Open access Journal International Journal of Emerging Trends in Science and Technology

Detection and Localization of Versatile spoofing Attackers in WSN

Authors

Kiran Kumar P N[1], Venkatesh D[2], T. Ramamohan[3]

[1]PG Student, Department of Computer Science & Engineering, Gates Institute of Technology Gooty, Anantapur.
Email: *pnkumar.cs036@gmail.com, kirancs.kss@gmail.com*

[2]Prof. & Dean, Department of Computer Science & Engineering, Gates Institute of Technology Gooty, Anantapur.
Email: *deancseit@gmail.com*

[3]Asst. Prof, Department of Computer Science & Engineering, Gates Institute of Technology Gooty, Anantapur
Email: *Ramamohan.tapasi@gmail.com*

Abstract

Wireless spoofing strikes are easy to launch and can dramatically significance the efficiency of networks. Although the recognition of a node might possibly be verified by means of cryptographic authentication, typical security approaches are not always desirable because of their extra specifications. In this paper ,We are suggest to use spatial knowledge, a physical character related with each node, hard to falsify, except for reliant on cryptography, considering the reason for one detecting spoofing attacks; Two discovering the number of attackers when multiple competitors masquerading as the similar node identity; and Three localizing multiple competitors. We are suggesting to use the spatial association of received signal strength (RSS) acquired from cord less nodes to discover the spoofing attacks. We then build up the trouble of discovering the number of attackers in form of a multiclass detection problem. Cluster-based strategies are designed to determine the number of attackers. As soon as the training facts are located, we examine using the Support Vector Machines (SVM) process to further improve the accuracy of discovering the number of attackers. In addition, we have designed an integrated recognition and localization strategy that can localize the positions of various attackers. We have ranked the strategies through two test beds using both a WiFi and ZigBee networks in two real workplaces. Our experimental results show that our proposed techniques can achieve over 90 percent Hit Rate and Accuracy while working out the array of attackers. Localization outputs implementing a standard couple of algorithms provide effective confirmation of high accuracy of localizing multiple competitors.

Keywords – *wireless spoofing attacks, localization, and cluster based strategies*

INTRODUCTION

As more wireless and sensor networks are deployed, they will increasingly become tempting targets for malicious attacks. Due to the openness of wireless and sensor networks, they are especially vulnerable to spoofing attacks where an attacker forges its identity to masquerade as another device, or even creates multiple illegitimate identities. Spoofing attacks are a serious threat as they represent a form of identity compromise and can facilitate a variety of traffic injection attacks, such as evil twin access point attacks found in [1]. It is thus desirable to detect the presence of spoofing and eliminate them from the network. The traditional approach to address

spoofing attacks is to apply cryptographic authentication. However, authentication requires additional infrastructural overhead and computational power associated with distributing, and maintaining cryptographic keys. Due to the limited power and resources available to the wireless devices and sensor nodes, it is not always possible to deploy authentication. In addition, key management often incurs significant human management costs on the network. In this paper [2], we take a different approach by using the physical properties associated with wireless transmissions to detect spoofing. Specifically, we propose a scheme for both detecting spoofing attacks, as well as localizing the positions of the adversaries performing the attacks. Our approach utilizes the Received Signal Strength to detecting spoofing attacks, determining the number of attackers when multiple adversaries masquerading as the same node identity and localizing multiple adversaries. The transmitted information from server is send to client in secure manner. If an intruder comes during transaction server discover and localize that specific system. Improves efficiency of the application, improves accuracy of detecting number of spoofing attackers and also improves efficiency of identifying the location of the attackers

Signal Strength (RSS) measured across a set of access points to perform spoofing detection and localization [6]. Our scheme does not add any overhead to the wireless devices and sensor nodes. By analyzing the RSS from each MAC address using K-means cluster algorithm, we have found in [4] that the distance between the centroids in signal space is a good test statistic for effective attack detection. We then describe how we integrated our K-means spoofing detector into real-time indoor localization system. Our K-means approach is general in that it can be applied to almost all RSS-based localization algorithms [3]. For two sample algorithms, we show that using the centroids of the clusters in signal space as the input to the localization system, the positions of the attackers can be localized with the same relative estimation errors as under normal conditions. In particular, The main contributions of our work are: 1) GADE: a 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; and 2) IDOL: an integrated detection and localization system that can both detect attacks as well as find the positions of multiple adversaries even when the adversaries vary their transmission power levels. In GADE, the Partitioning around Medoids (PAM) cluster analysis method is used to perform attack detection.

2. EXISTING SYSTEM

The recognition of a node may be certified by means of popular security approaches are not always desirable. Adversaries can easily purchase low-cost devices and use these commonly available platforms to launch a variety of attacks. Among various types of attacks, identity-based spoofing attacks are especially easy to launch and can cause significant damage to network performance. It is easy for an attacker to gather useful MAC address information during passive monitoring and then modify its MAC address. It can further facilitate a variety of traffic injection attacks, such as attacks on access control lists, rogue access point (AP) attacks, and eventually denial of service (DOS) attacks. The traditional approach to address spoofing attacks is to apply cryptographic authentication. However, authentication requires additional infrastructural overhead and computational power associated with distributing, and maintaining cryptographic keys.

Due to the limited power and resources available to the wireless devices and sensor nodes, it is not always possible to deploy authentication. In addition, key management often incurs significant human management costs on the network.

The existing system uses the cryptographic-based authentication, for example a secure and efficient key Management (SEKM) framework. SEKM builds a Public Key Infrastructure which uses periodic key refresh and host revocation to prevent the compromise of authentication keys the cryptographic authentication may not be always applicable because of the limited resources on wireless devices and lacking of a fixed key management infrastructure in the wireless network.

Disadvantages of Existing System

1. It requires reliable key distribution, management, and maintenance mechanisms. It is not always desirable to apply these cryptographic methods because of its infrastructural, computational, and management overhead.
2. It has no ability to determine the number of attackers when multiple adversaries use the same identity to launch attacks, which is the basis to further localize multiple adversaries after attack detection.
3. It can only handle the case of a single spoofing attacker and cannot localize the attacker if the adversary uses different transmission power levels.

3. PROPOSED SYSTEM

Formulate the problem of determining the number of attackers as a multiclass detection. Preside over a secure and efficient key management framework that builds a public key infrastructure. Cluster-based mechanisms are developed to determine the number of attackers. Explore using the support vector machines method to further improve the accuracy of determining the number of attackers. Determining the number of attackers when there are multiple adversaries collaborating to use the same identity to launch malicious attacks. This approach can accurately localize multiple adversaries. In this paper, we take a different approach by using the physical properties associated with wireless transmissions to detect spoofing. Specifically, we propose a scheme for both detecting spoofing attacks, as well as localizing the positions of the adversaries performing the attacks. Our approach utilizes the Received Signal Strength (RSS) measured across a set of access points to perform spoofing detection and localization [2]. Our scheme does not add any overhead to the wireless devices and sensor nodes. Under the spoofing attack, the victim and the attacker are using the same ID to transmit data packets, and the RSS readings of that ID is the mixture readings measured from each individual node (i.e., spoofing node or victim node). Since under a spoofing attack, the data packets from the victim node and the spoofing attackers are mixed together, this observation suggests conducting cluster analysis in order to detect the presence of spoofing attackers in wireless network. A 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. An integrated detection and localization (IDOL) system that can detect both attacks as well as find the positions of multiple adversaries even when the adversaries vary their transmission power levels.

Fig 1 shows the proposed system architecture. In GADE, the Partitioning around Medoids (PAM) cluster analysis method is used to perform attack detection. The problem of determining the number of attackers is formulated as a multiclass detection problem. Then cluster-based methods are applied to determine the number of attacker. Spoofing attack detection is performed using Cluster Analysis. As the wireless network is deployed as clusters, the attackers are identified in each and every cluster separately. Finally, an integrated system, IDOL, is utilizes the results of the number of attackers returned by GADE to further localize multiple adversaries.

A 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. An integrated detection and localization (IDOL) system that can detect both attacks as well as find the positions of multiple adversaries even when the adversaries vary their transmission power levels.

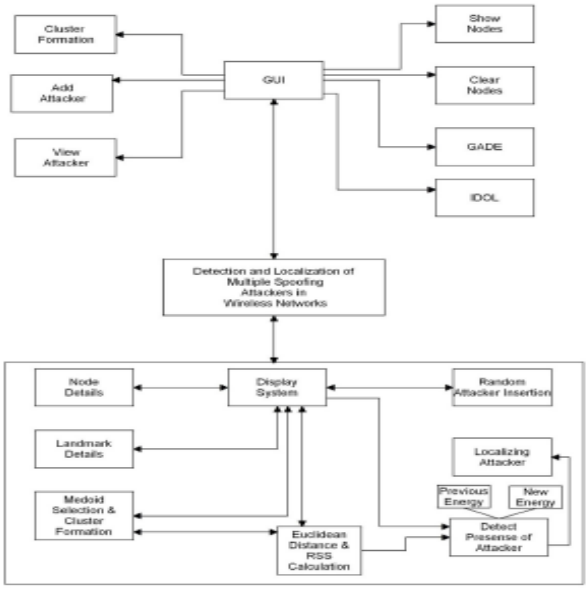


Fig 1: Proposed System Architecture

EXPERIMENTAL RESULTS

1. Generalized attack detection (GADE)

In order to evaluate the effectiveness of our spoofing detection mechanisms, we have conducted experiments using both an 802.11 (WiFi) network, using an Orinoco silver card, as well as an 802.15.4 (ZigBee) network, using a Telosb mote, on the 3rd floor of the Computer Science Department at Rutgers University. The floor size is 200x80ft (16000 ft2). Figure 2 (a) shows the 802.11 (WiFi) network with 4 landmarks deployed to maximize signal strength coverage, as shown in red squares.

The 802.15.4 (ZigBee) network is presented in Figure 2 (b) with 4 landmarks distributed in a squared setup in order to achieve optimal landmark placement as shown in red triangles. The small blue dots in the floor map are the locations used for spoofing and localization tests. For the 802.15.4 network, we used 300 packet-level RSS samples for each of the 100 locations. We utilized the actual RSS values attached to each packet. We have 286 locations in the 802.11 deployment. Unlike the 8 2.15.4 data, the RSS values are partially synthetic. We had access to only the mean RSS at each location, but to perform our experiments we needed an RSS value per packet.

To generate such data for 200 simulated packets at each location, we used random draws from a normal distribution. We used the measured RSS mean for the mean of the distribution. For the standard deviation, we computed the difference in the RSS from a fitted signal to distance function, and then calculated the standard deviation of the distribution from these differences over all locations. To keep our results conservative, we took the maximum deviation over all landmarks, which we found to be 5 dB. Much work has gone into characterizing the distributions of RSS readings indoors. It has been shown that characterizing the per-location RSS distributions as normal, although not often the most accurate characterization, still results in the best balance between algorithmic usability and the resulting localization error. In addition, we built a real-time localization system to estimate the positions of both the original nodes and the spoofing nodes.

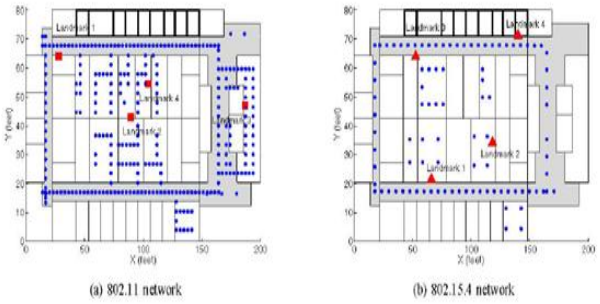


Fig 2: Landmark setups and testing locations in two networks

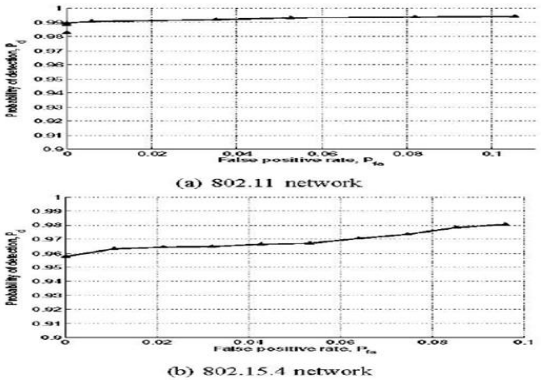


Fig 3: Receiver operating characteristics (ROC) Curves

When Pspoof (spoofing node) is about 13 feet away from Porg (original node) under T equals to 5.5dB. While for the 802.15.4 network, the detection rate is above 90% when the distance between Pspoof and Porg is about 20 feet by setting threshold T to 9dB. This is in line with the average localization estimation errors using RSS which are about 10-15 feet.

When the nodes are less than 10-15 feet apart, they have a high likelihood of generating similar RSS readings, and thus the spoofing detection rate falls below 90%, but still greater than 60%. However, when Pspoof moves closer to Porg, the attacker also increases the probability to expose itself. The detection rate goes to 100% when the spoofing node is about 45-50 feet away from the original node.

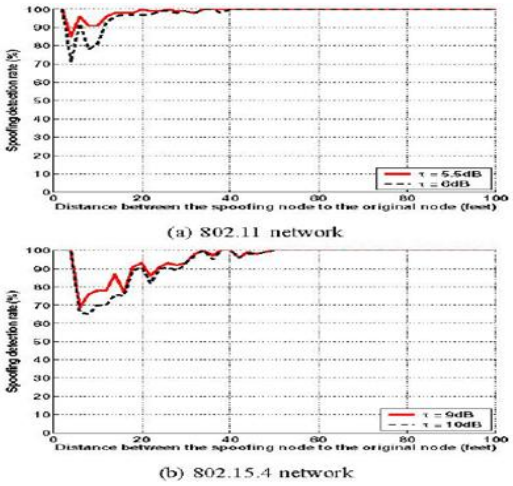


Fig 4: Detection rate as a function of the distance Between a spoofing node & original node

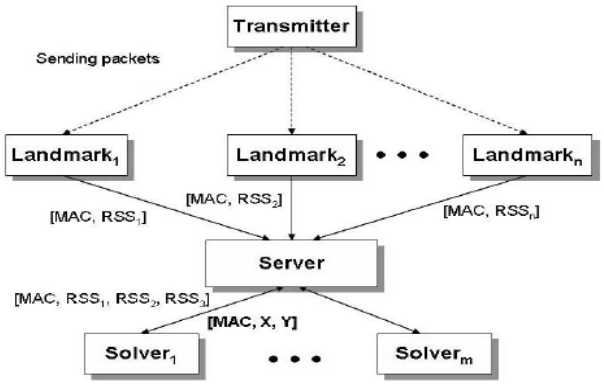


Fig 5: Localization system architecture

IDOL: INTEGRATED DATECTION AND LOCALIZATION

We have developed a general-purpose localization system to real-time indoor positioning. This system is designed with fully distributed functionality and easy to plug-in localization algorithms. It is built around 4 logical components: Transmitter, Landmark, Server, and Solver. The system architecture is shown in Figure 5. Transmitter: Any device that transmits packets can be localized. Often the application code does not need to be altered on a sensor node in order to localize it. Landmark: The Landmark component listens to the packet traffic and extracts the RSS reading for each transmitter.

It then forwards the RSS information to the Server component. The Landmark component is stateless and is usually deployed on each landmark or access point with known locations.

Server: A centralized server collects RSS information from all the Landmark components. The spoofing detection is performed at the Server component. The Server summarizes the RSS information such as averaging or clustering, then forwards the information to the Solver component for localization estimation. Solver: A Solver takes the input from the Server,

Performs the localization task by utilizing the localization algorithms plugged in, and returns the localization results back to the Server. There are multiple Solver instances available and each Solver can localize multiple transmitters simultaneously.

During the localization process, the following steps will take place:

1. A Transmitter sends a packet. Some numbers of Landmarks observe the packet and record the RSS.
2. Each Landmark forwards the observed RSS from the transmitter to the Server.
3. The Server collects the complete RSS vector for the transmitter and sends the information to a Solver instance for location estimation.
4. The Solver instance performs localization and returns the coordinates of the transmitter back to the Server.

If there is a need to localize hundreds of transmitters at the same time, the server can perform load balancing among the different solver instances. This centralized localization solution also makes enforcing contracts and privacy policies more tractable.

When our spoofing detector has identified an attack for a MAC address, the centroids returned by the K-means clustering analysis in signal space can be used by the server and sent to the solver for location estimation. The returned positions should be the location estimate for the original node and the spoofing nodes in physical space. Using a location on the testing floor as an example, Figure 6 shows the relationship among the original node Porg, the location estimation of the original node Lorg, the spoofing node Pspoof, and the localized spoofing node position L_{spoof}.

RADAR: Point-based methods return an estimated point as a localization result. A primary example of a point-based method is the RADAR scheme [5]. In RADAR, during the off line phase, a mobile transmitter with known position broadcasts beacons periodically, and the RSS readings are measured at a set of landmarks. Collecting together the averaged RSS readings from each of the landmarks for a set of known locations provides a radio map. At runtime, localization is performed by measuring a transmitter's RSS at each landmark, and the vector of RSS values is compared to the radio map.

The record in the radio map whose signal strength vector is closest in the Euclidean sense to the observed RSS vector is declared to correspond to the location of the transmitter. In this work, instead of using the averaged RSS in the traditional approach, we use the RSS centroids obtained from the K-means clustering algorithm as the observed RSS vector for localizing a MAC address.

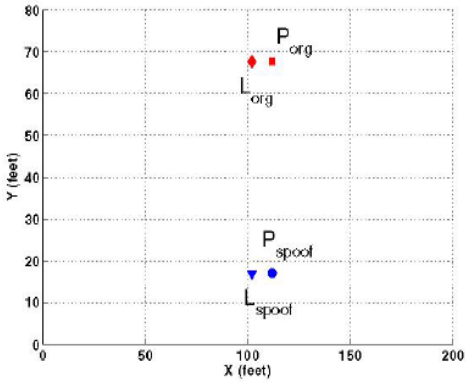


Fig 6: shows the relationship among the original node Porg, the location estimation of the original node Lorg, the spoofing node Pspoof, and the localized spoofing node position L,POOf

Area Based Probability (ABP): Area-based algorithms return a most likely area in which the true location resides. One major advantage of area-based methods compared to point-based methods is that they return a region, which has an increased chance of capturing the transmitter's true location. ABP returns an area, a set of tiles on the floor, bounded by a probability that the transmitter is within the returned area . ABP assumes the distribution of RSS for each landmark follows a Gaussian distribution. The Gaussian random variable from each landmark is

independent. ABP then computes the probability of the transmitter being at each tile L on the floor using Bayes' rule:

$$P(L_i|\mathbf{s}) = \frac{P(\mathbf{s}|L_i) \times P(L_i)}{P(\mathbf{s})}.$$

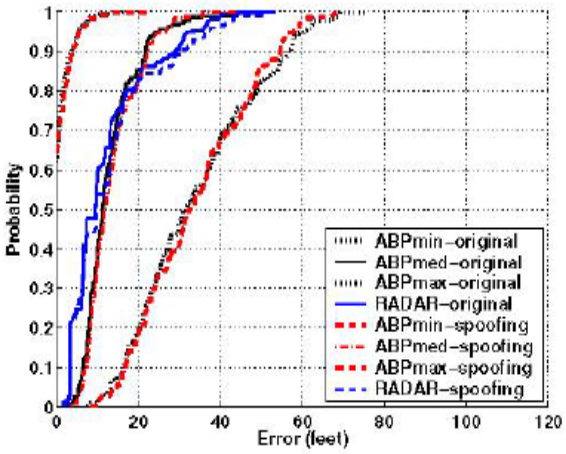
Given that the transmitter must reside at exactly one tile satisfying z 1 P(L| s) = 1, ABP normalizes the probability and returns the most likely tiles up to its confidence . Both RADAR and ABP are employed in our experiments to localize the positions of the attackers.

In order to evaluate the effectiveness of our localization system in finding the locations of the attackers, we are interested in the following performance metrics: Localization Error CD: We obtain the cumulative distribution function (CDF) of the location estimation error from all the localization attempts, including both the original nodes and the spoofing nodes.

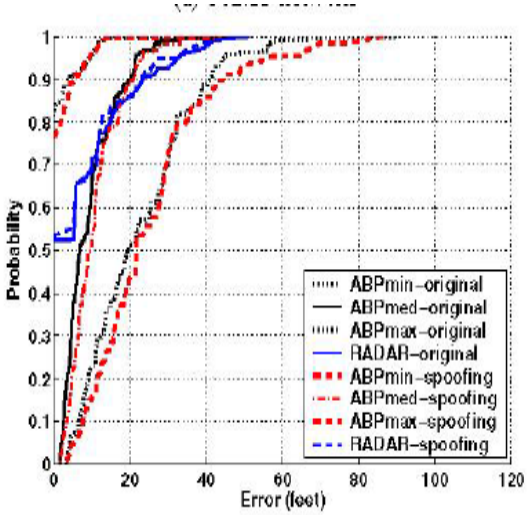
We then compare the error CDF of all the original nodes to that of all the possible spoofing nodes on the floor. For area based algorithms, we also report CDFs of the minimum and maximum error. For a given localization attempt, these are points in the returned area that are closest to and furthest from the true location. Relationship between the true and estimated distances: The relationship between the true distance of the spoofing node to the original node Porg -Pspoof and the distance of the location estimate of the spoofing node to that of the original node ILorg- Lspoof evaluates how accurate our attack localizer can report the positions of both the original node and the attackers. We first present the statistical characterization of the location estimation errors.

Figure 7 presents the localization error CDF of the original nodes and the spoofing nodes for both RADAR and ABP in the 802.11 network as well as the 802.15.4 network. For the area-based algorithm, the median tile error ABP-med is presented, as well as the minimum and maximum tile errors, ABP-min and ABP-max.

We would like to examine how accurate the localization system can estimate the distance between Porg and P5p,, f. Figure 8 displays the relationship between Lorg- Lspoof and IPorg - Pspoof across different localization algorithms and networks.



(a) 802.11 network



(b) 802.15.4 network

Fig 7. Localization error CDF across localization algorithms and networks

The blue dots represent the cases of the detected spoofing attacks. While the red crosses indicate the spoofing attack has not been detected by the Kmeans spoofing detector. Comparing with

Figure 4, i.e. the detection rate as a function of the distance between Porg and Pspoof, the results of the undetected spoofing attack cases represented by the red crosses are in line with the results in Figure 4, the spoofing attacks are 100% detected when IPorg -Pspoof equals to or is greater than about 50 feet. Further, the relationship between Lorg- Lspoof and Porg -Pspoof is along the 45 degree straight line.

This means that Lorg- Lspoof is directly proportional to Porg -Pspoof and indicates that our localization system is highly effective for localizing the attackers.

Analysis is effective in both identifying the spoofing attacks as well as localizing the attackers. The challenge in localizing the positions of the attackers arises because the system does not know the positions of either the original MAC address or the node with the masquerading MAC. hus, we would like to examine how accurate the localization system can estimate the distance between Porg and P5p,,f. Figure 8 displays the relationship between Lorg- Lspoof and IPorg - Pspoof across different localization algorithms and networks. The blue dots represent the cases of the detected spoofing attacks. While the red crosses indicate the spoofing attack has not been detected by the K-means spoofing detector. Comparing with Figure 4, i.e. the detection rate as a function of the distance between Porg and Pspoof, the results of the undetected spoofing attack cases represented by the red crosses are in line with the results in Figure 4, the spoofing attacks are 100% detected when IPorg -Pspoof equals to or is greater than about 50 feet. Further, the relationship between Lorg- Lspoof and Porg - Pspoof is along the 45 degree straight line. This means that Lorg- Lspoof is directly proportional to Porg -Pspoof and indicates that our localization system is highly effective for localizing the attackers.

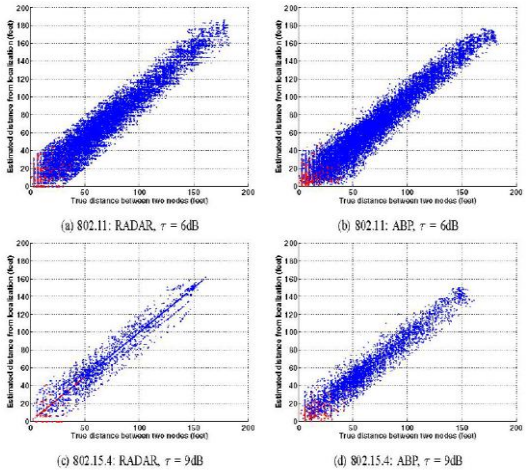


Fig 8. Relationship between the true distance and the estimated distance for the original node and the spoofing node across localization algorithms and networks

spoofing attacks are 100% detected when IPorg - Pspoof equals to or is greater than about 50 feet. Further, the relationship between Lorg- Lspoof and Porg -Pspoof is along the 45 degree straight line.

This means that Lorg- Lspoof is directly proportional to Porg -Pspoof and indicates that our localization system is highly effective for localizing the attackers.

CONCLUSION AND FUTURE ENHANCEMENT

The proposed work has used a generalized attack detection model that utilizes the received signal strength (RSS) 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 and cluster-based mechanisms are developed to determine the number of attackers. This approach can both detect the presence of attacks as well as determine the number of adversaries, spoofing the same node identity, so that any number of attackers can be localized. Further, based on the number of attackers determined by the mechanisms, an

integrated detection and localization system can localize any number of adversaries even when attackers using different transmission power levels.

In future, based on the outcome of this model, further we can find ways to eliminate those identified multiple adversaries, from the wireless network. Thus wireless networks will be more robust and less prone to attack. The proposed system further can increase the number of nodes and based area and verify this system in real time that means we can embed these algorithms on IC (Integrated Circuits) and link this IC's with real sensor and test in real time environment.

REFERENCES

1. D. Faria and D. Cheriton, "Detecting Identity-Based Attacks in Wireless Networks Using Signalprints," Proc. ACM Workshop Wireless Security (WiSe), Sept. 2006.
2. Jie Yang, Yingying (Jennifer) Chen and Wade Trappe, "Detection and Localization of Multiple Spoofing Attackers in Wireless Networks", IEEE Transaction on parallel and distributed system, Vol. 24, NO. 1, January 2013.
3. J. Yang and Y. Chen, "A Theoretical Analysis of Wireless Localization Using RF-Based Fingerprint Matching," Proc. Fourth Int'l Workshop System Management Techniques, Processes, and Services (SMTPS), Apr. 2008.
4. Y. Chen, W. Trappe, and R.P. Martin, "Detecting and Localizing Wireless Spoofing Attacks," Proc. Ann. IEEE Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON), May 2007.
5. F. Guo and T. Chiueh, "Sequence Number-Based MAC Address Spoof Detection," Proc. Eighth Int'l Conf. Recent Advances in Intrusion Detection, pp. 309-329, 2006.
6. M. Demirbas and Y. Song, "An rssi-based scheme for Sybil attack detection in wireless sensor networks," in Proceedings of the International Workshop on Advanced Experimental Activities on Wireless Networks and Systems, 2006.