Use of Sequential Hypothesis Testing for Detection of Mobile Replica Nodes in Wireless Sensor Network

Sandip D. Girase
Department of Computer Engineering,
SSBT's College of
Engineering & Technology,
North Maharashtra University,
Jalgaon, Maharashtra, India

Ashish T. Bhole
Department of Computer Engineering,
SSBT's College of
Engineering & Technology,
North Maharashtra University,
Jalgaon, Maharashtra, India

ABSTRACT
Network protection has grown to be a demanding area, previously tackled only by well qualified and familiar experts. Although more pupils become wired, an increasing number of pupils need to understand basis of security in the network world. The replica node attacks are hazardous as they allow attackers to leverage the compromise of a few nodes and exert control over much of the network. Earlier works on replica node recognition relied on set sensor locations and hence do not effort in mobile sensor network. The proposed method uses sequential probability ratio test for detection of mobile replica node. It has provided exclusive identity to the sensor nodes so that an adversary can not disturb the network. The proposed sequential hypothesis testing results in a better detection of mobile replica nodes within wireless sensor networks. The replica node attacks are hazardous as they allow the attacker to influence the compromise of a few nodes to make use of power over a lot of the network. A number of detection schemes have been proposed for static sensor networks, fixed sensor locations and it does not effort in mobile sensor networks. An effective mobile replica node detection scheme is proposed with Sequential Probability Ratio Test (SPRT).

Keywords
Wireless sensor network (WSN), network security, attack detection, node replication, tamper resistant hardware, sequential probability ratio test (SPRT)

1. INTRODUCTION
Now a day's robotics has advances which are responsible for raising several architectures for autonomous wireless sensor networks. The task such as static sensor network, adaptive sensor network and adaptive sampling is done by using mobile nodes with sensing and movement capabilities [1]. The advances are also useful for the applications such as military patrols, border monitoring and intruder detection. Wireless sensor network is also useful for military application and security monitoring. The wireless sensor network usually needs to be controlled remotely by the network operator they are often deployed in an unattended manner. The unattended nature of wireless sensor networks is exploited by adversaries. The adversaries takes the secret keying materials from a compromised node, generates a large number of attacker controlled replicas that shares the node's keying materials and ID and spreads these replicas throughout the network. With a single captured node, the adversary creates as many replica nodes as one has the hardware to generate.

One of the solutions for it is that the use of tamper resistance hardware to prevent adversary from extracting the keying material.

1.1 Tamper Resistance Hardware
For designing a secure computer system is ensuring that various cryptographic keys can be accessed only by their intended user(s) and only for their intended purposes. Keys stored inside a computer can be vulnerable to use, abuse, and or modification by an unauthorized attacker .For protecting the keys the appropriate way is to store them in a tamper-resistant hardware device. These devices can be used for applications ranging from secure e-mail to electronic cash and credit cards [2].

1.2 Sequential Hypothesis Test
The sequential hypothesis testing is statistical analysis where the sample size is not fixed in advance. Instead data are evaluated as they are collected, and further sampling is stopped in accordance with a pre-defined stopping rule as soon as significant results are observed. Thus a conclusion may sometimes be reached at a much earlier stage than would be possible with more classical hypothesis testing or estimation, at consequently lower financial and/or human cost.

1.3 Sequential Probability Ratio Test
The sequential probability ratio test (SPRT) is a specific sequential hypothesis test which is developed for use in quality control studies in the realm of manufacturing; SPRT has been formulated for use in the computerized testing of human examinees as a termination criterion [3]. It is nothing but a statistical hypothesis test which is a method of making decisions using data, whether from a controlled experiment or an observational study. In statistics, a result is called statistically significant if it is unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significance level. These tests are used in determining what outcomes of an experiment would lead to a rejection of the null hypothesis for a pre-specified level of significance; helping to decide whether experimental results contain enough information to cast doubt on conventional wisdom [4]. It is sometimes called confirmatory data analysis, in contrast to exploratory data analysis. Statistical hypothesis testing is a key technique of frequents statistical inference. The critical region of a hypothesis test is the set of all outcomes which cause the null hypothesis to be rejected in favor of the alternative hypothesis.

1.3.1 Null Hypothesis
The null hypothesis is nothing but a typically corresponds to a general or default position. Null hypothesis is typically paired with a second hypothesis, the alternative hypothesis, which

asserts a particular relationship between the phenomena [5]. The alternative need not be the logical negation of the null hypothesis it predicts the results from the experiment if the alternative hypothesis is true. The use of alternative hypotheses was not part of Fisher's formulation, but became standard. It is important to understand that the null hypothesis can never be expanded beyond the doubt. A set of data can only reject a null hypothesis or fail to reject it [6].

1.3.2 Alternative Hypothesis

The alternative hypothesis (or maintained hypothesis or research hypothesis) and the null hypothesis are the two rival hypotheses which are compared by a statistical hypothesis test. Sequential probability is a statistical decision process [7]. It consists of one dimensional random walk with lower and upper limit. A random walk is a mathematical formalization of a path that consists of a succession of random steps.

2. LITERATURE SURVEY

A literature survey is an evaluative report of studies found in the literature related to selected area. The survey should describe, summarize, evaluate and clarify the literature. It should give a theoretical basis for the research and help to determine the nature of research. It demonstrate a strong knowledge of the current state of research in the field then it show what issues are being discussed or debated and where research is headed and provide excellent background information for placing a program, initiative or grant proposal in context.

2.1 Background

A wireless sensor network consists of hundreds or even thousands of tiny nodes which are circulated over the network. These nodes sense the sensitive data from the locality and send the sensitive message to the base station the base station will authenticate the data and ID which is send by the sensor nodes [6]. These sensor nodes are deployed in unfriendly atmosphere and the nodes are unattended which makes an adversary to negotiation the sensor nodes and make many replicas of them. These replica nodes are hazardous to the network communication. Advances in robotics enlarge a variety of new architectures for self-governing wireless sensor networks. Mobile nodes in network communication are useful for network repair and event detection. These advanced Sensor network architecture could be used in variety of application including intruder detection, border monitoring, and military patrols. The compromised mobile nodes inject the fake data and disrupt network Operations and eavesdrop on network communications.

2.2 Related Work

Various replica node detection schemes have been proposed for the static sensor networks. Basic method used by these schemes is to have nodes report location claims which identify their positions and for other nodes to attempt to detect conflicting reports that single one node in multiple locations. But this approach requires the fixed node locations. It detects when nodes are expected to move.

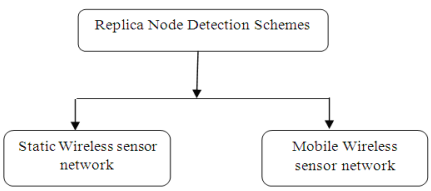

Figure 2.1 Replica Node Detection Schemes

Replica detection schemes are based on two ways that contains static wireless sensor network and mobile sensor network as shown in Figure 2.1. For static wireless sensor networks the techniques developed for node replication do not work when the nodes are likely to move as in mobile wireless sensor network, and thus they have turned out to be useless for mobile WSNs. As a result some techniques have also been developed for mobile WSNs to detect the replica or clone nodes. Replica node detection schemes are classified into two main classes as centralized and distributed.

2.2.1 Static Wireless Sensor Networks

Xing et al., in [7], proposed a social fingerprint which is computed for each sensor by using the neighborhood characteristics, and checks the legitimacy of the originator for each message by checking the enclosed fingerprint. Generation of fingerprint is depends on the superimposed s-disjunct code, which incurs a very light communication and computation overhead. The checking of fingerprint is conducted at both the base station as well as the neighboring sensors, which ensures a high detection probability. It also provides the real time clone detection in efficient as well as effective way. Unlimited clones were deploying by capturing and compromise nodes. These nodes can be involve such as that of legitimate node and have access the legitimate IDs and keys .If these remains inside the network and left undetected then the network get unshielded to attackers and clone attackers are spread over the entire network. Smart clone may try to hide from being detected by all means. So far they may collude to cheat the network administrator into believing that they are legitimate. Clone node may be serving by adversary in the network at anywhere. In this scheme main focus is on preventing technology rather than detecting technology. The scheme explores the superimposed s-disjunct code for timely clone attack detection. At very short bit stream, fingerprint can be encoded then which results in small message overhead. It also identify clone with high accuracy.

Parno et al., in [8], propose two new algorithms which are based on emergent properties that is properties that arises only during the collective action of multiple nodes. Randomized Multicast distributes node location information to randomly-selected witnesses, exploiting the birthday paradox to detect replicated nodes, while Line Selected Multicast uses the topology of the network to detect replication [14]. Both algorithms provide globally aware, distributed node replica detection, and Line Selected Multicast displays mainly tough presentation characteristics. Also shows that emergent algorithms represent a promising new approach to sensor network security.

2.2.2 Mobile Wireless Sensor Networks

Deng and Xiong, in [9], describe the new protocol for the detection of node replication attacks in the mobile wireless sensor networks. The number of pair-wise keys established by each node was collect by Counting Bloom Filters. This data is used by server to create a histogram. Also derive expression for the expected number of pair wise keys conventional by each node, and give the entrance for detection. Nodes whose number of pair wise keys exceeds the threshold value are considered to be replicas. Next the system can recover from the node replication attack by terminating connections established by the replicas. In the scheme each detection probability is determined by the total number of replicas that have been detected and the total number of nodes that have compromised by that time point. And each detection error rate

is determined by the total number of non-compromised nodes that have been incorrectly detected by that time point and the total number of non-compromised nodes in the network.

Balaji and Anihta, in [10], proposed the scheme XED and EDD. The proposed techniques developed solutions for a replica attack, challenge and response and encounter number, are basically dissimilar from the others. The proposed algorithm can resist node replication attacks in a localized fashion. Compared to the distributed algorithm, nodes perform the task without the intervention of the base station. The localized algorithm is a particular type of distributed algorithm. Each node in the localized algorithm can communicate with only its one hop neighbors. This attribute is useful in sinking the communication overhead significantly and enhancing the resilience against node compromise. The algorithm can identify replicas with high detection accuracy. The revocation of the replicas can be performed by each node without flooding the entire network with the revocation messages. The time of nodes in the network does not need to be synchronized.

3. PROPOSED WORK

Proposed work is focused on the use of sequential probability ratio test for detection of mobile replica nodes in efficient way.

3.1 Proposed Approach

Possibility made by robotics is to development of variety of new architecture for autonomous wireless network of sensors. A novel mobile replica detection scheme based on the sequential hypothesis test is proposed. Sensor network architectures are used for variety of applications such as military patrols, intruder detection and border monitoring.

3.1.1 The Sequential Probability Ratio Test

The sequential probability ratio test (SPRT) is a specific sequential hypothesis test which is developed for use in quality control studies in the realm of manufacturing, SPRT has been formulated for use in the computerized testing of human examinees as a termination criterion .It is nothing but a statistical hypothesis test which is a method of making decisions using data, whether from a controlled experiment or an observational study. In statistics, a result is called statistically significant if it is unlikely to have occurred by chance alone, according to a pre-determined threshold probability, the significance level. These tests are used in determining what outcomes of an experiment would lead to a rejection of the null hypothesis for a pre-specified level of significance; helping to decide whether experimental results contain enough information to cast doubt on conventional wisdom. It is sometimes called confirmatory data analysis, in contrast to exploratory data analysis. Statistical hypothesis testing is a key technique of frequents statistical inference. The critical region of a hypothesis test is the set of all outcomes which cause the null hypothesis to be rejected in favor of the alternative hypothesis. SPRT has been proven to be the best method in terms of the average number of observations that are required to reach a decision among all sequential and non-sequential test processes

3.1.2 Assumptions

The communication of nodes in mobile sensor network is with base station. Also nodes have way to correspond consistently to the base station on commonly basis. The assumptions in proposed work are as follows:

1. The Mobile Sensor Network is two dimensional.
2. The direct communication link between sensors nodesare bidirectional.
3. For each communication process, both source anddestination nodes are not malicious.
4. Nodes have fixed topology network.

3.2 Proposed System Architecture

Every time a mobile sensor node moves to a novel location, each of its neighbors asks for a signed claim containing its location and time information and decides probabilistically whether to forward the received claim to the base station or not as shown in Figure 3.1.

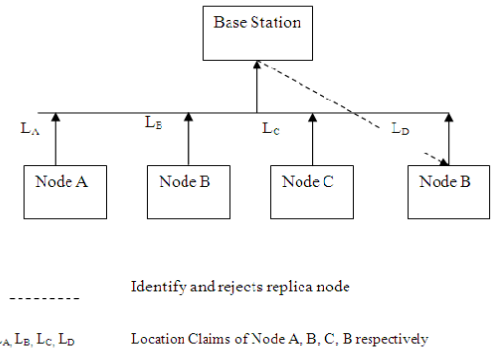

Figure 3.1 System architecture

The Figure 3.1 shows the architecture of the system. What should be its input material, how does it process and what is its desired output. Location ID is provided to each node. The base station computes the rate from every two consecutive claims of a mobile node and performs the sequential probability ratio test (SPRT) by considering speed as an observed sample. A little benefit is to the attacker of having a replica node in the same area as another compromised node. The compromised node can straightly report fake data, participate in local control protocol. The base station computes the rate from every two consecutive claims of a mobile node and performs the SPRT by considering speed as an observed sample. The compromised node can straightly report fake data, participate in local control protocol. Algorithm used for the proposed scheme is as shown in Figure 3.2.

Step 1: Let Number of Node n, Current Location L, Current Time T

Step 2: If Node n>0, compute speed for current_location L1, current_time T1 (n) and previous_location L0 and previous time T0 (n)

Step 3: If speed >Vmax, then replica detected

Step 4: Else accept test and terminate

Step 5: Prev_loc=cur_loc

Prev_time = cur_time.

Step6: Else go to step2

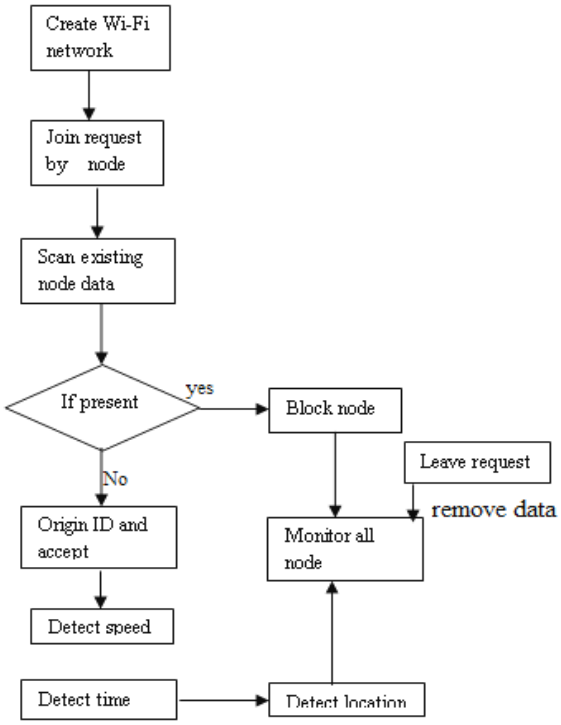

Figure 3.2 Flow of data

Each time a mobile sensor node u move each time to a new location, Lu represent the location of node u and discovers a set of near nodes N (u) as shown in Figure 3.2. Every neighboring node v ∈ N (u) asks for an true location claim from node u by sending its current time T to node u. after receiving T, node u checks whether T is valid or not. If yes then it proceed towards the location detection and finally it monitor the nodes.

4. RESULTS AND DISCUSSION

The result and discussion chapter describe detection of mobile replica nodes using sequential probability ratio test. Includes experimental setup, experimental results for existing detection schemes based on the performance with the proposed detection techniques based on reviews and discussion. The metrics used for the proposed work are detection rate, mobility rate and overhead.

4.1 Implementation Details

During the implementation, location ID is provided to each mobile sensor node and every mobile sensor node u generates location claim Cu = {u||Lu||T||Sigu} and sends it to a neighboring node v, where u, is the node identity, Lu is the Location, T is the Time and Sigu is the signature generated by node u's private key. Each time a mobile sensor node u moves to a new location, it first discovers its location Lu. Base station receive location claim from the mobile sensor nodes. Upon receiving a location claim, the base station verifies the authenticity of the claim with the public key of node u and discards the claim if it is not authentic. Threshold value for the maximum velocity of the mobile sensor node is given in base station. When a mobile sensor node moves from one location L1 to another location L2, the Euclidean distance is calculated between L1 and L2 (L2-L1). Similarly the time for the above location movement is measured using (T2-T1). Speed for a mobile sensor node is calculated using Speed S= (L2-L1) / (T2-T1). When the calculated speed S is less than the threshold, it is considered to the normal node, else it is considered to the replica node.

4.1.1 Data Structure

The data structure is a collection of data items stored in memory. Beside a number of operations are provided by the software to use that data structure. A data structure is some sort of relationship between the data items. Exactly determine what the relationships are and what type of data structure is being used. In computer programming, a data structure may be selected or designed to store data for the purpose of working on it. A data structure is a specialized format for organizing and storing data. Any data structure is designed to organize data to suit a specific purpose so that it can be accessed and worked with in appropriate ways.

Table 4.1 Data Structure of Packet Generated

Packet		**Packet Format**				
Data packet	Source ID	Destination ID	Packet type	Packet ID	Hop count	Trust value

The Table 4.1 shows the data structure of generated data packet by a sender. Source ID and destination ID is required for sending data packet from particular source to particular destination. Packet ID must be unique for each node's generated packet. Hop count is must to determine to shortest path to know the intermediate nodes. Trust value gets increment when a node forwards data packet to the next neighboring node. The Table 4.2 shows data structure of base station.

Table 4.2 Data structure of Base Station

Service	**Category**	**Node ID**	**Node name**	**Region**	**Speed**	**Key**	**Status**

It requires the field for creating the new client in the network. Node Id and node name is assign to the new client. Region is provided as location of new client.

4.2 Simulation Environment and Parameters

The replica detection scheme is simulated using Java swing and Net beans 7.1 environment. The system is compatible with Win7.The system is runs on a laptop with Intel(R) Core (TM) i3 CPU and 4-GB RAM. Simulation model having scenario of n (user defined) mobile sensor nodes and used to study the replica detection scheme and their performance. The parameters used for proposed work are as in Table 4.3.

Table 4.3 Simulation Parameters

Sr. No.	**Parameters**	**Types/Values**
1	Channel	Wireless
2	Routing Protocol	TCP/IP, UDP
3	Network Interface Type	Wireless Physical

4	MAC Type	Mac/IEEE802.11
5	Number of Reference Nodes	5 to 15
6	Number of Malicious Mode	2
7	Delay(milliseconds)	1.46

4.3 Performance Metrics

The simulation results are observed with respect to performance evaluation metrics. The performance evaluation metrics also contribute to study and analyze the sequential probability ratio test for detection of mobile replica nodes in wireless sensor network. The metrics that are considered for performance evaluation are as follows:

Detection of replica nodes is defined as the recognition of the replica nodes as shown in Equation 4.1.

$$r = 1 - \frac{\min(Pr, Pe)}{\max(Pr, Pe)} \qquad 4:1)$$

Where,

r = ratio difference

Pr = received signal strength

Pe =expected signal strength

Overhead is the average number of claims that are send or forwarded by nodes in network. Overhead is also the sum of number of malicious nodes and number of reference nodes, divided by delay as shown in Equation 4.2.

$$Overhead = \frac{number\ of\ malicious\ nodes + number\ of\ referance\ nodes}{delay} \qquad (4.2)$$

4.4 Experimental Results

The performance of sequential probability ratio test is evaluated in Java Swing. Here two parameters are considered for detection of mobile replica nodes in wireless sensor networks such as detection rate (replica nodes) and overhead.

The Table 4.5 shows the result of existing system and proposed sequential probability ratio test which shows the minimum delay for proposed work.

Table 4.5 Comparing Existing Attack Resilient System with Proposed Sequential Probability Ratio Test

Parameters	Proposed Sequential Probability Ratio Test (SPRT)	Existing Attack Resilient system
Detection Rate (Replica Nodes)	100%	99%
Overhead	4.79	0.11

The graph 4.1 shows the result of existing system and proposed sequential probability ratio test regarding detection rate (replica nodes).

The result of existing system and proposed sequential probability ratio test regarding overhead is shown in the graph 4.2.

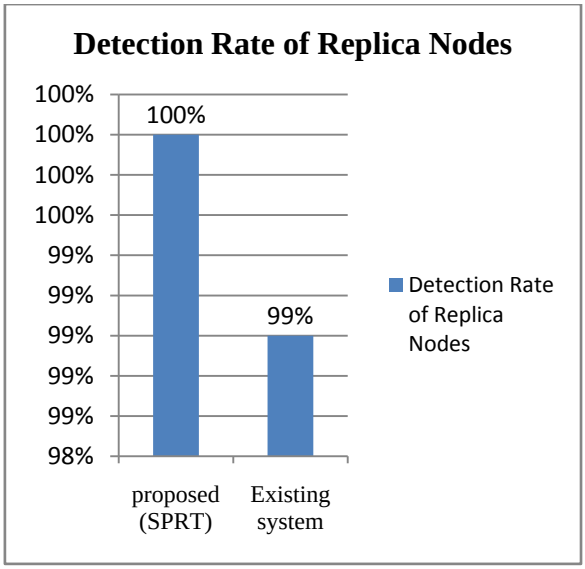

Figure 4.1 Graph of existing system and proposed SPRT regarding detection rate (replica nodes).

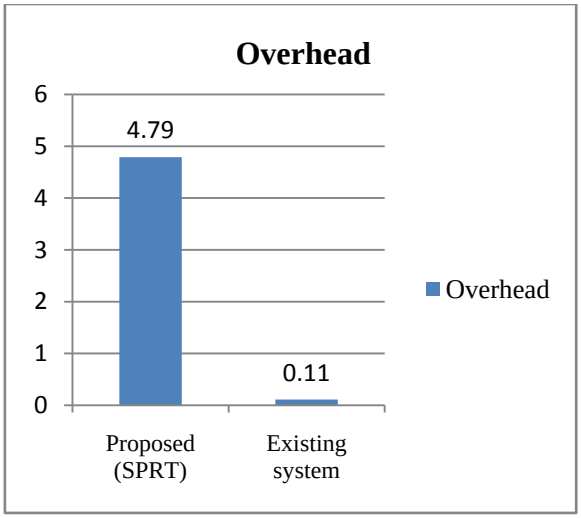

Figure 4.2 Graph of existing system and proposed SPRT regarding overhead.

5. CONCLUSION AND FUTURE WORK

The replica node detection scheme for mobile sensor network based on the SPRT is proposed. The experimentation showed limitations of a group attack strategy in which the attacker controls movements of a group of replicas. The work limits amount of time for which a group of replicas avoids detection and quarantine. The Proposed work contains interaction between detector and the adversary. The scenarios of proposed work simulates under a random movement attack strategy in which the attacker lets replicas randomly move in the network and under a static placement attack strategy in which one keeps the replicas from moving to best evade detection. The scheme quickly detects mobile replicas with a small number of location claims against either strategy.

In future, the theoretical capacity bounds may be used to prevent network overhead with rate control mechanism.

6. REFERENCES

[1] R. Di Pietro, L. V. Mancini, C. Soriente, A. Spognardi, and G.Tsudik, "Data security in unattended wireless sensor networks, IEEE Transactions on Computers, vol. 58, no. 11, pp. 15001511, 2009.

[2] Haafizah Rameeza Shaukat, Fazirulhisyam Hashim, Aduwati Sali, and M. Fadlee Abdul Rasid "Node Replication Attacks in Mobile Wireless Sensor Network: A Survey", In International Journal of Distributed Sensor Networks, Volume 2014, Article ID 402541, 15 pages.

[3] Navendu Jain, Mike Dahlin and Renu Tewari "TAPER: Tiered Approach for Eliminating Redundancy in Replica Synchronization" FAST'05:4th USENIX Conference on File and Storage Technologies.

[4] W. Znaidi, M. Minier, and S. Ubeda, "Hierarchical node replication attacks detection in wireless sensors networks," in Proceedings of the 20th IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC '09), pp. 82–86, Tokyo, Japan, September 2009.

[5] W.Zhang, G.Cao s, Optimizing tree reconfiguration for mobile target tracking in sensor networks in IEEE INFOCOM, 2004.

[6] Sandip D. Girase and Ashish T. Bhole Detection of Replica Nodes in Wireless Sensor Network: A Survey in IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN 2278-0661, p-ISSN: 2278-8727, Volume 17, Issue 2, Vol. I (Feb-Mar. 2015).

[7] J.Jung, V.Paxon, A.W.Berger and H. Balakrishnan Fast Portscan Detection Using Sequent Hypothesis Testing Proc.IEEE, Int`l Symp.Security and Privacy, pp.211-225, May 2004.

[8] Ramesh L, Dr. A.Marimuthu "Random direction based model for Intrinsic Secrecy in Wireless Sensor Network" International Journal of Science And Research (IJSR) ISSN (Online) 2319-7064.

[9] Ashok Kumar Das "An Identity Based Random Key Pre-Distribution Scheme for Direct Key Establishment to Prevent Attacks in Wireless Sensor Networks"International Journal of Network Security, Vol.6, PP.134-144, March 2008.

[10] Moirangthem Marjit, Ankita Singh and Jyotsna Kumar Mandal "Towards Technique of Detecting Node Replication Attacks in Static Wireless Sensor Networks" in International Journal of Information and Computation Technology. ISSN 0974-2239 Volume 4, Number 2(2014), pp. 153-164.

[11] K.Xing, X. Cheng, F. Liu, and D.H.C.Du, Real-time detection of clone attacks in wireless sensor networks, in Proceedings of the 28th International Conference on Distributed Computing Systems (ICDCS 08), pp. 3-10, Beijing, China, July 2008.

[12] B.Parno, A. Per rigand V.D Gligor Distributed Detection of Node Replication attacks in Sensors Networks, in Proc. IEEE Symp. Security and Privacy, pp.49-63, May 2005.

[13] X. M. Deng and Y. Xiong, A new protocol for the detection of node replication attacks in mobile wireless sensor networks, Journal of Computer Science and Technology, vol. 26, no. 4, pp. 732-743, 2011.

[14] Balaji. N and Anitha.M, Efficient and Distributed Detection of Node Replication Attacks in Mobile Sensor Networks, in International Journal of Research in Engineering and Technology pISSN: 2321-7308.