DEFENDING SPOOFING ATTACK IN WIRELESS SENSOR NETWORKS

Sasubilli Santoshi kumari[1], Mr. G.Rajendra kumar[2]

1. M.Tech, Department of CSE, Sivani College of Engineering, Srikakulam, Andhra Pradesh, India.
2. Ph.D.,(C.S.E), Andhra University, Visakhapatnam, Andhra Pradesh

Abstract: In running days Wireless Sensor Network is emerging technology with their limited energy, computation, and communication capabilities. In contrast of traditional network, wireless sensor networks are deployed in accessible areas, presenting a risk of physical attacks. Sensor networks interact closely with the physical environment because of these reasons current security approaches are inadequate in WSN .In order to facilitate applications that require packet delivery from one or multiple senders to one or multiple receivers must need proper security mechanism. Here we presented different types of security issues in WSN

Keywords*:* Attacks, Routing protocols, Wireless network security, spoofing attack, attack detection, localization.

I. Introduction

The main basic goals of wireless sensor network are to gather information from the surrounding environment in which they are deployed. WSN have attracted much attention due to its great potential to be used in various applications.WSN consist of battery - operated sensors devices with computing, data processing, and communicating components. The sensors networks are generally deployed where monitoring and surveillance are required. Sensors are deployed in large numbers which are often impractical to gather from the individual sensors, particularly from the energy consumption point of view. The prime challenge for sensor networks consists of two facts. First, sensors are extremely resource constrained. Second, in many applications sensor nodes will be randomly deployed. This random deployment raises issue of dimensioning the network. Scattering too few nodes may result in lack of field coverage and disconnection in the network. On the other hand, scattering many nodes may result in an efficient network due to increased medium access control (MAC) collision and interference. Because of the limited resource on sensor nodes, size and density of the networks, unknown topology prior to deployment, and high risk of physical attacks to unattended sensors, it is a challenge to security of the wireless sensor networks. The security requirement is to provide confidentiality, integrity, authenticity, and availability of all information in limited resource constraints. the need to ensure security and privacy is becoming imperatively important. For example, in an 802.11 network, its easy for an attacker to gather useful MAC address information and can modify. It can also facilitate a variety of injection attacks such as attacks on access control lists, Denial of- Service (DoS) attacks and rogue access point (AP) attacks. In a large-scale network, multiple attackers use the same identity and launch malicious attacks such as denial-of-service attack network resource utilization attack and quickly. In this paper we propose two models.

1) GADE: a generalized attack detection model (GADE) can detect spoofing attacks and determine the number of attackers using spatial correlation of received signal strength (RSS) among normal devices and attackers

2) IDOL an integrated detection and localization system that can both detect attacks as well as find the positions of multiple attackers. In GADE, the Partitioning around Medoids (PAM) method is used to perform attack detection as well determining the number of attackers and Scattering too few nodes may result in lack of field coverage and disconnection in the network

Sensibleness of Attacks

These detect motion through the principle of Doppler radar, and are similar to a radar speed gun. A continuous wave of microwave radiation is emitted, and phase shifts in the reflected microwaves due to motion of an object toward (or away from) the receiver result in a heterodyne signal at low audio frequencies. In this section we provide a brief overview of spoofing attacks and their impact. We then discuss the experimental methodology that we use to evaluate our approach of spoofing detection.

Burlesque Attacks

Due to the open-nature of the wireless medium, it is easy for adversaries to monitor communications to find the layer-2 Media Access Control (MAC) addresses of the other entities. Recall that the MAC address is typically used as a unique identifier for all the nodes 2 on the network. Further, for most commodity wireless devices, attackers can easily forge their MAC address in order to masquerade as another transmitter. As a result, these attackers appear to the network as if they are a different device. Such spoofing attacks can have a serious impact on the network performance as well as facilitate any forms of security weaknesses, such as attacks on access control mechanisms in access points [16], and denial of-service through a de-authentication attack [17]. Abroad survey of possible spoofing attacks can be found in [7], [10]. To address potential spoofing attacks, the conventional approach uses authentication. However, the application of authentication requires reliable key distribution, management, and maintenance mechanisms. It is not always desirable to apply authentication because of its infrastructural, computational, and management overhead. Further, cryptographic methods are susceptible to node compromise– a serious concern as most wireless nodes are easily accessible, allowing their memory to be easily scanned. It is desirable to use properties that cannot be undermined even when nodes are compromised. We propose to use received signal strength (RSS), a property associated with the transmission and reception of communication (and hence not reliant on cryptography), as the basis for detecting spoofing. Employing RSS as a means to detect spoofing will not require any additional cost to the wireless devices themselves– they will merely use their existing communication methods, while the wireless network will use a collection of base stations to monitor received signal strength for the potential of spoofing.

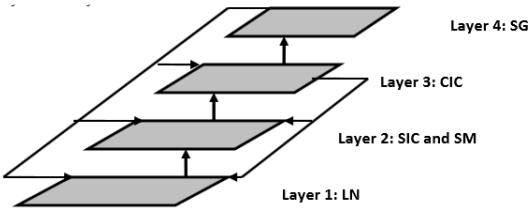
Experimental Methodology

In order to evaluate the effectiveness of our spoofing detection mechanisms, which we describe in the next section, we have conducted experiments using both an 802.11 (Wi-Fi)network as well as an network on the 3rd floor of the Computer Science Department at Rutgers University. The floor size is 200x80ft (Wi-Fi) network with 4 landmarks deployed to maximize signal strength coverage, as

shown in red triangles. The 802.15.4 (ZigBee) network is presented in Figure 1 (b) with 4 landmarks distributed in a squared setup in order to achieve optimal landmark placement [18] as shown in red squares. The small blue dots in the floor map are the locations used for spoofing and localization tests. 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. We randomly selected points out of the above locations as the training data for use by the localization algorithms. To test our approach's ability to detect spoofing, we randomly chose a point pair on the floor and treated one point as the position of the original node, and the other as the position of the spoofing node. We ran the spoofing test through all the possible combinations of point pairs on the floor using all the testing locations in both networks

Attack Detector

Most of the existing detection techniques have not met the requirements for practical deployment in wireless sensor network to mitigate sleep deprivation torture. In this section, a hierarchical model is proposed for wireless sensor network to detect the sensor nodes affected by sleep deprivation attack. It uses cluster based mechanism in an energy efficient manner. A dynamic detection model is designed here to overcome sudden death of IDS enabled sensor nodes. In this model responsibility of each node dynamically changes to reduce the burden of a single node. Our research focuses on distributed anomaly detection technique in order to provide a reliable and energy efficient heterogeneous wireless sensor network. Anomaly is detected by comparing the values with predefined parameters specified in normal profile. The proposed model uses anomaly detection technique in such a way so that false intrusion detection can be avoided. To mitigate the attack, proposed model physically excludes malicious nodes from the network and rejects fake packets. In heterogeneous sensor field, sensor nodes are categorized into various roles such as sink gateway (SG), cluster - in - charge (CIC), sector monitor(SM), sector - in - charge (SIC) and leaf node (LN) depending on their battery capacity. The roles of CIC, SM and SIC are changed dynamically to avoid exhaustion of nodes. Sink Gateway node is the honest gateway to another network or access point. SG is preset to perform gateway functionality. In this section we propose our spoofing attack detector. We first formulate the spoofing attack detection problem as one using classical statistical testing. Next, we describe the test statistic for spoofing detection. We then introduce the metrics to evaluate the effectiveness of our approach. Finally, we present our experimental results.



Formulation of Burlesque Attack Detection

RSS is widely available in deployed wireless communication networks and its values are closely correlated with location in physical space. In addition, RSS is a common physical property used by a widely diverse set of localization algorithms [13]–[15], [20]. In spite of its several meter-level localization accuracy, using RSS is an attractive approach because it can re-use the existing wireless infrastructure. We thus derive a spoofing attack detector utilizing properties of the RSS. The goal of the spoofing detector is to identify the presence of a spoofing attack. We formulate the spoofing attack detection as a statistical significance test, where the null hypothesis is

H0: normal (no attack).

In significance testing, a test statistic T is used to evaluate whether observed data belongs to the null hypothesis or not. If the observed test statistic Tobs

differs significantly from the hypothesized values, the null hypothesis is rejected and we claim the presence of a spoofing attack.

Test Statistic for Burlesque Detection

Although affected by random noise, environmental bias, and multipath effects, the RSS value vector, s ={s1, s2 ...sn} (n is the number of landmarks/access points(APs)), is closely related with the transmitter's physical location and is determined by the distance to the landmarks [15]. The RSS readings at different locations in physical space are distinctive. Each vector s corresponds to a point in a n-dimensional signal space [21]. When there is no spoofing, for each MAC address, the sequence of RSS sample vectors will be close to each other, and will fluctuate around a mean vector. As a result, the RSS sample readings from the attacked MAC address will be mixed with RSS readings from at least one different location. Based on the properties of the signal strength, the RSS readings from the same physical location will belong to the same cluster points in the n-dimensional signal space, while the RSS readings from different locations in the physical space should form different clusters in signal space. This observation suggests that we may conduct K means cluster analysis [22] on the RSS readings from each MAC address in order to identify spoofing. If there are M RSS sample readings for a MAC address, the K means clustering algorithm partitions M sample points into K disjoint subsets Sj containing Mj sample points so as to minimize the sum-of-squares criterion.

$$J_{min} = \sum_{j=1}^{K} \sum_{\mathbf{s_m} \in S_j} \|\mathbf{s_m} - \mu_j\|^2$$

where sm is a RSS vector representing the mth sample point and μj is the geometric centroid of the sample points for Sj in signal space. Under normal conditions, the distance between the centroids should be close to each other since there is basically only one cluster. Under a spoofing attack, however, the distance between the centroids is larger as the centroids are derived from the different RSS clusters associated with different locations in physical space. To illustrate, we use the following definitions, an original node P org is referred to as the wireless device with the legitimate MAC address, while a 4 spoofing node P spoof is referred to as the wireless device that is forging its identity and masquerading as another device. There can be multiple spoofing nodes of the same MAC address.

Thresholds Sell Up

A sensor network is composed of a large number of sensor nodes that are densely deployed [1]. These nodes have the ability to communicate either among each neighbor or directly to the base station [2]. Sensor nodes often have limited computation and communication resources and battery power. Sensor nodes are affected by physical attacks, potentially compromising the sensor's cryptographic keys, since they are deployed in hostile environments. [3] An adversary may use compromised nodes to inject false reports into the network. False reports may not only cause false alarms, but also the depletion of the serious amount of energy in each forwarding node [5]. To minimize critical damage, false reports should be dropped en-route as early as possible and the few elusive ones should be rejected at the base station. The early dropping of false reports leads to significant energy saving.

Staging Metrics

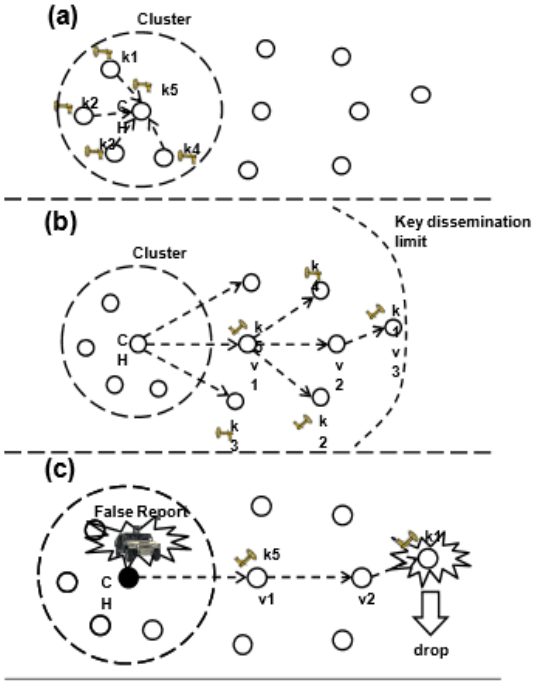
Wireless sensor networks (WSN) have emerged as one of the most exciting fields in Computer Science research over the past 15 years. Processors with on-board sensors are said to be nearing the size of a dust. Applications of WSN include military surveillance, habitat monitoring, structural monitoring and cargo tracking. The evolution of the field may be followed in research papers written by prominent personalities and institutions in Computer Science research. A macrocosm of topics in those papers is surveyed and their evaluation techniques are assessed in this paper. The topics include storage, routing, real-time communication, power management and architecture. These topics are discussed in the following sections. The discussion will be organized in five sections and each section will be focused on a research paper that presents an implementation relating to the topic. In each section, a general introduction to the topic is given. The introduction is followed by a summary of the evaluation of the implementation as presented in the research paper. This will span two subsections-experimental set-up and results. The following subsection, critique, reviews the evaluation techniques under four criteria data, workloads, factors and metrics selected.

The emergence of many kinds of networked data-centric sensor applications has given more importance to data generated by the sensors. Sensors in these applications probe the environment for useful data for analysis. To achieve a useful infrastructure for users, live data need to be processed, interpreted, filtered and archivedoften using stored data. Archival storage of past sensor data requires a storage system. A good storage system must address issues such as where the data is stored, whether the data is indexed and how the application can access this data in an energy efficient manner. One such storage system, Two-Tier Storage Architecture (TSAR), is analyzed in this section. TSAR is a two tier storage system which seeks to improve upon the existing homogenous storage system. The evaluation of TSAR was presented in the paper "TSAR: A Two Tier Storage Architecture Using Interval Skip Graphs"

Experimental evaluation

The outdoor routing experiment took place on a rectangular athletic field measuring approximately 225 (north-south) by 365 (east-west) meters. This field can be roughly divided into four flat, equal-sized sec-tions, three of which are at the same altitude, and one of which is approximately four to six meters lower. There was a short, steep slope between the upper and lower sections. In this section we present the evaluation results of the effectiveness of the spoofing attack detector. Table

I presents the detection rate and false positive rate for both the 802.11 network and the 802.15.4 network under different threshold settings. The corresponding ROC curves are displayed in Figure 3. The results are encouraging showing that for



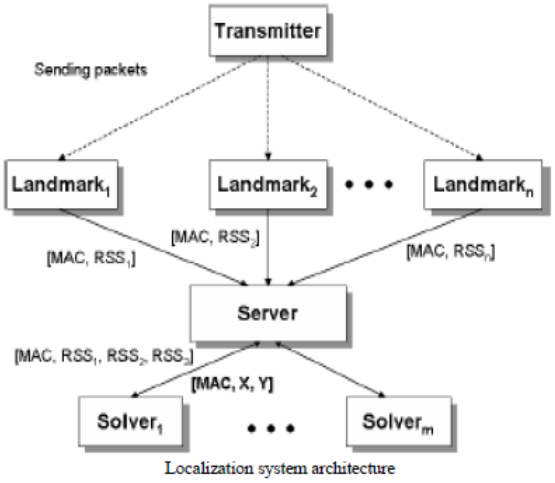
false positive rates less than 10%, the detection rates are above 95%. Even when the false positive rate goes to zero, the detection rate is still more than 95% for both 802.11 and 802.15.4 networks. We further study how likely a spoofing node can be detected by our spoofing attack detector when it is at varying distances from the original node in physical space. Figure 4 presents the detection rate as a function of the distance between the spoofing node and the original node. We found that the further away Pspoof is from Porg, the higher the detection rate becomes. For the 802.11 network, the detection rate goes to over 90%.

Localization System

We have developed a general-purpose localization system to perform 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.

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 system architecture

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 number 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.

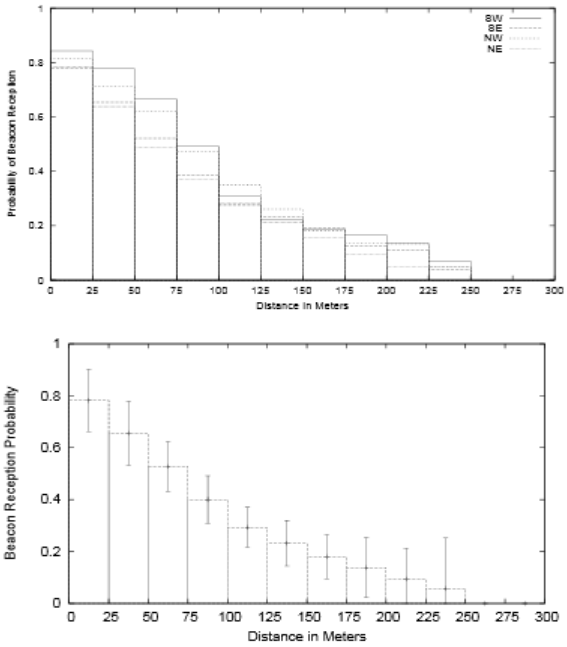
Attack Localizer

When our spoofing detector has identified an attack for a MAC address, the centroids returned by the Kmeans 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 5 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 Lspoof .

Experimental Evaluation

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 CDF

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.



Impact

We demonstrate above that the axioms are untrue, but a key question remains: what is the effect of these axioms on the quality of simulation results? In this section, we begin by comparing the results of our outdoor experiment with the results of a best-effort simulation model, and then progressively weaken the model by assuming some of the axioms. The purpose of this study is not to claim that our simulator can accurately model the real network environment, but instead to show quantitatively the impact of the axioms on the simulated behavior of routing protocols. Clearly, analytical or simulation research in wireless networking must work with an abstraction of reality, modeling the behavior of the wireless network below the layer of interest. Unfortunately, overly simplistic assumptions can lead to

misleading or incorrect conclusions. Our results provide a counter-example to the notion that these axioms are sufficient for research on ad hoc routing algorithms. We do not claim to validate, or invalidate, the results of any other published study. Indeed, our point is that the burden is on the authors of past and future studies to

a) Clearly lay out their assumptions, b) demonstrate

whether those assumptions are reasonable within the context of their study, and c) clearly identify any limitations in the conclusions they draw. While others have used simulation to explore the impact of different radio propagation models [TMB01,ZHKS04], we use the identical implementation of the routing protocol in both the simulator and the experiment [LYN+04], use a large number of nodes in an outdoor experiment [GKN+04],and are able to compare our simulation results with the actual experiment

We begin by comparing the results of the outdoor experiment with the simulation results obtained with our best signal propagation model and a detailed 802.11 protocol model. The best signal propagation model is a stochastic model that captures radio signal attenuation as a combination of two effects: smallscale fading and large-scale fading. Small-scale fading describes the rapid fluctuation in the envelope of a transmitted radio signal over a short period of time or a small distance, and primarily is caused by multipath effects. Although small-scale fading is in general hard to predict, wireless researchers over the years have proposed several uccessful statistical models for small-scale fading, such as the Rayleigh and Ricean distributions. Large-scale fading describes the slowly varying signal-power level

over a long time interval or a large distance, and has two major contributing factors distance pathloss and shadow fading. The distance path-loss models the average signal power loss as a function of distance. the receiving signal strength is proportional to the distance between the transmitter and the receiver raised to a given exponent. Both the free-space model and the two-ray ground reflection model mentioned earlier can be classified as distance path-loss models. The shadow fading describes the variations in the receiving signal power due to scattering; it can be modeled as a zero-mean log-normal distribution. Rappaport [Rap96] provides a detailed discussion of these and other models.For our simulation, given the light traffic used in the real experiment, we used a simple SNR threshold approach instead of a more computational intensive BER approach. Under heavier traffic, this choice might have substantial impact [TMB01]. For the propagation model, we chose 2.8 as the distance Comparing packet delivery ratios between real experiment and simulation. log normal standard. These values, which must be different for different types of terrain, produce signal propagation distances consistent with our observations from the real network. This duplicated the 7 crashed nodes from the real experiment, and allowed us to reproduce the same traffic pattern. shows the difference in the overall packet delivery ratio (PDR)—which is the total number of packets received by the application layer divided by the total number of packets sent—between the real experiment and the simulation. The simple propagation model produced relatively good results, the relative errors in predicted PDR were within 10% for all three routing protocols tested. We caution, however, that one cannot expect consistent results when generalizing the simple stochastic radio propagation model to deal with all network scenarios. After all, this model assumes some of the axioms we have identified, including flat earth, omni-directional

radio propagation length, and symmetry. Thus this model, our best, nonetheless assumes some of the same axioms we discount in the preceding section. This ironic situation is testimony to the difficulty of detailed radio and environment modeling; in situations where such assumptions are clearly invalid—for example, in an urban area—we should expect the model to deviate further from reality. On the other hand, this approximation is sufficient for the purposes of this paper, because we can still demonstrate how the other axioms may affect performance. On the other hand, since the model produced good results amenable to our particular outdoor experiment scenario, we use it in this study as the base line to quantify the effect of the axioms on simulation studies. As we show, these assumptions can significantly undermine the validity of the simulation results

Result

First, we look at the reception ratio of the beacon messages, which were periodically sent via broad casts by the beacon service program on each node. We calculate the reception ratio by inspecting the entries in the beacon logs, just as we did for the real experiment. Plots the beacon reception ratios during the execution of the AODV routing protocol. The choice of routing protocol is unimportant in this study since we are comparing the results between the real experiment and simulations. We understand that the control messages used by the routing protocol may slightly skew the beacon reception ratio due to the competition at the wireless channel. Compared with the two simple models, our best

Conclusions

The great majority of these papers rely on overly simplistic assumptions of how radios work. Both widely used radio models, “flat earth” and ns-2 “802.11” models, embody the following set of axioms: the world is two dimensional; a radio’s transmission area is roughly circular; all radios have equal range; if I can hear you, you can hear me; if I can hear you at all, I can hear you perfectly; and signal strength is a simple function of distance.Others have noted that real radios and ad hoc networks are much more complex than the simple models used by most researchers [PJL02], and that these complexities have a significant impacton the behavior of MANET protocols and algorithms [GKW+02]. In this paper, we enumerated the set of common assumptions used in MANET research, and presented a real-world experiment that strongly contradicts these “axioms.” The results cast doubt on published simulation results that implicitly rely on these assumptions, In this work, we proposed a method for detecting spoofing attacks as well as localizing the adversaries 9 in wireless and sensor networks. In contrast to traditional identity-oriented authentication methods, our RSS based approach does not add additional overhead to the wireless devices and sensor nodes. We formulated the spoofing detection problem as a classical statistical significance testing problem. We then utilized the K-means cluster analysis to derive the test statistic. Further, we have built a real-time localization system and integrated our K-means spoofing detector into the system to locate the positions of the attackers and as a result to eliminate the adversaries from the network. We studied the effectiveness and generality of our spoofing detector and attacker localizer in both an 802.11 (Wi-Fi) network and an 802.15.4 (ZigBee) network in a real office building environment. The performance of the K-means spoofing detector is evaluated in terms of detection rates and receiver operating characteristic curves. Our spoofing detector has achieved high detection rates, over

95% and low false positive rates, below 5%. When locating the positions of the attackers, we have utilized both the point-based and area-based algorithms in our realtime localization system. We found that the performance of the system when localizing the adversaries using the results of K-means cluster analysis are about the same as localizing under normal conditions. Usually the distance between the spoofing node and the original node can be estimated with median error of 10 feet. Our method is generic across different localization algorithms and networks.Therefore, our experimental results provide strong evidence of the effectiveness of our approach in detecting the localizing the positions of the adversaries and spoofing attacks.

References

1. M. Takai, R. Bagrodia, K. Tang, and M. Gerla. Efficient wireless network simulations with detailed propagation models. Wireless Networks , 7(3):297–305, May 2001.

2. M. bohge and W. Trappe, “An authentication framework for hierarchical ad hoc sensor networks,” in Proceedings of the ACM Workshop on Wireless Security (WiSe), 2003, pp. 79–87.

3. B. Wu, J. Wu, E. Fernandez, and S. Magliveras, “Secure and efficient key management in mobile ad hoc networks,” in Proceedings of the 19th EEE International Parallel and Distributed Processing Symposium (IPDPS), 2005.

4. A.Wool, “Lightweight key management for ieee 802.11 wireless lans with key refresh and host revocation,” ACM/Springer Wireless Networks, vol. 11, no. 6, pp. 677–686, 2005.

5. K. Pawlikowski, H.-D.J Jeong, and J.-S.R. Lee. On credibility of simulation studies of telecommunication networks . IEEE Communications , 40(1):132–139, January 2002

6. S. Zhu, S. Xu, S. Setia, and S. Jajodia, “Lhap: A lightweight hop-by-hop authentication protocol for ad-hoc networks,” in Proceedings of the IEEE International Workshop on Mobile and Wireless Network (MWN), 2003, pp. 749–755.

7. T. Aura, “Cryptographically generated addresses (cga),” RFC 3972, IETF, 2005.

8. Q. Li and W. Trappe, “Relationship-based detection of spoofingrelated anomalous traffic in ad hoc networks,” in Proceedings of the Third Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), September 2006.

9. E. Kempf, J. Sommerfeld, B. Zill, B. Arkko, and P. Nikander, “Secure neighbor discovery (send),” RFC 3971, IETF, 2005.

10. Q. Li and W. Trappe, “Light-weight detection of spoofing attacks in wireless networks,” in Proceedings of the 2nd International Workshop on Wireless and Sensor Network Security (WSNS), October 2006.

11. Z. Li, W. Xu, R. Miller, and W. Trappe, “Securing wireless systems via lower layer

enforcements," in Proceedings of the ACM Workshop on Wireless Security (WiSe), 2006.

12. D. Faria and D. Cheriton, "Detecting identitybased attacks in wireless networks using signalprints," in Proceedings of the ACM Workshop on Wireless Security (WiSe),s.