A Survey on Detection of Clones in Wireless Sensor Networks

Akhila Daniel
Computer Science and Engineering department
SCT College of Engineering, Trivandrum, Kerala
Affiliated to Kerala University

Preeja. V
Computer Science and Engineering department
SCT College of Engineering, Trivandrum, Kerala
Affiliated to Kerala University

ABSTRACT
Now a days wireless sensor networks (WSN) have wide applications in variety of fields such as military purposes, environmental monitoring, and gathering sensing information in inhospitable locations. But with the increase in use of wireless sensor network the risk of data leakage also increased. The adversary can launch different types of attacks to the network which may subvert the overall functioning of the network. Sometimes the attacker captures some nodes from the network and their credentials will be collected. Some clones of the captured nodes will be created with same credentials and they will be deployed to the network. These nodes act as eavesdroppers of the network data and they will exploit all the privileges of the original captured nodes. So these clones should be detected before they can do much harm to the network. Several algorithms are developed for this purpose.

Here a study is carried out on the various novel techniques introduced to detect replicas in wireless sensor networks and their efficiency and performances are analyzed.

General Terms
Wireless Sensor Networks (WSN), Clones, Node replication, Replicas

1. INTRODUCTION
Wireless sensor network is one of the emerging fields of 21st century with wide applications. Wireless sensor network can monitor a particular area from a remote location with high accuracy. The network is composed of a number of small inexpensive devices called sensor nodes which has the capabilities of sensing, computing and wireless communication. These sensor nodes are scattered in the area where we want to monitor. Hence the position of the sensor nodes cannot be predetermined. Various algorithms provide self organizing capacity to sensor nodes and all the sensor nodes work in co-ordination to measure a physical environment correctly

Challenges faced by WSN:

1. ENERGY UTLIZATION OF SENSOR NODES: Each sensor node has a particular finite amount of energy which will be provided by some batteries inside the sensor node. There is no external energy providing source. So the power should not be wasted in any manner. The communication and processing should be minimized for optimal energy utilization.
2. DEPLOYMENT AND SELF ORGANIZATION: Usually the sensor network will be deployed in inhospitable locations. So the deployment cannot be done manually (throwing the nodes from some vehicles like aeroplane). Hence proper placement of each node is not practical. Therefore the algorithms and protocols must provide the self organization of nodes for its proper functioning.
3. ROBUSTNESS: The sensor nodes should be adaptable to the various environmental changes. They should also reconfigure in case of any cases.
4. PRIVACY AND SECURITY: The data transmitted through the wireless sensor networks may be extremely sensitive in case of some military applications. So effective mechanisms should be adapted to prevent data leakage.

2. SECURITY OF WIRELESS SENSOR NETWORKS
With the increasing use of wireless sensor networks for data collection the risks of data transmission also increased. In fact, wireless networks are usually more vulnerable to various security threats as the unguided transmission medium is more susceptible to security attacks than those of the guided transmission medium [10]. Several attacks in WSN are summarized in the Table 1.

3. NODE REPLICATION ATTACK-A THREAT TO SENSOR NETWORK SECURITY
Node replication attack is a severe attack in wireless sensor networks. In this attack the attacker will capture some nodes, make copies of them and deploy those nodes in to the network shown in Figure 1. Since the replicated nodes use the identities of the captures nodes they cannot be detected easily. They will also misuse all the privileges of the legitimate nodes. In this way an attacker can leak the sensitive data flowing through the network and can also insert false data in to the network.

Table 1: Attacks in WSN

Attack	Description
Wormhole attack	Attacker uses tunneling mechanism to establish path between himself and two nodes by confusing the routing protocol
Black hole attack	Attacker places himself in the network with high capacity resources so that all data passes through the attacker.
"Hello flood" attack	Specific type of Denial of Service attack in which the attacker broadcasts hello packets in the network. The receiving node assumes that this is the nearest node and send packets to that node and thereby creates congestion in the network
Denial of service	Attacker sends extra unnecessary packets in the nodes and thereby exhausting the resources of the sensor nodes so that a legitimate user will not get adequate service from the network.
Attack of information in transit	Information in transit may be altered, spoofed or replayed again by the attacker.
Sybil attack or node replication attack	Attacker gains identity of a sensor node and behaves like that node and misuse the privileges of the replicated node

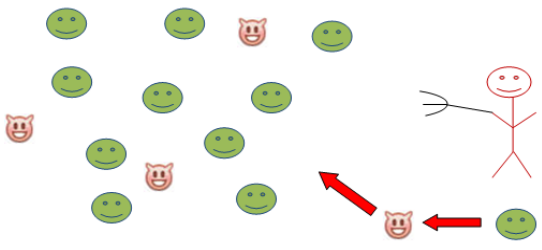

Figure 1: Attacker injecting clones to the network

4. COUNTERMEASURES TO DETECT REPLICATED NODES

The replicated nodes should be detected as early as possible before the attacker can subvert the functioning of the network. Also the detection of nodes should not cause much overhead to the network. Some sensor networks are stationary and the nodes may be fixed at a location (static WSN) but in some other wireless sensor networks the nodes may be continuously moving from one location to another (mobile WSN). Separate algorithms are there to detect clones in stationary and mobile sensor networks. The techniques to detect node replication attacks can be broadly classified in to two types: Centralized and Distributed. In centralized approach there will be central entity called a base station. It is the task of the base station to detect replicated nodes. But in distributed approach this tedious task of the base station is distributed among several nodes and each node will check for the existence of replicas. Several algorithms are developed to detect clones in wireless sensor networks. Each has its own pros and cons. Table 2 shows the some of the algorithms to detect replicated nodes.

Table 2: Various algorithms to detect clones

Sl.No	Algorithm	Type of WSN	Type of approach
1	Random key pre-distribution technique	Static	Centralized
2	Sequential analysis method	Mobile	Centralized
3	A memory efficient technique	Static	Distributed
4	Random walk based approach	Static	Distributed
5	Localized multicast approach	Static	Distributed
6	Randomized Efficient and Distributed approach	Static	Distributed
7	A Range based detection method	Mobile	Distributed
8	Trust based detection method	Static	Distributed
9	Using Localized Extremely Efficient Distributed(XED) algorithms	Mobile	Distributed

4.1 Random Key Pre-distribution Technique

This is one of the early replica detection techniques which rely completely on the base station [1]. The main data structure this algorithm maintains is a key pool. Before the sensor nodes are deployed a key pool of P cryptographic keys are generated. Soon after the key generation each of the nodes in the network randomly selects a subset of k keys from the key pool. This subset of k keys is called the key ring and they cannot be replaced. These keys act as the authentication tokens. There is also a base station and a system wide symmetric key is used for communication with the base station. After the deployment of the sensor nodes the next step is to establish links between every possible node. It is done through two phases, key discovery phase and path key establishment phase. In key discovery phase every two pairs of the node tries to establish a link between them if they found a common key in their set of k keys. Further communication between those two nodes will be using that shared key. But there may be some nodes they are within the communication range but they may not have any keys in common. The will try to establish a direct link by exchanging a key. We can determine the presence of clones in the network when the key usage distribution is skewed. Since the clones also uses the same key that the normal nodes are using. So the keys are present on a greater number of nodes than normal. The base station also maintains a copy of the key pool. The nodes communicate with the base station using the shared key and the base station thus collects the key usage statistics form the network. Analyzing the usage statistics the base station determines the cloned keys. For collecting the key usage statistics, counting bloom filters are used. There exists a gossip protocol also for broadcasting purpose. On detecting the cloned node all connections are terminated. The main problem with this technique is its single point of failure and communication overhead.

4.2 Sequential Analysis

This Sequential method is based on sequential probability ratio test. This scheme is based on the assumption that an uncompromised mobile node should move at speeds in excess of the system configured maximum speed [2]. There will be a system configured maximum speed. If the speed of a node is identified to be greater than the maximum speed that node will be considered as a clone. So the main challenge here is to configure the upper and lower speed limits so that we can choose the right hypothesis for fast and accurate detection.

Here the system is designed in such a way that every mobile sensor nodes communication is bidirectional. Then Sequential Probability Ratio Test is applied to the moving sensor nodes. Mobility provides us with a clue that can help to resolve the mobile replica detection problem [2]. A random walk is performed through the sensor nodes to detect the clones. Before the random walk starts a null and alternate hypothesis are defined in such a way that the null one is associated with the lower limit and alternate one associated with upper limit [2]. A random walk starts from a point between two limits and moves towards the upper or lower limits.

This protocol has mainly two steps, (i) claim generation and forwarding (ii) detection and revocation. In claim generation and forwarding phase each node starts its random walk. When it reaches a new location it identifies the location as well as its neighbors. All neighbor nodes ask for a location claim and the node sends the location claim if the request doesn't timeout. The neighbors will then authenticate the claim. If any of the neighbor nodes fail to authenticate the node, it will be removed from the neighbors list of that particular node otherwise the neighbor will forward the claim to the base station. The role of base station comes to play in the next phase, detection and revocation phase. In this phase the base station authenticates the claim by extracting the time and location information. It then calculates the speed from this information and compares with the previous location claim of the same node. If any mismatch occurs in the speed the node with that identity is revoked. The main problem with this approach is that several messages are to be sent which may cause great communication overhead.

4.3 A Memory Efficient Technique

As the name indicates the main objective of this Memory Efficient Approach[3]is to minimize the memory utilization for detecting clones. Two new techniques called cell forwarding and cross forwarding are used here to improve detection probability and to reduce memory and energy consumption. It uses the technique of propagation of location claim as in random key predistribution to locate the clones. The main defects with that system are single point of failure because of the existence of a base station and large communication overhead. But the base station has no role here. Four replica detection protocols are proposed in this work [3]. The bloom filters mentioned in the random key predistribution technique is also used in first protocol but the difference is that the location claim information is encoded in the bloom filter. In the second approach with the intention of solving cross over problem a cell forwarding mechanism is introduced which improves detection probability and reduces memory overhead. The crowded centre problem is resolved using the cross forwarding technique in the third protocol. Then cell forwarding and cross forwarding techniques are integrated in the fourth approach to achieve best performance.

4.4 A Random walk Based Approach

The Random Walk Based Approach rectified the disadvantages of existing methods in which all of them needs a central control and cannot defend against smart attacks. Smart attack means the adversary finds out the witness nodes that will detect the replicas and only compromise those witness nodes to avoid detection [4]. The main features of this approach are its fully distributed and nondeterministic nature. Here all nodes can be equally witness of a node and is also resistant to smart attack. It naturally distributes the responsibility of witness node selection to every passed node of random walks, and then adversaries cannot easily find out the critical witness nodes [4]. Two protocols are suggested one is Random walk based detection and Table assisted random walk which detects the replicas with the assistance of a table. In these protocols soon after a node broadcasts it location claim each of its neighbors forwards the claim to some randomly selected nodes. After that a random walk is initiated by this randomly selected node and each passed nodes are selected as witness nodes and they will store the claim. The replicated nodes are revoked if different location claims are received for a same node ID.

4.5 Localized Multicast Approach

Localized Multicast Approach is also a distributed replica detection method for detecting clones based on the location claim. This approach [5] is also based on the witness finding strategy in which the witness nodes are randomly selected from the nodes which are located in a geographically limited region (cell). First the node ID is mapped to one or more cells and using randomization witness nodes are selected. Witness

node performs the usual location verification procedure to detect replicas.

4.6 Randomized Efficient and Distributed Approach

The Randomized Efficient and Distributed protocol [6] is applied to the nodes which are stationary and it works in two steps. Like some of the previous methods it also relies on a trusted entity which broadcast a random value or seed to all the nodes which constitute the first step. For ensuring more security we can use some election mechanism to choose the trusted entity. Also some mechanisms are adopted to make sure that nodes do not lie about their physical location. In the second step each node will locally broadcast its claim (ID and location). The neighbors again send the claim to some pseudo randomly selected network locations. The claim is not sent to some selected node Id's because there is a chance that some nodes may not be present in the network and those claims will be lost. The witness nodes of a particular node are selected using the ID of a node, the current random value and randomly selected locations. As usual the witness nodes will verify each location claim and checks with the previous location claim of the same node or store the claim if it the one with a new node ID.

4.7 Range Based Detection Method

The Range Based Detection method [7] is based on the idea of ranging between nodes. It uses system synchronization time and precise node location information as key factors for the detection of clones in mobile sensor networks. The fundamental idea is the unique identification property. That is there will not be two nodes in the network with same identity. Each node categorizes its neighbors to close neighbors and far neighbors depending on the distance to them. Each node maintains the details about their neighbors in a neighbor information table which will be periodically broadcasted through the network. Unique Identification property will be applied to expel the clones. This method defines three criterions: Local Unique ID Criterion (LUIC), Neighbor Unique ID Criterion (NUIC) and Global Unique ID Criterion (GUIC)

4.8 Trust Based Detection Method

The Trust Based detection method [8] is focused on a trust based witness finding strategy with a trust factor. Here each node overhear its neighboring nodes sensing data and compare it with its own result to measure their trustworthiness. Each node also maintains a neighbor behavior table with the details of neighbor's behavior with a consistency count, inconsistency count, sensing success and sensing failure. In this protocol there is a major role for the centralized base station. A random seed will be generated and broadcasted by the base station. Each node will forward a signed location claim to some witness nodes. The witness nodes are selected by a node from its neighbors by comparing their trust with the threshold. If any witness node receives different location claims for a particular node that node will be considered as a replica. This method also has great communication overhead.

4.9 Localized Extremely Efficient Detection(XED) Algorithms

This Localized algorithms[6] proposes an Extremely Efficient Detection approach which resists to node replication attacks in a localized fashion. The main idea of the algorithm is the exchange of some code between two nodes. There will be a random number shared between two nodes. Every time they meet each node authenticate the other using that random number. The random number is generated using some cryptographic hash functions. In this method there is a greater memory requirement because each node has to store the authentication key of the nodes they meet. Each node also stores a set containing black listed nodes for which the authentication step fails and the probability of them being a clone is high. Hence a node refuses to communicate with the black listed nodes.

5. PERFORMANCE EVALUATION

Efficiency of all these algorithms are evaluated based on the communication and storage overhead. Table 3 shows the communication and storage overhead of the above algorithms.

Table 3: Performance evaluation

ALGORITHM	COMMUNICATION OVERHEAD	STORAGE OVERHEAD
Random Key Predistribution	$O(nlogn)$	$O(n)$
Sequential analysis	$O(n\sqrt{n})$	O(n)
Amemory efficient technique	$O(n^2)$	O(k)
Random Walk	$O(\sqrt{n}logn)$	O(1)
Localized multicast	$O(d.p_f.\sqrt{n}) + O(s)$	$O(s.p_s)$
Randomized Efficient and Distributed	$O(r.\sqrt{n})$	$O(w.\sqrt{n})$
A Range based detection method	$O(n\sqrt{n})$	$O(\sqrt{n})$
Trust based detection method	$\frac{SS_i - SF_i}{SS_i + SF_i}$	–
Using Localized XED algorithm	$O(1)$	$O(n)$

where,

n	Number of sensor nodes
k	Number of random line segments in the network
d	Number of neighbors of a node
p_f	Probability that a node forwards a claim
s	Number of sensors in network
p_s	Probability that a node stores the location claim
w	Percentage of witness nodes
r	Number of witness nodes
SS_i	Sensing success count of node i
SF_i	Sensing failure count of node i

6. CONCLUSION

A study on different algorithms for replica detection in wireless sensor network has been done and their performances are evaluated. The factors which determines the efficiency of an algorithm is communication, computation and storage overhead. Detection accuracy is also a main factor. Distributed algorithms are always efficient than the centralized approach because the problem of single point of failure can be eliminated. The detection of clones in mobile sensor network is more tedious than the detection in static sensor networks because we cannot use the location claim there. The main objective of every algorithm is to detect the clones with minimum communication and computation overhead with better detection accuracy.

7. REFERENCES

[1] R. Brooks, P. Y. Govindaraju, M. Pirretti, N. Vijaykrishnan, and M. T. Kandemir, 2007, "On the detection of clones in sensor networks using random key predistribution"., IEEE Trans. Syst., Man, Cybern. C, Applicat. Rev., vol. 37, no. 6, pp. 1246–125.

[2] J. Ho,M.Wright, and S. K. Das. 2009, "Fast detection of replica node attacks in mobile sensor networks using sequential analysis,". in Proc. IEEE Int. Conf. Computer Communications (INFOCOM), Brazil,, pp. 1773–1781.

[3] M. Zhang, V. Khanapure, S. Chen, and X. Xiao. 2009 "Memory efficient protocols for detecting node replication attacks in wireless sensor networks," in Proc. IEEE Int. Conf. Network Protocols (ICNP), Princeton, NJ, USA, pp. 284–293.

[4] Y. Zeng, J. Cao, S. Zhang, S. Guo, and L. Xie. 2010 "Random-walk based approach to detect clone attacks in wireless sensor networks",IEEE J. Sel. Areas Commun., vol. 28, no. 5, pp. 677–691,

[5] B. Zhu, S. Setia, S. Jajodia, S. Roy, and L. Wang. 2010, "Localized multicast: Efficient and distributed replica detection in large-scale sensor networks",IEEE Trans. Mobile Comput., vol. 9, no. 7, pp. 913–926.

[6] M. Conti, R. D. Pietro, L. V. Mancini, and A. Mei. 2011 "Distributed detection of clone attacks in wireless sensor networks", IEEE Trans. Depend. Secure Comput., vol. 8, no. 5, pp. 685–698.

[7] Huang Jian1, Xiong Yan1+, Li Ming-xi1, Miao Fu you2. 2012 "A Range-based Detection Method of Replication Attacks in Wireless Sensor Networks", International Conference on Information and Computer Networks (ICICN).

[8] Manjula and C. Chellappan. 2012",Trust based node replication attack detection protocol for wireless sensor networks"(2012), Journal of Computer Science , 8 (11), 1880-1888, ISSN 1549-3636

[9] Chia-Mu Yu, Yao-Tung Tsou, Chun-Shien Lu, Member, IEEE, and Sy-Yen Kuo. Fellow,IEEE. 2013"Localized Algorithms for Detection of Node Replication Attacks in Mobile Sensor Networks", IEEE transactions on information forensics and security.

[10] Sakib Khan Pathan, Hyung-Woo Lee and Choong Seon Hong," security in wireless sensor networks: issues and challenges".