MULTI ZONE BASE SPOOFING ATTACK DETECTION AND LOCALIZATION FOR WIRELESS SENSOR NETWORK USING NOVEL HYPOTHESIS TESTING

1.T.Priya M.Sc., Research scholar, Department of computer Science, Kongu Arts and Science College,Erode.

2. C.P.Balasubramaniam M.Sc.,M.Phil.,Assistant Professor,Department of computer Science, Kongu Arts and Science College, Erode.

Abstract: Wireless spoofing attack is a combine physical and application level attacker, spoofing attack model first affected node compromised one node MAC address and group of node attack in wireless sensor network. So ireless spoofing attacks are easy to launch and can extensively impact the performance of network. In the new methodology there are three phases of work. First one is, to implement to receive signal strength (RSS) and property associated with the transmission and reception of communication for detecting spoofing. Second one is, group of node request between server to inform all kinds of neighbors details send and update database. In this method simulation applied for K-Means clustering algorithm. Third one is, to implement the new Chronological Probability Fraction Testing applied for multiple zones based wireless sensor networks. In this paper the new testing, statistical framework based on Chronological Probability Fraction Test (CPFT) hypothesis testing for modeling, analyzing, and evaluating statistical source anonymity in wireless sensor networks has been analyzed. The proposed system introduces the notion of number of multiple zone in distinguish ability to model source location privacy. It shows that the current approaches for designing statistically anonymous systems introduce correlated in multiple zones. In addition, a fast and effective mobile replica node detection scheme is proposed using the chronological probability ratio test. The Location results are representing using Enhanced CPFT algorithm that provides strong evidence of high accuracy of localizing multiple adversaries.

Key Words: Spoofing attack, RSS,CPFT, Replica Detection, K-Means clustering, localization.

I.INTRODUCTION

A Wireless Sensor Network(WSN) of spatially distributed autonomous sensor to monitor physical or environmental conditions, such as temperature, sound, pressure, etc. and pass their data through the network to main location. Wireless Sensor Network (WSN) is an emerging technology that shows great promise for various futuristic applications both for mass public and military. The sensing technology combined with processing power and wireless communication makes it lucrative for being exploited in abundance in future. Wireless sensor networks are characterized by severely constrained computational and energy resources, and an ad hoc operational environment. Wireless sensor networks (WSN) are currently receiving significant attention due to their unlimited potential. However, it is still very early in the lifetime of such systems and many research challenges exist.

WIRELESS SENSOR NETWORK ATTACK

Wireless Sensor Networks (WSN) are emerging as both an important new tier in the IT ecosystem and a rich domain of active research

 All Rights Reserved © 2014 IJARCET

involving hardware and system design, networking, distributed Algorithms, programming models, data management, security and social factors. Wireless sensor network applications include ocean and wildlife monitoring, manufacturing machinery performance monitoring, building safety and earthquake monitoring, and many military applications.

An even wider spectrum of future applications is likely to follow, including the monitoring of highway traffic, pollution, wildfires, building security, water quality, and even people's heart rates. A major benefit of these systems is that they perform in-network processing to reduce large streams of raw data into useful aggregated information. Because sensor networks pose unique challenges, traditional security techniques used in traditional Networks cannot be applied directly.

First, to make sensor networks economically viable, sensor devices are limited in their energy, computation, and communication capabilities. Second, unlike traditional networks, sensor nodes are often deployed in accessible areas, presenting the added risk of physical attack. And third, sensor networks interact closely with their physical environments and with people, posing new security problems. Consequently, existing security mechanisms are inadequate, and new ideas are needed.

II. RELATED WORK

The system uses a minimum number of trusted nodes it is not so applicable to sensor networks where the nodes are randomly spread out. In other words, it is possible that under certain conditions nodes cannot find the minimum number of neighboring nodes in order to be named trusted. One solution for locationized anomaly detection in a group of nodes is suggested in [7]. Every node gets the localization information from the neighboring nodes and also computes the localization information itself and compares these two values. If the difference is small enough, that node decides there is no adversary around causing the localization problem in its location. In mobile ad hoc networks, the computational load and complexity for key management is strongly subject to restriction of the node's available resources and the dynamic nature of network topology. Author [8] proposed a secure and efficient key management framework (SEKM) for mobile ad hoc networks. SEKM builds PKI by applying a secret sharing scheme and an underlying multicast server group. In SEKM, the server group creates a view of the certification authority (CA) and provides certificate update service for all nodes, including the servers themselves. A ticket scheme is introduced for efficient certificate service. In this paper [9], Wireless networks are vulnerable to spoofing attacks, which allows for many other forms of attacks on the networks. Although the identity of a node can be verified through cryptographic authentication, authentication is not always possible because it requires key management and additional infrastructural overhead. In this paper, to propose a method for both detecting spoofing attacks, as well as locating the positions of adversaries performing the attacks. They first proposed an attack detector for wireless spoofing that utilizes K-means cluster analysis. Next, it describes how they integrated the attack detector into a real-time indoor localization system, which is also capable of localizing the positions of the attackers.

 All Rights Reserved © 2014 IJARCET

SPOOFING ATTACK

Spoofing is when an attacker pretends to be someone else in order gain access to restricted resources or steal information. This type of attack can take a variety of different forms; for instance, an attacker can impersonate the Internet Protocol (IP) address of a legitimate user in order to get into their accounts. Also, an attacker may send fraudulent emails and set up fake websites in order to capture users' login names, passwords, and account information.

GADE: Generalized Attack Detection Model (GADE) that can both detect spoofing attacks as well as determine the number of adversaries using cluster analysis methods grounded on RSS-based spatial correlations among normal devices and adversaries is used in existing system.In GADE, the Partitioning Around Medoids (PAM) cluster analysis method is used to perform attack detection. It formulates the problem of determining the number of attackers as a multiclass detection problem. It then applied cluster-based methods to determine the number of attackers.

To study Received Signal Strength (RSS), a property closely correlated with location in physical space and is readily available in the existing wireless networks. The RSS readings at the same physical location are similar, whereas the RSS readings at different locations in physical space are distinctive. Thus, the RSS readings present strong spatial correlation characteristics.

The Received Signal Strength value vector as s = (s1, s2,... sn) where n is the number of landmarks/access points that are monitoring the RSS of the wireless nodes and know their locations. Generally, the RSS at the ith landmark from a wireless node is distributed as

$$s_i(d_j)[dBm] = P(d_0)[dBm] - 10\gamma \log\left(\frac{d_j}{d_0}\right) + X_i,$$

where P(d0) represents the transmitting power of the node at the reference distance d0, dj is the distance between the wireless node is the path loss exponent, Xi is the shadow j and the ith landmark, and fading which is given as input. For simplicity, we assume the wireless nodes have the same transmission power.

If the received signal strength does not match in successive RSS values, then the node is said to be malicious.

THEORETICAL ANALYSIS OF THE SPATIAL CORRELATION OF RSS

The challenge in spoofing detection is to devise strategies that use the uniqueness of spatial information, but not using location directly as the attackers' positions are unknown. The proposed study RSS, a property closely correlated with location in physical space and is readily available in the existing wireless networks. Although affected by random noise, environmental bias, and multipath effects, the RSS measured at a set of landmarks (i.e., reference points with known locations) is closely related to the transmitter's physical location and is governed by the distance to the landmarks. The RSS readings at the same physical location are similar, whereas the RSS readings at different locations in physical space are distinctive. Thus, the RSS readings present strong spatial correlation characteristics.

 All Rights Reserved © 2014 IJARCET

ATTACK DETECTION USING CLUSTER ANALYSIS

The non-hierarchical method initially takes the number of components of the population equal to the final required number of clusters. First, the final required number of clusters is chosen such that the points are mutually farthest apart. Next, it examines each component in the population and assigns it to one of the clusters depending on the minimum distance.

The centroid's position is recalculated every time a component is added to the cluster and this continues until all the components are grouped into the final required number of clusters. K-means (MacQueen, 1967) is one of the simplest unsupervised learning algorithms that solve the well known clustering problem.

The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because of different location causes different result.

So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age is done.

At this point they need to re-calculate k new centroids as barycenters of the clusters resulting from the previous step. After they have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop they may notice that the k centroids change their location step by step until no more changes are done. In other words centroids do not move any more.

The k-means approach to clustering performs an iterative alternating fitting process to form the number of specified clusters. The k-means method first selects a set of n points called cluster seeds as a first guess of the means of the clusters. Each observation is assigned to the nearest seed to form a set of temporary clusters. The seeds are then replaced by the cluster means, the points are reassigned, and the process continues until no further changes occur in the clusters.

The K-means Algorithm

- The dataset is partitioned into K clusters and the data points are randomly assigned to the clusters resulting in clusters that have roughly the same number of data points.
- For each data point
- Calculate the distance from the data point to each cluster.
- If the data point is closest to its own cluster, leave it where it is. If the data point is not closest to its own cluster, move it into the closest cluster.
- Repeat the above step until a complete pass through all the data points results in no data point moving from one cluster to another. At this point the clusters are stable and the clustering process ends.

 All Rights Reserved © 2014 IJARCET

- The choice of initial partition can greatly affect the final clusters that result, in terms of inter-cluster and intracluster distances and cohesion.

III. METHODOLOGY

The proposed system work is motivated from mitigating the limitations of previous schemes. In particular, the new system proposes a method in which the nodes are fixed as well as in movement. A reputation-based trust management method is designed to facilitate fast detection of compromised nodes. The key idea of the method is to detect untrustworthy zones and perform software attestation against nodes in these zones to detect and revoke the ones that are compromised.

Specifically, it first divides the network into a set of zones, establish trust levels for each zone, and detect untrustworthy zones by using the Chronological Probability Fraction Test (CPFT). The CPFT decides a zone to be untrustworthy if the zone's trust is continuously maintained at low level or is quite often changed from high level to low level.

Once a zone is determined to be untrustworthy, the base station or the network operator performs software attestation against all nodes in the untrustworthy zone, detects compromised nodes with subverted software modules, and physically revokes them.

In addition, a novel mobile replica detection scheme is proposed based on the Chronological Probability Fraction Test (CPFT). The new system uses the fact that an uncompromised mobile node should never move at speeds in excess of the system-configured maximum speed. As a result, a caring mobile sensor node's measured speed will nearly always be less than the system-configured maximum speed as long as it employs a speed measurement system with a low error rate.

On the other hand, replica nodes are in two or more places at the same time. This makes it appear as if the replicated node is moving much faster than any of the caring nodes, and thus the replica nodes' measured speeds will often be over the system-configured maximum speed.

The scheme is based on the Chronological Probability Fraction Test which is a statistical decision process. The CPFT can be thought of as one- dimensional random walk with the lower and upper limit.

Before the random walk starts, null and alternate hypotheses are defined in such away that the null hypothesis is associated with the lower limit while the alternate one is associated with the upper limit. A random walk starts from a point between two limits and moves toward the lower or upper limit in accordance with each observation. If the walk reaches (or exceeds) the lower or upper limit, it terminates and the null or alternate hypothesis is selected, respectively.

It is believed that the CPFT is well suited for tackling the mobile replica detection problem since we can construct a random walk with two limits in such a way that each walk is determined by the observed speed of a mobile node. The lower and upper limits can be configured to be associated with speed less than and in excess of Vmax, respectively.

 All Rights Reserved © 2014 IJARCET

Enhanced CPFT Algorithm:

DECLARATION: $n1$=0,w_n=0

INPUT: location information *P1* and time information *C*

OUTPUT: accept the hypothesis H_0 or H_1

node_curr_loc= *P1*

node_curr_time=*C*

if *n1*>0 then

compute $C_0(n1)$ and $C_1(n1)$

compute speed 0 from node_curr_loc and node_prev_loc, node_curr_time and node_prev_time

if 0>V_{max} then

$w_n=w_n+1$

end if

if $w_n>=C_1(n1)$ then

Accepts the hypothesis h_1 and terminate the test

end if

if $w_n<=C_0(n1)$ then

initialize $n1$ and w_n to 0 and accepts the hypothesis H_0

return;

end if

end if

$n1=n1+1$

node_prev_loc=node_curr_loc

node_prev_time=node_curr_time

end

The CPFT is applied to the mobile replica detection problem as follows: Each time a mobile sensor node moves to a new 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.

The base station computes the speed from every two consecutive claims of a mobile node and performs the CPFT by considering speed as an observed sample. Each time the mobile node's speed exceeds (respectively, remains below) Vmax, it will expedite the random walk to hit or cross the upper (respectively, lower) limit and thus lead to the base station accepting the alternate (respectively, null) hypothesis that the mobile node has been (respectively, not been) replicated. Once the base station decides that a mobile node has been replicated, it revokes the replica nodes from the network.

IV. RESULTS

The following **Table 1** describes experimental result for existing system error rate analysis. The table contains zone id, time interval (per sec), node id, affect number of node id details and error rate percentage details are shown

Table 6.2 Performances Result for Existing System

Zone ID	Times (sec)	Node ID	Total No. of. Attacker Node	Error Rate (%)
1	**60**	N0	**45**	**75.00**
1	**120**	N4	**43**	**71.66**
1	**180**	N18	**26**	**43.33**
1	**240**	N1	**22**	**36.66**
1	**300**	N7	**20**	**33.33**
1	**360**	N19	**18**	**30.00**
1	**420**	N15	**11**	**18.33**
1	**480**	N9	**5**	**8.33**
1	**520**	N20	**4**	**6.66**
1	**580**	N11	**3**	**5.00**

 All Rights Reserved © 2014 IJARCET

The following **Table 6.3** describes experimental result for proposed system error rate analysis. The table contains zone id, time interval (per sec), node id, affect number of node id details and error rate percentage details are shown

Table 6.3 Performances Result for Proposed System

Number of Zone ID	Times (sec)	Node ID	Total No. of. Attacker Node	Error Rate (%)
1	**60**	N0	**52**	**86.66**
1	**120**	N4	**51**	**85.00**
2	**180**	N18	**32**	**53.33**
1	**240**	N1	**28**	**46.66**
1	**300**	N7	**25**	**41.66**
2	**360**	N19	**24**	**40.00**
5	**420**	N15	**13**	**21.66**
5	**480**	N9	**8**	**13.33**
5	**520**	N20	**6**	**10.00**
1	**580**	N11	**5**	**8.33**

V. CONCLUSION AND FUTURE ENHANCEMENT

This proposed work to use received signal strength-based spatial correlation, a physical property associated with each wireless device that is hard to falsify and not reliant on cryptography as the basis for detecting spoofing attacks in wireless networks. It provided theoretical analysis of using the spatial correlation of RSS inherited from wireless nodes for attack detection. It derived the test statistic based on the cluster analysis of RSS readings. The approach can both detects the presence of attacks as well as determine the number of adversaries, spoofing the same node identity, so that we can localize any number of attackers and eliminate them. In addition, a zone-based node compromise detection scheme is proposed used and furthermore, several possible attacks are described against the proposed scheme and proposed counter-measures against these attacks. The scheme is evaluated in simulation under various scenarios. The experimental results show that the scheme quickly detects untrustworthy zones with a small number of zone-trust reports. This approach identify malicious nodes through trust management schemes and thereby revoke the compromise node and find out the most affected i.e., compromised zones.

In future, the scheme may evaluate against various types of attacker models. It is believed that a game theoretic model is suited for this evaluation. A variety of strategies novel hypothesis testing applying for multiple zone base network to finding number of spoofing attacker detected and that may be taken by detector and adversary.

VI. REFERENCES

1. J.Bellardo and S. Savage, “802.11 Denial-of-Service Attacks. Real Vulnerabilities and Practical Solutions”. In Proceedings of the 12th USENIX Security Symposium, Washington, D.C., August 4-8, 2003.
2. F. Ferreri, M. Bernaschi, and L. Valcamonici, “Access Points Vulnerabilities to Dos Attacks in 802.11 Networks,” Proc. IEEE Wireless Comm. and Networking Conf., 2004.

 All Rights Reserved © 2014 IJARCET

3. D. Faria and D. Cheriton, "Detecting Identity-Based Attacks in Wireless Networks Using Signalprints," Proc. ACM Workshop Wireless Security (WiSe), Sept. 2006.
4. W. Du, L. Fang, and P. Ning, "LAD: Localization Anomaly Detection for Wireless Sensor Networks," the 19th International Parallel and Distributed Priocessing Symposium (IPDPS'05), April 3 – 8, 2005, Denver, Colorado, USA.
5. Q. Li and W. Trappe, "Relationship-Based Detection of Spoofing-Related Anomalous Traffic in Ad Hoc Networks," Proc. Ann. IEEE Comm. Soc. on IEEE and Sensor and Ad Hoc Comm. and Networks (SECON), 2006.
6. Y. Chen, K. Kleisouris, X. Li, W. Trappe, and R.P. Martin, "The Robustness of Localization Algorithms to Signal Strength Attacks: A Comparative Study," Proc. Int'l Conf. Distributed Computing in Sensor Systems (DCOSS), pp. 546-563, June 2006.
7. Sheng.Y, KTan.K, Chen.G, Kotz.D, and Campbell.A, "Detecting 802.11 MAC layer spoofing using received signal strength," in Proc. IEEE INFOCOM, April 2008.
8. Wu. B, Wu.J, E. Fernandez, and S. Magliveras, "Secure and Efficient Key Management in Mobile Ad Hoc Networks," Proc. IEEE Int'l Parallel and Distributed Processing Sym. (IPDPS), 2005.
9. Jie Yang, Yingying Chen, and Jerry Cheng, "Detection and Localization of Multiple Spoofing Attackers in Wireless Networks" in IEEE 2012.

 All Rights Reserved © 2014 IJARCET