IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

Anomalies and Selective Forwarding Attacks in WSN-A Survey

Arya I S[1], Dr Binu G S[2]

PG Scholar, Department of Electronics and Communication Engineering, NSSCE, India [1]

Associate Professor, Department of Electronics and Communication Engineering, NSSCE, India [2]

Abstract: Wireless Sensor Networks **(**WSN networks) are becoming an important topic in the eyes of researchers because of its vast variety of applications in various fields. WSNs are application specific networks which have a large number of small, randomly deployed self-configurable sensor nodes, capable of sensing an event, computing and communicating the event details with the neighboring sensor nodes. In many technologies, like Internet of Things (IoTs), WSNs act as a digital skin by which the information about the outside world is obtained. WSNs are used in many applications, in particular, Personal applications like home automation, Business applications like sales tracking, Industrial applications, and military applications etc. So the task of WSN is not only sensing but also the computation of data which is their responsibility and hence the quality of data is very crucial. Since there are no particular criteria for the deployment of sensor nodes, the hostile environments in which they are deployed and their energy constraints, they are very much vulnerable to anomalies and attacks. Any deviation from the normal sensing profile can be considered as an abnormality in the network. This paper gives an overview of such cases in a WSN.

Keywords: Wireless Sensor Networks, Anomalies, Selective Forwarding Attack.

I. INTRODUCTION

Wireless Sensor Network (WSN) consists of a large number of tiny, low cost, low energy self-configurable sensor nodes that are randomly deployed in order to monitor the events in the environment. They are domain specific in a way such that sensing and computations are done based on the scenario for which they are deployed. Since the size of sensor nodes is small, the battery associated with the sensor nodes is also very small. Hence the battery usage should be minimum as possible in order to prolong the lifetime of a wireless sensor network. Two main architectures that WSNs prefer are the flat based structures and the cluster based structures [4]. In flat based structures, all the sensor nodes broadcast data to its neighboring nodes and the data reaches the sink by multi-hop communication.

In cluster based approach, the total area is divided into clusters. Since the sensor nodes are randomly deployed, the number of nodes in each cluster will not be same. A cluster head is selected based on the highest amount of energy a sensor node has in a cluster. In order to increase the lifetime of the network, the role of cluster head is switched between the nodes in a cluster. Every node in a cluster communicates with the cluster head and the cluster head forwards the data to the base station. Cluster based approach is preferred because it can meet energy constraints. The data that is collected by WSN network can be either univariate or multivariate. Data types are similar in univariate data whereas the data types are different in multivariate data. Each data type is called an attribute or a feature. If the relations between all attributes are correct in a multivariate data then high accuracy can be obtained. The relations can be estimated by taking spatial and temporal correlations [20]. Spatial correlations provide the relation between sensor nodes located in the same geographical area and temporal correlations provide the relation between the data collected during different time periods. By using these techniques any mismatch in the data collected by the sensor nodes can be detected from the normal sensing profile. If any abnormality occurs without the cause of an external entity then it is called an Anomaly[14].

Anomalies are mainly classified as Node anomaly, Network anomaly, and Data anomaly. Node anomalies occur when nodes are deployed in harsh environments and are related to the battery issues of a node. Network anomalies occur to a group of nodes. Every node communicates with their neighbors. If the communication is interrupted by any manner then it can be considered as a network anomaly. Network anomalies mostly occur because of attacks. Data anomalies are the irregularities in the sensed data. These irregularities can be found out by spatial-temporal correlations.

Apart from the abnormalities happening within a network, abnormalities can happen because of other factors, mainly because of attacks. The main attacks that happen in a WSN network are Tampering, Sybil attack, Hello flood attack, Jamming, exhaustion, wormhole attack, Identity replication attack and selective forwarding or sinkhole attack[17].In Tampering, attacker access a node to obtain the cryptographic parameters especially key. In Sybil attack, the attacker node produces fake multiple identities. In Hello flood attacks, the attacker floods messages which prevent other

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

messages being exchanged between the sensor nodes. In Jamming, the attacker disturbs the radio channel by keeping it busy by sending useless information on the frequency band. In Exhaustion, the attacker attacks a node in such a way that it makes the node to perform unnecessary operations so that the node consumes all its energy and leads to node failure.

In Wormhole attack, the attacker node acts as a wormhole and drops the data that is passed through it. In Identity replication attack, the same identity is given to different nodes. In selective forwarding attacks or sinkhole attacks, the attacker node act as a compromised node with an improved routing pattern [18]. It attracts the neighboring node and all the neighboring nodes forward data to the attacker node misunderstanding that it is the shortest distance to the base station. Because of the broadcasting nature of wireless sensor networks, it isn't difficult to get the routing information of the nodes. Therefore WSNs are more prone to sinkhole attacks.

II. CHALLENGES TO IMPLEMENT DETECTION MECHANISMS IN WSN

A. Random Deployment Of Sensor Nodes

Because of the randomness in sensor node deployment, WSN nodes are very much vulnerable to attacks. So it is difficult to keep an anomaly/attack detection system to monitor the network since the number of nodes within each cluster is different.

B. Energy/Power constraints in WSN

The size of sensor nodes is very small and so the size of the battery associated with it is very small. WSN nodes perform sensing, computation and communication equipped with low energy resources and hence is an important parameter in WSN. In order to increase the network lifetime, the energy consumption should be minimized.

C. Limited Memory of Sensor Nodes

In early days off-line methods were used to detect the abnormalities/attacks. Here the data collected is stored in memory. When new data arrives, it is compared with the previously stored data and based on the comparison results, detection is done. Since WSN nodes have limited memory storage, such systems aids the inherent difficulty to implement detection systems.

D. Continuous Streaming of Data

Wireless Sensor Networks operates in real time. So there is a continuous flow of data between the sensor nodes. Hence the system should update itself continuously. It is difficult to keep a detection system based on the predetermined data since the network collects multivariate data.

E. Heterogeneity of Sensor Nodes

Nodes are deployed in a large area. Because of the adverse and hostile environments, the nodes may exhibit different configurations. They have different energy levels, power consumptions and lifetimes.

III. ANOMALY DETECTION TECHNIQUES DESIGNED FOR WSN

Statistical Based Anomaly Detection

It is considered as the oldest anomaly detection technique [14,21]. It is used in one-dimensional data. Here probability distribution is adopted which represents the reference model. Any deviation from the reference model is considered as an anomaly.

They are classified into parametric and non-parametric methods. In the parametric method, the data is generated from a known distribution and then the parameters are estimated from the data. In Non-parametric method, estimation techniques, histograms are used to estimate the distribution models from that reference models are created. Since WSNs are operating in real time, prior distribution of data is not possible.

Histogram based approaches are useful in univariate data but difficult to use in multivariate data. Since WSNs are deployed over a large area the probability distribution from different areas is not the same. These are the main drawbacks of Statistical approaches.

Because of the limitations, new approaches emerge in the field of anomaly detection.

Nearest-Neighbor Based Anomaly Detection

It is based on the concept that the normal patterns of data are always found in a dense neighborhood while the anomalous ones are usually far from the neighborhood [15,16]. These are based on the use of similarity measures which measures the degree of the normal pattern [22,21].

Neighborhood characterization is done based on the amount of data gathered from the neighborhood. This can be done by cross-correlating between the local sensor and aggregated information. The size of the neighborhood is also considered. It can be estimated by correlating the network density and checking how well the aggregated information correlations depend on the size of the neighborhood. PLE (path loss exponential) is also considered. It dictates the decay in signal strength over distance. Evaluation matrices are used to determine the network performance.

The confusion matrix is used in anomaly detection [32]. It is a 2*2 matrix used in the predictive analysis. True positive value occurs when an anomaly is present and node detects it. False positive occurs when node detect an anomaly when there is no anomaly. True negative occurs when no anomaly has occurred and node detects none. False negative occurs when node detect no anomaly even if there is some.

The main disadvantages are scalability problems and the computation of the distance between data patterns of multivariate data is difficult.

Clustering Based Anomaly Detection

Here the data with similar patterns are grouped together as a cluster and then clustering technique is employed [14, 21]. A cluster is said to be anomalous if it is smaller

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

than or distant from other clusters in the data set. The main disadvantages are that clustering cannot cope with continuous data changes. This method is very expensive in the case of multivariate data. Here fixed width clusters are used, so fixing the width of each cluster adds more difficulty.

A cluster can be expressed as a centroid and effect influence radius [33]. This format has a centroid field, radius field, and type field. Centroid field represents the centric vector of a cluster. Radius is calculated by euclidean distance measurements. Type shows whether the vector is normal or malicious. To find whether a vector is in cluster region or not, the radius is checked, if the radius is less than the coverage of a cluster from the centroid then it is considered to be in the cluster. The class with less purity has a number of attacks and class with high purity has less number of attacks.

Classification Based Anomaly Detection

Here a classifier is used. The classifier is trained by using training data patterns and used to classify unknown packets into different types [14,21]. So any data pattern that does not match the training patterns can be considered as anomalous. The spatial and temporal characters of WSN is mapped to Bayesian network parameters and then the interference provided with suitable conditions are made for computation. By using SVM (Support Vector Machine) and sliding window approach, classification based anomaly detection can be done.

Influence function: Influence function is a mathematical tool that is used to find the influence of an object. In general, it is a distance function [31]. Square law influence function, electric field based influence function, and Gaussian influence functions are the frequently used influence functions.

IV. SELECTIVE FORWARDING ATTACK IN WSN

In WSN the nodes advertise their information to their neighboring nodes. When sinkhole attack happens the affected node advertises an improved routing pattern and all the information from the neighboring nodes is send to the sinkhole node [19]. When an attacker node receives this message they either act as a wormhole or they selectively forward the data [23]. The Sinkhole node is placed at the center when it acts as a wormhole.

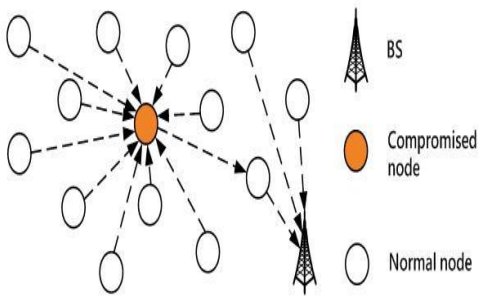

Fig.1.Sinkhole attack in the WSN network [23]

The energy efficiency and security can be achieved by using good routing protocols. Geographic routing protocols are used; it is referred as position based routing.

In WSN a base station and geographically distributed sensors are present. These sensor nodes have a unique ID [9]. Because of the 'many to one' nature of WSN, they are very vulnerable to sinkhole attack. If a node is provided with high-quality single hop link to the base station, then it can act as a sinkhole. In WSN network, it should be identified as which node is good and which node is bad.

Security solutions in WSN networks can be done in two categories. [5]

1) Prevention based approaches

Prevention based approaches use cryptographyand authentication. But it is not practical in a WSN because these methods increase the computational complexity and also because of the broadcast nature, attackers can easily get the key.

2) Detection based approaches

These are based on the system behaviors; based on the homogeneity or heterogeneity of the systems.

Proposed Detection Techniques in WSN

Rule-Based Approach

Rule-based approaches with forward chaining are adopted. In rule-based approach, the inferences were done with predefined rules [4]. If the condition satisfies then the conclusion is made. In forward chaining approach, the conclusion from the cause can be determined. But interfering with new levels is not possible. In backward chaining approach cause from the conclusion is determined. All possible causes for a particular conclusion can be considered.

Back Propagation Networks (BPN)

It generates trained data called epoch by processing in an environment where it gets input as well as the environment where it gets input as well as the output data [4]. It mainly has three layers: Input layer from which it gets information. The middle layer where all processing is done. Output layer where output is obtained. Many epochs are generated until the target value is obtained.

Adaptive Resonance Theory (ART) Networks

It has only input variable and does unsupervised learning. It uses the clustering information to know from where the new data comes from. If the data is not from any current cluster, then a new cluster is formed. It has two layers: Input layer for input variables and output layer for output variables. Then the trained data is generated until it converges to the targeted value. Since the attack happens at different time instants, online learning is adopted [4].

Detection in Sink

It can be done using anomaly detection module and misuse detection module. Anomaly detection module acts as a filter. Here the predetermined characters are the

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

parameters. Based on what is normal, that is the Detection in Sink

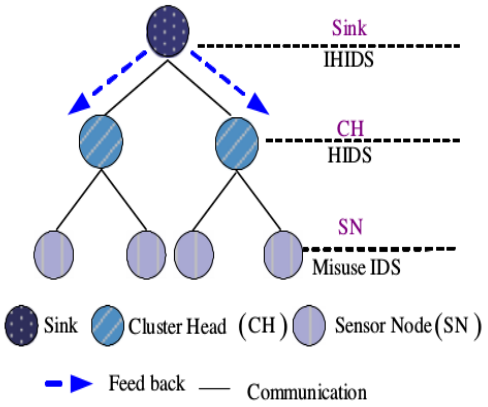

Fig. 2. Architecture [4]

It can be done using anomaly detection module and misuse detection module. Anomaly detection module acts as a filter. Here the predetermined characters are the parameters. Based on what is normal, that is the key features, energy levels of sensor nodes etc.., abnormalities are identified. Those packets are forwarded to misuse detection module for further analysis. Misuse detection module uses BPN. Here input-output variable relations are analyzed and corresponding weight is determined. Decision making is based on the rule-based approaches.
Since there is an advancement in technology, the existing IDS (Intrusion Detection System) system should be improved. So IHIDS must be an intelligent detection system. BPN sometimes avoids less frequent packets. So filtering is done initially. BPN data are processed to match each other and then send to the system.

Intrusion in CH
Here learning mechanism is not present because of the limited resources of CH. There is a feedback with the sink. Sink performs the learning mechanism and feedback in data. Using this data the misuse detection module is retained and it can identify new attacks.

Intrusion in SN
Only misuse detection is adopted because of the fewer resources. Rule based approach is done. It checks the abnormal behavior and detects the attack.

V. NEW APPROACHES TO DETECT SINKHOLE ATTACK

One of the main features of sinkhole attack is selective forwarding. When base station notices consistent data loss from a location, it is known that data is missing from that particular node. Then the area is selected in such a way that the attacked nodes are within the radius.

By identifying the routing pattern in the affected area, any change in the routing pattern can found out. Format of the request message is <ID_V,$ID_{next\ hop}$,Cost> [9]. Since the next hop ID and the cost are decided by the routing protocol, the reply message has the reverse path of that of flooding which is not exactly altered by the intruder. Consider 'a' is sending information to 'b'. 'c' is the cost. Hop information is used by the base station to realize the routing pattern. Based on this a tree network is generated. So when the routing pattern of the intruder node is changed, it can be known. When there are more than one malicious node then one of the exact sinkhole node which is represented as SH and the other colluding node which is represented as SH'.

Table 1

i	-2	-1	0	1	2
Count [i]	0	14	8	6	0

Table 2

i	-2	-1	0	1	2
Count[i]	0	1	20	6	0

The Count [ ] is analysed [9]. In table 1, 8 nodes agree that SH' is the intruder but 14 says that it is one hop distant from SH. Due to majority decision, a correction algorithm is run. In table 2, 20 nodes agree SH is the intruder node. Since it is the majority, SH is considered as the sinkhole node and this is how detection is done when there are a number of nodes. Hash chains can be used to provide authenticity.

Hash functions are used for providing authenticity in mobile agent-based approaches [7]. A mobile agent is a self-controlling segment which navigates from node to node, not only transmitting but also doing the computation. They are actually task-specific in nature. Mobile agents are activated periodically or on demand. By using a mobile agent scalability, energy awareness and reliability of the wireless sensor network can be increased. A mobile agent can only communicate with nodes. Agent packet is encapsulated in an agent packet object. Agent program contains #function named #AgentHashFunc() and two unique codes named code3 and data code. Code3 is the output of a hash operation performed on Code2 by #NodeHashFunction(). A copy of them is stored on the node memory. Nodes have Code1 and Code2. Code1 is not ciphered. Code2 is the output of performing #AgentHashFunction() on Code1. Agent node actually has the mobile agent. That means agent node is a normal sensor node that has received an agent packet from the base station. This mobile agent can be used to deliver data in case if we detect a sinkhole attack in a path.
Some of the hash functions used are MD5, RIPEMD, SHA-1. some modifications are made in MD5 and SHA-1 [12] to increase the security by introducing a padding logic. The constants are fed into the compression algorithms once the initialization is done. It is repeated

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

several number of times. And a message digest of 160 bits is generated. Double Davis Mayer Scheme is adopted. When a node is malicious, it alters the message. To check whether the message is transmitted through the exact route or transmitted through a newly advertised attacker node, the two packets are compared at the base station. If any change occurs in the packet, it can be detected. By using hash functions any small change in the input can have large changes in the output. So if any modifications are done then it can be easily identified. The main disadvantage is time wastage and energy wastage in transmitting same packets through different paths.

By using QoS based multi-path routing [2] the sinkhole attack can be detected. In multi-path routing, alternate paths existing between source and destinations are found. Data is segmented into sub-packets. These are transmitted to sink adopting different paths. The main disadvantage is the increase in communication overhead.

In RSSI (Received Signal Strength Indicator) [7] based systems, some Extra-Monitor nodes (EM Nodes) apart from the ordinary nodes are used to determine the position of all sensor nodes and thereby creating a visual geographic map of the wireless sensor networks. Whenever any sensor node in the network sends its message to the network, then all EM nodes will receive the message and then will send RSSI value to the RSSI based sinkhole detector to localize the position of sensor nodes on the map. If the flow of received message does not correspond with the normal flow of map, then sinkhole attack will be detected. In early days information from a single layer was taken,and with advancement in technology,cross-layer techniques emerge [24]. Here the information from five layers is shared. That is by taking information from different layers system performance can be improved. For sinkhole attack detection which is a network layer attack, network and mac layer data are used [1]. Network layer gives the information about the routing patterns and MAC layer informs about the medium that is accessed frequently or the number of nodes that actively take part in communication. By taking Packet Drop Ratio (PDR) of the system, any loss in the packet can be checked. If there is a sufficient decrease in PDR, then that area is analysed. If a path is found active more time than expected and if the routing pattern is improved than others, then it is the sinkhole node. In such cases, a mobile node can be used to deliver packets. In order to get data from a node, three-way handshaking procedure is performed using hash functions.

Cross-layer intrusion technique considering three layers of OSI layers, physical, MAC and network layers can also be considered [25]. It is done in such a way that when a node receives route requests to send a frame. It checks the routing table to know the existing transmitting node. If no transmitting node is found, then intrusion test is done. If the path exists RSSI is measured in order to check authentication. If it is not from an authorized node, intrusion detection can be done.

By checking the hope count [6] also helps to detect sinkhole attack. All nodes forward data to the node that is at a shorter distance from the sink. By changing the hop count to 1 then that is the node is the nearest neighbor of the node and the compromised node lure entire traffic from the affected area. So by checking the hop count also sinkhole attack can be detected.

By adopting different routing protocols, sinkhole attack can be prevented. Such type of a simple routing protocol is Min-route approach [5]. It works based on link quality factor(LQI). Every node calculates the LQI value from the Packet Delivery Ratio (PDR). Every node sends route update message periodically. When a node gets that message it updates the value based on PDR and broadcast the route update to the neighbors. The node with high LQI value is taken as the parent. If two nodes have same LQI then the one appears first in the neighbor table is selected as the parent. The main assumption is that the system is homogeneous. That is the nodes have a the same capability. They have flat based architecture. All the nodes communicate with the sink node in a single hop. The nodes make calculations and generate an alarm message for the sink. The sink makes decisions based on the alarm messages. In the proposed approach a detection mechanism is present in each node. In normal networks, some watchdogs monitor the network characters. Each node maintains a neighbor table that is updated based on LQI calculated using PDR. To detect the sinkhole the sender ID should not be the same as its node ID in the route_update, the sensor ID should be one of its neighbor ID. If any predefined rules in the route_update is violated then sink gets the alarm message, the attacker ID is broadcasted to remove the sensor ID from the neighbor tables of neighbor nodes. Rule-based approaches are used.

Sinkhole attack can also be detected by using neighbors information for Localized Encryption Authentication Protocol (LEAP) based WSN [13]. Localized encryption and authentication protocol has mainly four keys.

a) Individual key,Every node has an individual key to secure communication between the node and the base station.

b) Cluster key, Every node share a cluster key with its neighbors to secure broadcast information.

c) Pairwise key, It is shared with the immediate neighbor nodes.

d) Group key, Base station encrypt the data and give it to all sensor nodes in the network. This key is shared by the base station to all sensor networks.

Here the key is exposed and so it is difficult to detect the internal attacks. In such cases, Fuzzy logic systems are incoperated. "Numerical interpolation to address nonlinear problems in rule-based systems are used". The first step is the fuzzy matching step. Here the input is matched with the conditions of the fuzzy logic. The second step is to calculate the degree of the match based on the interference rule. The third step is the combining of all inference rules. The fuzzy rules, parameters, and member functions are calculated by iterative trial and error methods. The accuracy can be increased if neural networks or a genetic algorithm are used.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

Some of the cryptographic protocols to fight sinkhole attacks on tree-based routing was introduced in WSN [10]. This proposes two resilient and simple topology based reconfiguration protocols RESIST-1 and RESIST-0. The network model is designed based on the tree routed to sink. "Tree route is the aggregation of the shortest path to the sink from each node based on the cost metric" [34]. Public key cryptographic methods are present in WSNs. Sink has a public key. Every other node trusts the public key of the sink. Every node has a private-public key combination which shows their identity. Risk factor is calculated. Risk factor is a probability by which a compromised node gets a message from its neighbors. Resit-h protocols prevent malicious nodes to modify within a range of h-hops [8].

RESIST-1: Every node sends hello packet. Hello packet is {epoch,token}. Epoch is a strictly increasing time-stamp. There are two cases. First, if epoch is new, then the hello message is sent to the neighbors removing the smallest token value. Second, if epoch is known the nodes update by selfish approach or by gossip approach where the hello packets are transmitted to the neighbors. The malicious node will not change the smallest value so it is considered as the shortest path. If multiple malicious nodes are present then the number of hops in-between will be the number of malicious nodes.

RESIST-0: Procedure is same as that of RESIST-1. Here challenge message is sent,[challenge(k,epoch)]. The node that receives challenge sends challenge reply. Token is verified by using the public key. Then decryption is done by the private key of the node. If the node is compromised it can not understand the private key used and cannot reply to the challenge message [30].

Trustability based on beta distribution [11] is also used for detecting the abnormality in WSNs. The beta distribution is used where uncertainty is present. In order to get the overview of the network, the node receives the recommendation from other nodes along with its own findings[11]. A target node sends queries to the assessing node and then the information are taken together to model the network. A trust value is determined based on the current and previous value. Final trust value is generated from neighboring information.

To calculate the trust value of nodeB, the number of nodes that give a recommendation on the performance of Node B is taken. The current and past satisfied and dissatisfied values are calculated based on the good and bad interactions. Then the target node calculates the weight values. It is used to determine how much direct/indirect observations and recommendation values are used. Event based biasing," the method of accelerating simulation of useful events at the expense of accurate fluctuations" is used to make a comparison between current and last observations so that the nodes can punish bad behaviors faster leading to efficient malicious detection. Here assessing node uses observing node to get information about target node based on honest/dishonest recommendation value from the final value. Assessing node creates a weight by comparing the recommending node with other recommending nodes. A threshold is set for comparison by trail-error method.

Trust and repudiation models are also used. It maintains minimum security between the transmission entities in a communication system. The probability of an agent performing a particular task is called trust and expecting a node performing a particular task based on its previous behavior is called repudiation [26].

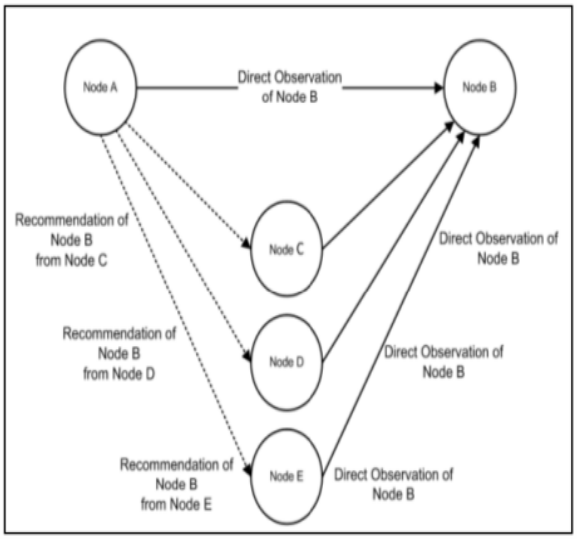

Fig. 3. Network Model [11]

Swarm optimization techniques are employed in trust and repudiation systems. Mainly ant colony optimization is used in bio-inspired security systems. Here pheromone value represents the trustworthiness of a sensor node. Pheromones represent the function of quality of the solution found.

In bio-inspired WSN, pheromone is used to determine whether a sensor is trustworthy based the average pheromone value of the solution path from the client sensor to the sensor selected. Similarly, the quality of each solution is evaluated based on average pheromone value and the length of those solution paths as well as the number of ants which have found the same solution.

Enhanced bio-inspired WSN have some modifications. The average pheromone value is compared with the predefined threshold value. The trust value of that sensor is computed by simplified peer trust algorithm [35]. 0-1 values are taken for satisfaction and dissatisfaction in order to get stable values. Path quality and trust value can be calculated.

To achieve energy efficiency clustering technique is employed. Low-energy adaptive clustering hierarchy (LEACH) is a routing protocol used for clustering. But it tells nothing about security. A secure cluster based multi-path routing protocol (SCMRP) [27]. It is a proactive routing protocol since the nodes are considered to be static. During the time of deployment, every node must

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

have a unique ID, a certificate from base station (BS), a shared key with the base station and a public key of the base station.

Neighbor detection phase in which [27] very node broadcast an NBR_DET packet. It has the node ID and the certificate. The receiver node checks the certificate. If it is authenticated the source node is added to the neighbor_list. Otherwise, the packet is dropped. After completing the broadcasting procedure neighbor information is send to the base station. When an intermediate node receives a packet it checks the authenticity of the sender by verifying the certificate. Then if it is authenticated then the packet is rebroadcasted. If a node receives the same packet twice (can be known by checking the node ID)then it decides whether to rebroadcast or drop the packet. Every node maintains a received_packet table. When the base station receives the packet it checks the integrity and authenticity by using message authentication code(MAC) and encrypt the neighbor information with a shared key between the sender and base station. And hence base station will have an idea about the topology of the network based on which it generates a pairwise key.

$$K_{xy}=h(secret,ID_X,Id_y)$$

where the secret key is the random number produced by the base station[27]. Then it unicasts the pairwise key to all the sensor nodes. When a node receives the packet it verifies the certificate of the base station with a public key. Then the sequence number and node pair of the packet in received_packet table are checked. The sequence number and node pair are stored if it is not in the received_packet table. If the destination ID is same as that of its own ID then it encrypts the pairwise key and it is sent to the intended user with a nonce. This is called a challenge message. Neighbor node decrypts the message using shared key and gets the pairwise key. Then the pairwise key is decrypted and sent as challenge reply. If the node doesn't receive challenge reply then it will report to the base station that the node is fake.

For cluster formation, the base station sends initiation packet to the CH to calculate the shortest distance from the base station to respective CH. It can be done by

Next hop ID is verified. If it is same as that of its ID then decrypt the "path" field in the packet to find the next hop from the routing table. Else it is dropped.

Store the packet type and sequence number if it is not present in the received_packet table. Store in the routing table as next hope-previous hop for packet forwarding back to the base station.

Encrypt the path and sequence number using the pairwise key and broadcast this packet which is modified.

After sometime, the base station will stop receiving a packet from the CH. Then base station performs this again. The path is selected in such a way that it should have greater residual energy and lesser hop count.

For selecting members, the first CH sends CH_ADV advertisement packet to the nodes. Nodes check the certificate. Based on the strength of the received packet and the existence of a pairwise key with the advertised ID it sends CH_Join packets to CH. CH forwards the information to the base station. The base station generates a TDMA schedule based on the number of sensor nodes.

Sensor nodes send the encrypted packet to cluster heads. It forwards the aggregated messages to the base station. The base station uses a unique shared key to decrypt and then the threshold value of CH is checked. When CH energy goes below the threshold then another node is chosen to be the CH.

By using ETT-IBS(Identity Based digital Signature) and ETT-IBOOS(Identity Based Online-Offline Signature) security aspects of LEACH can be improved. ETT-IBS is an ID-based cryptography technique. The private key is generated from node ID and master key and the public key is generated from the master key. ETT-IBOOS authentication is given by digital signature. In [28] they use SHA-1 and AES techniques are used.

A detection based redundancy mechanism is also used to check the sinkhole attack in wireless sensor networks [29]. Sink node has more energy. It calculates the shortest distance between two nodes. Nodes are represented as (x_i,y_i). The coordinates of sink node are (0,0). The shortest distance between two nodes is the straight line distance and is calculated by,

$$y = \frac{yb - ya}{xb - xa} * x - xa + ya$$

Effective area is selected in a rectangular form so and all the nodes near the straight line are included in the rectangle. The chosen nodes are selected periodically and saved in [N] array. M shortest paths are calculated by the sink node by Dijkstra's algorithm. If M increases then detection increases. When node A gets the reply packet, it checks the next hop address. If it is the same as node B it is not processed. Otherwise it is send to sink node. It checks whether the suspicious node is malicious or not.

Decision Table

S_{sus}	P_1	P_2	.....	.....	P_3	FLAG
1 2 . . . . . N_{sus}						

N_{sus} is the number of suspicious nodes. Flag represents the number of malicious nodes. After suspicious nodes handshake with the credential nodes, the table is filled with their identity. Now check the number of nodes whose values are same to its x-axis. It is represented as NumOfBi.

$$\frac{\text{NumOfBi}}{M} \geq \text{alpha}$$

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

If the content does not match to its x-value then it is included in S_x .

$$\text{beta} \leqslant \frac{\text{Nx}}{\text{M} * \text{Nsus}}$$

Beta checks whether the elements in suspicious group are malicious or not.
These are some of the approaches to detect selective forwarding attacks in WSNs.

VI. CONCLUSIONS

WSN networks are widely used for many applications. Because of the random deployment of the sensor nodes in hostile areas and also because of the wireless nature of the WSNs they are so vulnerable to attacks. Many advanced routing protocols and methods are introduced in order to secure the network along with conserving energy. But 100% security has not been achieved so far. Anomalies and attacks are the main problems faced by WSNs. Anomalies are abnormalities in the networks. They can be detected by using the statistical methods, cluster-based approaches, neighborhood-based approaches and classification based approaches. Sinkhole attacks take place when a compromised node lures the entire traffic from an area. It can be detected by checking the sequence number or hop count or routing pattern of each sensor node. AODV, an improved routing protocol of DSDV, performs connection establishment by sending route request-route reply packets[3]. So the attacker can easily access the packets and alter the information. Many types of research are going on in the field of secure wireless sensor networks.

REFERENCES

[1] Gandhimathi, L., and G. Murugaboopathi. "Cross layer intrusion detection and prevention of multiple attacks in Wireless Sensor Network using Mobile agent." Information Communication and Embedded Systems (ICICES), 2016 International Conference on. IEEE, 2016.
[2] Kalnoor, Gauri, and Jayashree Agarkhed. "QoS based multipath routing for intrusion detection of sinkhole attack in wireless sensor networks."Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on. IEEE, 2016.
[3] Balla, Rajeshwar L., and Venugopal Kotoju. "Sinkhole Attack detection and prevention in MANET & Improving the performance of AODV Protocol." Compusoft 2.7 (2013): 210.
[4] Wang, Shun-Sheng, et al. "An integrated intrusion detection system for cluster-based wireless sensor networks." Expert Systems with Applications 38.12 (2011): 15234-15243.
[5] Rassam, Murad A., et al. "A sinkhole attack detection scheme in mint-route wireless sensor networks." Telecommunication Technologies (ISTT), 2012 International Symposium on. IEEE, 2012.
[6] Ibrahim, Abdullah, Mohammad Muntasir Rahman, and Mukul Chandra Roy. "Detecting Sinkhole Attacks in Wireless Sensor Network using Hop Count." International Journal of Computer Network and Information Security 7.3 (2015): 50.
[7] Hamedheidari, Sina, and Reza Rafeh. "A novel agent-based approach to detect sinkhole attacks in wireless sensor networks." Computers & Security 37 (2013): 1-14.
[8] Le Fessant, Fabrice, et al. "A sinkhole resilient protocol for wireless sensor networks: Performance and security analysis." Computer communications 35.2 (2012): 234-248.
[9] Ngai, Edith CH, Jiangchuan Liu, and Michael R. Lyu. "On the intruder detection for sinkhole attack in wireless sensor networks." 2006 IEEE International Conference on Communications. Vol. 8. IEEE, 2006.
[10] Papadimitriou, Anthonis, et al. "Cryptographic protocols to fight sinkhole attacks on tree-based routing in wireless sensor networks." Secure Network Protocols, 2009. NPSec 2009. 5th IEEE Workshop on. IEEE, 2009
[11] Cohen, Dylan, et al. "Trustability based on beta distribution detecting abnormal behaviour nodes in WSN."2013 19th Asia-Pacific Conference on Communications (APCC). IEEE, 2013.
[12] Sharmila, S., and G. Umamaheswari. "Detection of sinkhole attack in wireless sensor networks using message digest algorithms." Process Automation, Control and Computing (PACC), 2011 International Conference on. IEEE, 2011.
[13] Zhu, Sencun, Sanjeev Setia, and Sushil Jajodia. "LEAP+: Efficient security mechanisms for large-scale distributed sensor networks." ACM Transactions on Sensor Networks (TOSN) 2.4 (2006): 500-528.
[14] Bhojannawar, Satish S., Chetan M. Bulla, and Vishal M. Danawade. "Anomaly Detection Techniques for Wireless Sensor Networks-A Survey." International Journal of Advanced Research in Computer and Communication Engineering 2.10 (2013).
[15] Zhang, Yang, Nirvana Meratnia, and Paul Havinga."Outlier detection techniques for wireless sensor networks: A survey." IEEE Communications Surveys & Tutorials 12.2 (2010): 159-170.
[16] Abid, Aymen, Abdennaceur Kachouri, and Adel Mahfoudhi. "Anomaly detection through outlier and neighborhood data in Wireless Sensor Networks."Advanced Technologies for Signal and Image Processing (ATSIP), 2016 2nd International Conference on. IEEE, 2016.
[17] Messai, Mohamed-Lamine."Classification of Attacks in Wireless Sensor Networks."arXiv preprint arXiv:1406.4516 (2014).
[18] Chaudhry, Junaid Ahsenali, et al. "Dealing with sinkhole attacks in wireless sensor networks." Advanced Science and Technology Letters 29.2 (2013): 7-12.
[19] Shafiei, Hosein, et al. "Detection and mitigation of sinkhole attacks in wireless sensor networks." Journal of Computer and System Sciences 80.3 (2014): 644-653.
[20] O'Reilly, Colin, etal. "Anomaly detection in wireless sensor networks in a non-stationary environment." IEEE Communications Surveys & Tutorials 16.3 (2014): 1413-1432.
[21] Rassam,MuradA.,Anazida Zainal, and Mohd Aizaini Maarof."Advancements of data anomaly detection research in wireless sensor networks: a survey and open issues." Sensors 13.8 (2013): 10087-10122.
[22] Bosman, Hedde HWJ, et al. "Spatial anomaly detection in sensor networks using neighborhood information." Information Fusion 33 (2017): 41-56.
[23] Chaudhry, Junaid Ahsenali, et al. "Dealing with sinkhole attacks in wireless sensor networks." Advanced Science and Technology Letters 29.2 (2013): 7-12.
[24] Boubiche, Djallel Eddine, and Azeddine Bilami. "Cross layer intrusion detection system for wireless sensor network." International Journal of Network Security & Its Applications 4.2 (2012): 35.
[25] Boubiche, Djallel Eddine, and Azeddine Bilami. "Cross layer intrusion detection system for wireless sensor network." International Journal of Network Security & Its Applications 4.2 (2012): 35.
[26] Marzi, Hosein, and Mengdu Li. "An enhanced bio-inspired trust and reputation model for wireless sensor network." Procedia Computer Science 19 (2013): 1159-1166.
[27] Kumar, Suraj, and Sanjay Jena. "SCMRP: Secure cluster based multipath routing protocol for wireless sensor networks." Wireless Communication and Sensor Networks (WCSN), 2010 Sixth International Conference on. IEEE, 2010.
[28] R. Yasmin, E. Ritter, and G. Wang, "An Authentication Framework for Wireless Sensor Networks using Identity-Based Signatures," in Proc. IEEE CIT, 2010.
[29] Zhang, Fang-Jiao, et al. "Sinkhole attack detection based on redundancy mechanism in wireless sensor networks." Procedia Computer Science 31 (2014): 711-720.
[30] Le Fessant, Fabrice, et al. "A sinkhole resilient protocol for wireless sensor networks: Performance and security analysis." Computer communications 35.2 (2012): 234-248.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 1, January 2017

[31] Yang, Hongyu, Feng Xie, and Yi Lu. "Clustering and classification based anomaly detection." International Conference on Fuzzy Systems and Knowledge Discovery. Springer Berlin Heidelberg, 2006.
[32] Fawzy, Asmaa, Hoda MO Mokhtar, and Osman Hegazy. "Outliers detection and classification in wireless sensor networks." Egyptian Informatics Journal 14.2 (2013): 157-164.
[33] Khurana, Mehak, and Ashish Payal. "An improvement of Centroid Algorithm based on distance in Wireless Sensor Network." International Journal of Smart Sensors and Ad-Hoc Networks 1.1 (2011).
[34] Qiu, Wanzhi, Efstratios Skafidas, and Peng Hao. "Enhanced tree routing for wireless sensor networks." Ad hoc networks 7.3 (2009): 638-650.
[35] Momani, Mohammad, and Subhash Challa. "Survey of trust models in different network domains." arXiv preprint arXiv:1010.0168 (2010).

BIOGRAPHIES

Arya I S is a M.Tech scholar in Communication Engineering, Kerala Technical University. She received B. Tech in Electronics and Communication Engineering degree in the year 2015 from the University of Kerala, India. Her research interest is in Wireless Sensor Networks.

Dr. Binu G S is an Associate Professor, Dept of ECE, NSS College of Engineering, Palakkad. She has many papers in National and International journals to her credit. She has more than 15 years of academics and research experiences. Her research area of interest is in Wireless Sensor Networks.