A Mobile Agent based Sybil Attack Detection Algorithm for Wireless Sensor Network

Kamdeo Prasad
M.Tech, Student
Department of Computer Science & Engineering
Centurion University of Technology and Management, Odisha

Chandrakant Mallick
Assistant Professor
Department of Computer Science & Engineering
Centurion University of Technology and Management, Odisha

ABSTRACT
In the recent years it has been a great challenge for the researchers to ensure security in Wireless Sensor Network (WSN). The traditional prevention based cryptographic algorithms cannot meet the security goals in a resource constrained wireless sensor network; these cannot also prevent the insider attacks. So a second line of defense called intrusion detection systems are developed to detect and alarm on various kinds of insider as well as outsider attacks in Wireless Sensor Networks. WSNs are susceptible to many kinds of attacks because of their open and harsh operating environment. Sybil attack is one of the dangerous attacks in terms of resource usages and poses threats to many security goals. In this paper, an algorithm called Sybil Attack Detection Algorithm (SDA) is proposed to detect and prevent the Sybil attack in the WSNs. Our proposed SDA algorithm is dynamic and accurate in detecting the Sybil attack that uses Mobile agent, threshold value, random key pre-distribution & random password generation. We are using random password and threshold value to distinguish and then for the confirmation of a legitimate node and a Sybil node. Moreover our algorithm helps in transmission of data in a more secured way by avoiding the Sybil attacks. We have simulated the proposed algorithm in NS2. We have checked the throughput and packet delivery ratio hence verified the detection performance of SDA in wireless sensor network.

Keywords
Wireless Sensor Networks, Mobile Agent, Sybil attack Detection Algorithm, Intrusion Detection.

1. INTRODUCTION
1.1 Wireless Sensor Network
Wireless Sensor Network can be regarded as a collection of sensor nodes that work in a collaborative manner to sense the environment phenomena. The applications of sensor networks typically range from surveillance systems to battle field [1]. Sensors are typically portable, self-possessed, battery-powered and low cost devices that measure physical parameters like light, temperature, or pressure in a particular application area. The basic task of sensor networks is to sense the events, collect data and send it to their requested destination [2]. In case of wireless sensor network, the communication among the sensors is done using wireless transceivers. The magnificent features like flexibility, fault tolerance, Scalability, high sensing capabilities, low cost, and rapid deployment of the sensor networks enable the WSNs for many exciting applications. However its stringent constraints such as limited resources of sensor nodes, multi-hop, insecure, short ranged radio communication, node and link failures, mobility of nodes and change of network topology, broadcast network and data redundancy, power consumption etc. pose many new challenges. These challenges, attracted many researchers to work on various issues and challenges. In the current scenario, the routing strategies and wireless sensor network modeling are getting much preference; at the same time security issues are yet to receive extensive and potential focus [3].

1.2 Intrusion Detection System
An intrusion-detection system (IDS) can be defined as a set of tools, methods and resources to identify and report unauthorized or unapproved network activity. The purpose of intrusion detection is to serve as an alarm mechanism for a computer system or a network [3]. The functionality of IDS is similar to that of a firewall but with only difference. A firewall acts as a hedge that protects the information flow and prevents intrusions from outside where as IDS starts detection when the network is under attack or when the security enforced by the firewall is penetrated. Together firewall and IDS enhance the security of network [4]. The intrusion detection systems act as a second line of defense in a wireless sensor networks. It also monitor network activity, analyze data integrity, audit network and system configurations for vulnerabilities.

1.3 Intrusion Detection Policies
Generally, intrusion detection policies are categorized into two types: misuse detection and anomaly detection.

1.3.1 Rule-Based/Signature-based IDS
Misuse detection algorithms detect attacks based on the known attack signatures. They are effective in detecting known attacks with low errors. However, they cannot detect newly attacks that do not have similar properties to the known attacks [5].

1.3.2 Anomaly-based IDS
The anomaly based IDS is based on the hypothesis that the attacker behavior differs to that of a normal user. They classify traffic as an attack if the characteristics of the traffic are far from those of normal traffic patterns. Anomaly detection algorithms can be useful for new attack patterns, but they are not as effective as rule based detection methods in the detection rate for known attacks [5].

2 ATTACKS
A human who exploits vulnerability perpetrates an attack on the system [17]. Vulnerability is a weakness in the security system. For instance, an unauthorized data manipulation in system is due to mis-verification of user's identity before allowing data access.

There are different types of attacks such as sinkhole attack, blackhole attack, DoS attack, Wormhole attack etc. in wireless sensor networks. We are mainly dealing with the Sybil attack.

2.1 Sybil Attacks
Sybil attack is first described by Microsoft researcher John Douceur. Sybil attack is named after the case study of a woman with multiple personality disorder. Sybil node is the process of creating two or more duplicate nodes with similar identity i.e. same node id as shown in Fig.1.
A Sybil attack is a type of security threat. A threat is a set of circumstances that has the potential to cause loss or harm [17]. Particularly, wireless sensor networks are more prone to Sybil attack because of the open and broadcast communication medium and the same frequency is being shared among all nodes. In Sybil attack, attacker makes multiple illegitimate identities in sensor networks either by fabricating or stealing the identities of legitimate nodes. So the base station cannot distinguish the legitimate and the forged node. This confuses the base station and other nodes and the network performance degrades. The different ways in which a Sybil attack could be performed are as followings [17]

I) Direct Communication
In direct communication the attack is performed through Sybil node by communicating directly with legitimate nodes. When a legitimate node sends a radio message to a Sybil node, one of the malicious devices listens to the messages. Similarly, messages can be sent from one of the malicious devices [15].

II) Indirect Communication
In the attacks through indirect communication, no legitimate node is able to communicate directly with the Sybil nodes. Here one or more of the malicious devices claims to reach the Sybil nodes. Messages sent to a Sybil node are routed through one of the malicious nodes in network, which pretends to pass on the message to a Sybil node [6].

III) Fabricated Identities
In this type of attack, the attacker can simply create arbitrary new Sybil identities. For instance, if each node is identified by an 8-bit integer, the attacker can simply assign each Sybil node a random 8-bit value or it creates a new identity for itself based on the identities of the legitimate nodes, that is, if legitimate nodes have an ID with length 8 bit integer, it randomly creates ID of 8 bit integer. These nodes have fabricated identities which pretend them as legitimate nodes [17].

IV) Stolen Identities
In stolen identities, attacker gets access to the identity of a legitimate node and then uses it. The attack may go unidentified if the node whose identity has been stolen is destroyed. Identity replication occurs when the same identities are used many times in the same places and the Sybil nodes need to be identified [17].

V) Simultaneous
In simultaneous type of attack, all the Sybil identities participate in the network at the same time. Since only one identity appears at a time, practically cycling through identities will make it appear simultaneous. The attacker may try to have his Sybil identities all participate in the network at once. While a particular hardware entity can only act as one identity at a time, it can cycle through these identities to make it appear that they are all present simultaneously. The number of identities the attacker uses is equal to the number of physical devices; each device presents different identities at different times [17].

V) Non-Simultaneous
It is opposed to simultaneous type of attack i.e. the attacker have a large number of identities over a period of time, but only fewer number of identities act at any given time. The attacker can do this by having one identity seem to leave the network, and have another identity join in its place. A particular identity might leave and join multiple times, or the attacker might only use each identity once [17].

3 RELATED WORK
3.1 Position based Sybil Attack Detection
Manjunatha et al. [9] proposed that network verifies the physical position of each node. Sybil nodes can be detected using this approach because they will appear to be at exactly the same position as the malicious node that generates them. By placing a limit on the density of the network, in-region verification can be used to tightly bind the number of Sybil identities that a malicious node can create.

3.2 Ant Colony based Sybil Detection
Zeng et al. [10] proposed a novel protocol to limit the influence of Sybil attacks by combining ant colony optimization (ACO) algorithm, where a node does random walks and it will leave trails on the path. However, with the random work, the trails of the first node left on each node become diluted, at the end of routing it just fades away. Based on the nature of the ACO and limit the number of attack edge in an efficiently and conveniently. Hence, system can ensure an honest node would accept and be accepted by other honest nodes in system with high probability, also, reject Sybil nodes with great probability.

3.3 Random Password based Sybil Detection
Amuthavalli et al. [11] proposed an algorithm to detect Sybil nodes in WSN called Random Password Comparison (RPC). This algorithm deploys and controls the position of node thereby preventing the Sybil attack. The RPC method is dynamic and accurate in detecting the Sybil attack. This method improves data transmission in the network and will also increase the throughput.

Technique of the RPC algorithm: It gives three values c1, c2, c3 for every node and they are stored in the r-table. When a route is discovered from source node to destination node, every nearest node should submit their id, time, password (pwd) which is then compared with the r-table values. If the node is legitimate then it is chosen as the nearest neighbor and added to the route. The RPC algorithm is divided into two parts. DISROUT (node n, node m): The algorithm compares the node id, password and time of generation of password with the r-table. CHKROUT (node n, node m): The sub procedures DISROUT and CHKROUT detect the route from source to destination before transmitting the data. These comparisons require the RPC method to assign random values to all the nodes to check if the nodes are normal nodes or not.

3.4 Threshold based Sybil Detection
Sharmila et al. [12] proposed an algorithm to find a Sybil node in WSN. Their technique is divided into three phase in first phase they send ID and power value to the head nodes, the head node checks for nodes with power value below the threshold value. In second phase, the distance between the receivers and sender are zero, and then the node suffers from Sybil attack after that if the nodes are very close, then the nodes will be detected as Sybil nodes even if they are not. In third phase, the routing procedure in the cluster is checked to verify
if there was a hop between the Sybil identities, and if there exists a hop between the Sybil identities, then the nodes are not Sybil nodes.

3.5 Mobile Agent based Sybil Detection
Sheela et al. [13] proposed an algorithm to detect clones and any malicious nodes in wireless sensor networks. Their system has two algorithms based on mobile agent. MARCAD (Mobile Agent based Cloning Attack Detection) is to detect clones and any malicious nodes. Another is MASAD (Mobile agent based sinkhole Attack Detection) algorithm is to tell how a node uses the global network information to route data packets by avoiding sinkhole attack.

3.6 Clone based Sybil Detection
Rapinder et al. [14] proposed an algorithm Mobile Agent Based clone Attack Detection Algorithm (MACAD). In MACAD algorithm, the system is designed to make every node aware of location and identity of many nodes (Say n) so that each neighbor of node A verifies the signature and checks the plausibility of Location of A. When a node finds a collision different location claims with the same ID. It broadcasts the two conflicting claims as evidence to revoke the replicas.

3.7 Key Pre-distribution Sybil Detection
Newsome et al. [15] proposed random key pre-distribution, assign a random set of keys or key-related information to each sensor node, so that in the key set-up phase, each node can discover or compute the common keys it shares with its neighbors; the common keys will be used as a shared secret session key to ensure node-to-node secrecy.

Those ideas are:

1. Associating the node identity with the keys assigned to the node.
2. Key validation, i.e., the network being able to verify part or all of the keys that an identity claims to have.

3.8 Authentication based Sybil Detection
Karen et al. [16] proposed Key-based authentication. This was extensively studied in terms of its storage overhead, computational efficiency and resiliency against fractional compromise of the network. It has a nice property that the amount of overhead can often be modeled as a function of security resiliency. A node may have only one master key shared with everyone, a group key shared with a group of nodes, a cluster key shared with all its neighbors, or a pairwise key shared with each immediate neighbor.

4 PROPOSED WORK
We have proposed a Sybil Detection Algorithm (SDA) to detect multiple identity nodes i.e. Sybil nodes in wireless sensor network (WSN). Through this technique we are resolving the problem of duplicate nodes in Sybil attack. In the proposed work we detect the Sybil nodes based on Mobile Agent and using random key pre-distribution, random password and threshold value taking together.

4.1 Components

4.1.1 Mobile Agent
Mobile Agent is an independent computer program that executes continuously in cross-platforms. It has the abilities of self-control moving, imitating human behavior and relationships, and providing certain Artificial Intelligent services. It can autonomously move in heterogeneous network according to certain rules and searching for suitable computing resources, information resources or software resources [7].

The advantages of mobile agent are space savings, reduction in network traffic, asynchronous autonomous interaction, interaction with real-time systems, robustness and fault tolerance, support for heterogeneous environments, on-line extensibility of services, convenient development paradigm and client customization etc [8].

There are many applications of mobile agent. The intrusion detection process is our main concern. A lot of problems such as centralization or partial distribution, static reconfiguration, vulnerability to direct attacks and limited response capabilities etc are associated with the traditional intrusion systems (IDSs). But mobile agent technology offers the potential to overcome a number of limitations intrinsic to existing Intrusion Detection Systems (IDSs).

Defining and collecting data which is used as input to the intrusion detection engine is a main issue in IDSs. Using a mobile agent retrieval of information and tracing intrusion by IDSs becomes easy [18]. After intrusion detection, the intrusion must be either brought to the attention of the system administrator or to an automatic response system to take some action i.e. blocks the detected intrusion packets [8]. Mobile agents can helps in IDSs by accomplishing the tasks like monitoring the WSNs, decision-making, notification as well as reaction to attempted intrusions.

4.1.2 Threshold Value
According to recent experimental studies, the non-ideal unreliable communication links in sensor network have the impact on variable transmission power. So taking the power into consideration we can detect the Sybil nodes. The threshold value (THg) contains the minimum PRR (packet reception rate) requirement for a good link decision. The links with quality better than THg, in only one direction are defined as asymmetric links or weak links [20].

We choose a node with minimum packet drop, called as “Faith node (F_a)”. The F_a node become a lead node with a group of its own member nodes. The member node sends its ID and power value to the lead nodes. The lead node checks the nodes with power value below the threshold value. If the power value is lesser than threshold value, those nodes are detected as Sybil nodes.

4.1.3 Random Key Pre-distribution
Random key pre-distribution for WSNs was first proposed by Eschenauer and Gligor [19]. In a random key pre-distribution scheme, each node is assigned a set of keys drawn from a much larger key pool. In our proposed model, we use the random password generator in random key pre-distribution process. This process is accomplished in three phases.

I) Key Generation: It takes place prior to network deployment. A large set of random keys along with node ids are generated.

II) Key Distribution: Each node is assigned a random key, drawn from the set of random keys at random, without replacement and without repetition.

III) Shared-key discovery: It takes place during network setup. All nodes have a specific node id as well as a random key. Those keys will be discovered and will be used along with the pre-random password generated by the Random Password Generator (RPG) Algorithm to generate a random password.

4.1.4 Random Password Generator
A random password generator generates a new password which is a random value, every time when the sensing process starts. These values will be stored in the pwd_table. Through the RPG Algorithm a random password will be generated. When a source node communicates with the destination node, the source node’s id, the random password and the threshold value if the source node is compared by the mobile agent. If the node id, corresponding to the threshold value along with the password is matched, the source node is considered as a
legitimate node and allowed to send data otherwise the node is considered as a Sybil node and blocked and this information is sent to the base station through the corresponding cluster head.

4.1.5 Random Password Generation (RPG) Algorithm

Random Password Generation (RPG) Algorithm describes the process of generating the random password.

Step 1
Record the packet sensing time of a particular node in 00:00 format.

Step 2
Generate ten random values in the respective column of the particular node staring from 0 to 9 in pwd_table in Table 3.

Step 3
Note the individual digit sequentially from left to right of above described format of recorded time.

Step 4
Pick the random value from the respective row of pwd_table for individual digit (Table3).

Step 5
Using the chosen random value, generate a pre-random password.

Step 6
Generate a random password using pre-random password along with the random key.

4.1.6 Random Password Generation Process

Here in the process of generating the random password, we will take help of some components. Those are i) Time of packet sensing ii) Random password/value generator.

The pwd_table of cluster head contains the node_id and a sequence of 0-9 in row as shown in table3.

A random value is generated for the particular sensing node when the sensing process starts. Ten values from 0 to 9 is generated and saved.

The time of packet sensing of the particular node is recorded. Suppose the time is 12.30 then all the individual number i.e. 1, 2, 3, 0 are considered, which we call “pick value”. The values from the respective node_id column will be chosen for each pick value. So the values are R, A, C, B for node_id1. Using all the above values a pre-random password is created. Finally the pre-random password along with the random key will generate a random password.

The random value in the pwd_table for a particular node will reset again for next sensing process.

4.2 The Database Tables

In dealing with Sybil node detection we are using few database tables at the mobile agent, cluster head, base station and sensor nodes. Those are namely ptable, ctable, and pwd_table.

4.2.1 ptable

This ptable is the table of mobile agent. This table contains the following data as shown in table 1.

<table>
<thead>
<tr>
<th>Node_id</th>
<th>Threshold Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>4.5</td>
</tr>
<tr>
<td>N2</td>
<td>3.4</td>
</tr>
<tr>
<td>..........</td>
<td>.............</td>
</tr>
</tbody>
</table>

4.2.2 ctable

This ctable is the table of mobile agent. This table contains the following data as shown in table 2.

<table>
<thead>
<tr>
<th>Node_id (Sybil node)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
</tr>
<tr>
<td>N2</td>
</tr>
<tr>
<td>..........</td>
</tr>
</tbody>
</table>

4.2.3 pwd_table

This pwd_table is present in the cluster head and the sensor node. These are the following data as shown in table 3.

<table>
<thead>
<tr>
<th>Node_id</th>
<th>Node_id 1</th>
<th>Node_id 2</th>
<th>..........</th>
<th>Node_id n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>B</td>
<td>A</td>
<td>..........</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>R</td>
<td>F</td>
<td>..........</td>
<td>N</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>P</td>
<td>..........</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>Z</td>
<td>..........</td>
<td>E</td>
</tr>
<tr>
<td>9</td>
<td>K</td>
<td>I</td>
<td>..........</td>
<td>R</td>
</tr>
</tbody>
</table>

4.2.4 btable

This btable is the table of base station. This table obtains the data from the ctable and keeps the record of the Sybil nodes in the network. The data is as shown in the following in table 4.

<table>
<thead>
<tr>
<th>Sybil node_id</th>
<th>Cluster_id</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>C3</td>
</tr>
<tr>
<td>N2</td>
<td>C6</td>
</tr>
</tbody>
</table>

4.3 Sybil attack Detection Algorithm (SDA) Algorithm

In our proposed model (Fig.2) we are using the three main components in wireless sensor network (WSN). Those are sensor nodes, cluster head nodes and base station. The cluster head nodes and base station have two tables namely ctable and btable respectively. The mobile agent has a table named as ptable. The algorithm goes through the following steps.

Step 1
Generate random_key as well as node_id. //Before network deployment

Store these in ptable of mobile agent.

Step 2
Assign the node_id and random_key to each node.

Step 3
Random password is generated in reference to the random_key.

Step 4
The sensor node sends data through the mobile agent, to the cluster node along with random_password.
Step 5
Mobile agents check the node_id in its stored database (ptable).

Step 5.1
If node_id matches Check for random_password as well as threshold value.

Step 5.1.1
if random_password and threshold value matches The legitimate node is confirmed. Data received by respective cluster head from legitimate node securely.

Step 5.1.2
else
The Sybil node is detected and blocked by Mobile agent. ctable data is sent to base node. Store the data in b table.

We also represent the steps of Sybil detection algorithm for the implementation point of view.

Pseudocode
Set Node_ID=n(n1,n2,n3,......),Random_Key=r (r1,r2,r3,........),Threshold_value=n_th(n_th1,n_th2, n_th3,.......) store in ptable
Generate Random_Password p (p1,p2,p3,........).
mobile_agent(n,p,n_th) {
  if(n==n') / /n'=node_id stored in ptable of mobile agent
  if (p==rp && n_th==threshold_value) //rp and threshold_value stored in ptable.
    data sent to destination.
  }
else
  {“Detect Sybil node” and Store n in ctable
   Send ctable data to btable
  }
}

Fig 2: Flow diagram of SDA

4.4 Procedure
In the process (Fig.3) of Sybil node detection, we are first generating the random_key as well as node_id before network deployment. The generated random_key as well as node_id is assigned to each individual node.

When the sensing process starts, a random password is generated using the Random Password Generator (RPG) algorithm. The generated random passwords of respective sensing nodes are circulated through the Mobile agent.

When the data packet travels in the network to the destination (Base station), mobile agent tracks it and checks for the node identity. So it checks the node_id. If the node_id matches then the random_password and threshold value for that node is checked. The threshold value stored in the database of mobile agent is compared, if the value matches then that node is considered as the legitimate node. The rest of the nodes with same node_id but unmatched random_password and threshold
value, are detected as Sybil nodes. The detected Sybil node is stored in ctable of cluster head. The Sybil node information from the ctable is transmitted to the btable of the base station, so that on further communication the compromised node can be easily detected and Sybil nodes are blocked. The data from legitimate node is sent through the respective cluster head to base station. The base station then keeps the record of the Sybil nodes and informs the same to other nodes to alert. The user can analyze the network performance, throughput and packet delivery ratio to conclude that the network suffers from Sybil attack.

Fig 3: Working of SDA Algorithm

5 SIMULATION

The proposed Sybil Detection algorithm is implemented in NS2. The process of Sybil nodes detection in wireless sensor network is simulated. The performance of the algorithm is compared in terms of packet delivery ratio and throughput before and after the attack.

The parameters used in our simulation are shown in Table 5.

Table 5. Simulation parameter

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>10000m X 10000m</td>
</tr>
<tr>
<td>Nodes</td>
<td>50</td>
</tr>
<tr>
<td>Packet size</td>
<td>512bytes</td>
</tr>
<tr>
<td>Transmission protocol</td>
<td>UDP</td>
</tr>
<tr>
<td>Simulation time</td>
<td>80 Seconds</td>
</tr>
<tr>
<td>Propagation</td>
<td>Two Ray Ground</td>
</tr>
<tr>
<td>Application Traffic</td>
<td>CBR</td>
</tr>
<tr>
<td>Queue type</td>
<td>Drop tail</td>
</tr>
<tr>
<td>Antenna Model</td>
<td>Omni directional area</td>
</tr>
<tr>
<td>Routing Protocol</td>
<td>AODV</td>
</tr>
<tr>
<td>Initial energy</td>
<td>100 joules</td>
</tr>
</tbody>
</table>

5.1 Simulation Results

Fig 4 and Fig.5 shows the network deployment and malicious Sybil node detection respectively. The number of nodes created is 50 and node 0 is the base station. After the assignment of node_id and random_key to each node the data packet was sent to cluster head. Through mobile agent the node_id, random_password and threshold_value is checked to detect the Sybil node. The Sybil nodes detected are 7, 29 and 34.

Fig-6 and Fig-7 represents the network throughput before and after Sybil nodes detection.

Similarly Fig-8 and Fig-9 represent the packet delivery ratio before and after Sybil nodes detection.

It has been observed that the network throughput has been consistently improved after the detection and isolation of the compromised nodes. Likewise the packet drop has been decreased and hence the packet delivery ratio has been improved after the detection and blocking the Sybil nodes.
CONCLUSION AND FUTURE WORK

Sybil attack is one of the serious security issues for wireless sensor network. In wireless sensor networks, similar identities or duplicate nodes confuse the entire network and degrade the network performance significantly. The network may even collapse. So, these attacks create major challenges to the security in WSNs. Our proposed SDA method based on mobile agent uses three major parameters namely, random key pre-distribution, random password, and threshold value to detect the Sybil node. The Sybil node detected is informed to the base station and the base station keeps the record of Sybil nodes for security purpose and alerts the network nodes. The network performances before and after the Sybil attack detection, is evaluated in terms of throughput and packet delivery ratio to analyze the adverse effect of Sybil attack.

In our future work, we will compute and compare the detection rates (i.e true positives and false positives) and the detection ratio and analyze the performance for a different set of simulation parameters. We will also verify the energy efficiency of the proposed SDA method to be cost effective.

REFERENCES


