node *n meets* a node *m* should be restricted to a threshold value within a given interval. When a network has replicas then this number will be greater than the threshold value. With this observation the replica node is detected.

Before deployment the length of the interval and the threshold values are selected. During the movement for each move of a sensor node the number of nodes encountered is compared with the predefined value. If it is greater, then a replica is detected.

A slight variation to EDD is SEDD. In EDD monitoring of all the nodes is done for calculating the number. But in SEDD, only a subset is monitored. That set is called as *monitor set.* When number of elements in the set is $, the number of nodes to be monitored by each node is randomly chosen $ unique nodes from *1 to n*, where *n* is the number of nodes in the network.

4.2. Fast detection of node replication attack using sequential analysis

It is based on the fact that a benign node should not move with speed

which is greater than the system configured speed V_s [6]. When a node's speed is greater than V_s then we can consider that at least 2 nodes are there with same identity.

For this Sequential Probability Ratio Test (SPRT) is performed. This can be considered as a one dimensional random walk with lower and upper bounds. The null and alternate hypotheses are defined. The random walk starts at a random point that lies in between the two bounds. When it reaches the bounds any one of the hypothesis is selected based on the bound crossed.

Each node sends the location and time information as a claim to its neighbours when it moves to a new location. Then they make decision on whether to forward the claim or not. The BS applies statistical analysis to check for the speed that is measured from the last claim and the current claim. If it is greater than the system configured speed, then replica is detected.

5. PERFORMANCE ANALYSIS OF THESE SCHEMES

Wireless sensor networks are designed with very limited battery power, limited memory size and limited processing capabilities. So any of protocol that is

used in WSN should have restrictions in the usage of battery power and memory. A variety of attacks can be mounted in the network using replica nodes [9][10][11]. The protocols suggested for these types of attacks should also have minimum energy and storage consumption.

In SET [1] the communication cost due to the transfer of messages is O(N). In this each root node is having N/T identifiers each of which is only 12bits. They occupy only limited space according to the number of trees formed.

For RED [2] only a very few nodes are storing more than 10 messages. So the storage overhead is reduced. In terms of energy consumption ,it needs only the witness nodes perform the verification of the signature, which in turn reduces the energy consumption. It also has better detection probability.

In XED [3] the location of each not is not recorded and only a constant communication cost O(1) is needed for exchanging the random numbers, since each node is capable of finding the replica per each move.

Real time detection of replica attack [4] is based on generation of finger print and forwarding. It has O(N) message transmission for fingerprint generation. If the total number of regular data messages is *Num* then the total messages message transmission cost is num.$\sqrt{N}$.

EDD [5] provides distributed detection, individual detection and network wide revocation avoidance. The additional communication overhead incurred by EDD and SEDD method is only *b* that is the beacon resulting in O(1). It has the reasonable storage overhead as O(N) and SEDD has improved storage overhead as O($) where $ is the number of elements in the monitor set.

In fast detection of replica using SPRT [6], communication overhead in the worst case will be O(N.$\sqrt{N}$).BS needs to store only one claim per node. So at most N claims are to be stored in the base station.

Centralized schemes

Protocols	Communication Cost
SET	O(N)
RED	O(√N)
Real Time detection	O(num_m. √N)
Random Key Predistribution	O(N)

num_m – Number of messages

N – Number of NodesDistributed Schemes

Protocols	Communication Cost
Node to Network Broadcast	O(N^2)
Random Multicast	O(N^2)
Line Selected Multicast	O(N√N)
XED	O(1)
Active Detection	O(√N)

6. RESULTS AND DISCUSSIONS

When we compare the communication cost due to message transfer, XED,EDD and SEDD are having the same cost in the order of O(1). Real time detection of replica node scheme is having the highest cost even considering minimum number of nodes related to others. RED and SET are having the average cost O(N) and O(N. √N). The SPRT scheme is having lowest cost even in the worst case O(√N).

7. CONCLUSION

Due to the unattended nature of sensor nodes, they are vulnerable to compromise. We cannot find out all the nodes which are the targets of the adversary. So a distributed detection is necessary to protect the network from the attacks from the compromised nodes. In this paper we analyze the most dangerous attack the node replication attack. From the analysis we come to the conclusion that performing sequential analysis for replica detection is the comparatively best one in mobile sensor networks. We are interested in changing the parameters for the hypothesis test and to compare the performance with the existing one.

We provide the overview of the existing schemes for the detection of replica nodes. Each one has its own advantages and limitations. Detection of replica nodes requires more research work. The detection method should consider the limited battery power and limited memory.

8. REFERENCES

[1] H. Choi, S. Zhu, and T.F. La Porta, "SET: Detecting Node Clones in Sensor Networks," Proc. Int'l Conf. Security and Privacy in Comm. Networks and the Workshops (SecureComm '07), pp. 341-350, 2007.

[2] M. Conti, R.D. Pietro, L.V. Mancini, and A. Mei, "A Randomized,Efficient, and Distributed Protocol for the Detection of Node Replication Attacks in Wireless Sensor Networks," Proc. ACM MobiHoc, pp. 80-89, Sept. 2007.

[3] K. Xing, F. Liu, X. Cheng, and D. Du. Real Time Detection of Clone Attack in Wireless Sensor Networks, In *IEEE ICDCS*, 2008.

[4] C.-M. Yu, C.-S. Lu, S.-Y. Kuo. Mobile Sensor Network Resilient Against Node Replication Attacks. In *IEEE SECON*, 2008. (poster)

[5] C.-M. Yu, C.-S. Lu, and S.-Y. Kuo, "Efficient and Distributed Detection of Node Replication Attacks in Mobile Sensor Networks," Proc. IEEE Vehicular Technology Conf. Fall (VTC Fall), Sept. 2009.

[6] J. Ho, M. Wright, and S.K. Das, "Fast Detection of Replica Node Attacks in Mobile Sensor Networks Using Sequential Analysis," Proc. IEEE INFOCOM, pp. 1773-1781, Apr. 2009.

[7] M. Bellare, R. Canetti, and H. Krawczyk. Keying Hash Functions for Message Authentication. Lecture Notes in` Computer Science Vol. 1109, 1996.

[8] W. Du, J. Deng, S. Han, and P. Varshney. A pairwise key predistribution scheme for wireless sensor networks. In Proceedings of ACM Conference on Computer and Communications Security, 2003.

[9] B. Parno, A. Perrig, and V.D. Gligor, "Distributed Detection of Node Replication Attacks in Sensor Networks," Proc. IEEE Symp.Security and Privacy, pp. 49-63, May 2005.

[10] A. Becher, Z. Benenson, and M. Dornseif. Tampering with Motes: Real-World Physical Attacks on Wireless Sensor Networks. International`Conference on Security in Pervasive Computing, 2006.

[11] S. Capkun and J.-P. Hubaux. Secure Positioning of wireless devices with application to sensor networks. IEEE Infocom, 2005.

[12] S. Zhu, S. Setia, and S. Jajodia. LEAP: Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks. ACM conference on Computer and communications security, 2003.

[13] Richard Brooks, P.Y.Govindaraju, Mathew Pretti, N.Vijaykrishnan, M.T.Condemir, "On the detection of clones in sensor using random key predistribution", IEEE transactions on system, man, and cybernetics-part c applications and reviews, vol 37, no. 6, November 2007

[14] C.A.Melchor et.al "Active detection of node replication attacks", IJCSNS International Journal of Computer Science and Network Security, vol 9,no.2, February 2009.