Thwarting Spoofing Attacks Using RARE Model

[1]Swapnil Dutta[2]Sidharth Shishodia [3]Sagar Agarwal [4]Mahalakshmi S
[1]neal.m2014@gmail.com,[2]sidsidhu121@gmail.com,[3]sagar.agl@gmail.com,[4]maha.shanmugam@gmail.com
[1 2 3]Student, [4]Assistant Professor
Department of Information Science Engineering
BMS Institute of Technology, Yelahanka, Bengaluru 560064

***Abstract*— Wireless sensor networks (WSNs) have many potential applications. In many scenarios WSNs are of interest to adversaries and they become susceptible to some types of attacks since they are deployed in open environments and have limited resources. Spoofing attack is one of the most common network attacks. WSNs are in peril to spoofing attacks which leads to many other forms of attack on the network. In this paper, we propose a Recognition And REvelation[RARE] model which uses spatial information, a physical property associated with each node, as the basis for recognizing spoofing attacks and revealing multiple adversaries that masquerade with the same node identity. We present Recognition[R] model to use the Medoid based clustering mechanism for a clustering with a node whose average dissimilarity to all the objects in the cluster is minimal. We propose the concept of received signal strength (RSS) conceived from wireless nodes to detect the spoofing attacks. In addition, we developed Revelation[RE] model that can determine the number of attackers and reveal the positions of the multiple attackers. Our revelation mechanism uses a representative set of algorithms that can provide positive results for recognizing and revealing multiple adversaries.**

Keywords—Wireless Network, Network Security, RSS, Recognition, Revelation, Clustering, Spoofing Attacks, RARE Model, PAM, SILENCE, SVM

I. INTRODUCTION

The IEEE 802.11 is the adopted standard for WLANs. 802.11 is a set of media access control (MAC) and physical layer (PHY) specifications protocol for implementing wireless local area network (WLAN) computer communication. Wireless networks have gained popularity as compared to the wired network due to the flexibility, low cost and easy deployment layouts. The combination of free spectrum, efficient channel coding and cheap interface hardware has made 802.11-based access networks extremely popular.

WLAN are widely used by laptop users on the corporate and educational environments. However, some fundamental weaknesses of the wireless access medium make wireless networks more vulnerable to attacks. The widespread deployment makes wireless networks an attractive target for potential attackers. For a couple hundred dollars a user can buy an 802.11 access point that seamlessly extends their existing network connectivity for almost 100 meters.

Among various types of attacks, identity-based spoofing attacks are especially easy to launch and can cause significant damage to network. Web spoofing allows an attacker to create a "shadow copy" of the entire World Wide Web. Accesses to the shadow Web are funnelled through the attacker's machine, allowing the attacker to monitor all of the victim's activities including any passwords or account numbers the victim enters. The attacker can also cause false or misleading data to be sent to Web servers in the victim's name, or to the victim in the name of any Web server. In short, the attacker observes and controls everything the victim does on the Web.

There are two ways that IP spoofing attacks can be used to overload targets with traffic:

1. Simply flood a selected target with packets from multiple spoofed addresses. This method works by directly sending a victim more data than it can handle.
2. Spoof the target's IP address and send packets from that address to many different recipients on the network. When another machine receives a packet, it will automatically transmit a packet to the sender in response. Since the spoofed packets appear to be sent from the target's IP address, all responses to the spoofed packets will be sent to (and flood) the target's IP address.

Spoofing attacks can further facilitate a variety of traffic injection attacks [1], [2], such as attacks on access control lists, rogue access point attacks, and eventually Denial-of- Service (DoS) attacks. A broad survey of possible spoofing attacks can be found in [3], [4]. Moreover, in a large-scale network, multiple adversaries may masquerade as the same identity and collaborate to launch malicious attacks such as network resource utilization attack and denial-of-service attack quickly.

Therefore, it is important to

- recognize the presence of spoofing attacks,
- determine the number of attackers, and
- reveal multiple adversaries and eliminate them.

The traditional approach employs cryptographic schemes to address potential spoofing attacks [5], [6]. However, the application of cryptographic schemes 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. Further, cryptographic methods are susceptible to node compromise, which is a serious concern as most wireless nodes are easily accessible, allowing their memory to be easily scanned. 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.

This paper proposes to use RSS-based spatial correlation, a physical property associated with each wireless node that is hard to falsify and not reliant on cryptography as the basis for recognizing spoofing attacks. Since the concern is on the attackers who have different locations than legitimate wireless nodes, utilizing spatial information to address spoofing attacks has the unique power to not only identify the presence of these attacks but also reveal adversaries.

An added advantage of employing spatial correlation to recognize spoofing attacks is that it will not require any additional cost or modification to the wireless devices themselves. The focus is on static nodes in this work, which are common for spoofing scenarios [7]. The works that are closely related are [3], [7], [9]. [3] Proposed the use of matching rules of signal prints for spoofing recognition, [7] modeled the RSS readings using a Gaussian mixture model and [9] used RSS and K-medoid cluster analysis to recognize spoofing attacks. However, none of these approaches have the ability to determine the number of attackers when multiple adversaries use a same identity to launch attacks, which is the basis to further reveal multiple adversaries after attack detection. Although [9] studied how to localize adversaries, it can only handle the case of a single spoofing attacker and cannot localize the attacker if the adversary uses different transmission power levels.

The main contributions of the work are:

RARE Model: We formulate a Recognition And Revelation [RARE] model which utilizes the theoretical analysis of exploiting the spatial correlation of the RSS inherited from wireless nodes for attack detection. In our theoretical analysis, we first derived the mathematical relationship between the distance of RSS in signal space and the node distance in physical space. We then developed the analytical expression of the detection rate, false-positive rate, and accuracy of determining whether two nodes reside at the same location based on the RSS distance in signal space which uses spatial information, a physical property associated with each node, as the basis for recognizing spoofing attacks and revealing multiple adversaries that masquerade with the same node identity. We present Recognition[R] model to use the Medoid based clustering mechanism for a clustering with a node whose average dissimilarity to all the objects in the cluster is minimal. We propose the concept of received signal strength (RSS) conceived from wireless nodes to recognize the spoofing attacks. In addition, we developed Revelation [RE] model that can determine the number of attackers and reveal the positions of the multiple attackers.

In Recognition[R] model, the Partitioning around Medoids (PAM) cluster analysis method is used to perform attack detection. The problem of determining the number of attackers as a multi-class detection problem is formulated. Then cluster based methods are applied to determine the number of attacker. Further a mechanism called SILENCE is used for testing Silhouette Plot and System Evolution with minimum distance of clusters, to improve the accuracy of determining the number of attackers. Additionally, when the training data is available, Support Vector Machines (SVM) method [21] is used to further improve the accuracy of determining the number of attackers.

Moreover, Revelation [RE] model is used which utilizes the results of the number of attackers returned by [R] model to further reveal multiple adversaries. The recognition and revelation mechanism uses a representative set of algorithms that can provide positive results for recognizing and revealing multiple adversaries.

II. Related work

The traditional approach to prevent spoofing attacks is to use cryptographic-based authentication [5], [6], [10]. Wu et al. [5] have introduced a secure and efficient key management (SEKM) framework. SEKM builds a Public Key Infrastructure (PKI) by applying a secret sharing scheme and an underlying multicast server group. Wool [6] implemented a key management mechanism with periodic key refresh and host revocation to prevent the compromise of authentication keys.

Recently, new approaches utilizing physical properties associated with wireless transmission to combat attacks in wireless networks have been proposed. Based on the fact that wireless channel response deco relates quite rapidly in space, a channel-based authentication scheme was proposed to discriminate between transmitters at different locations, and thus to detect spoofing attacks in wireless networks [11]. Brik et al. [12] focused on building fingerprints of 802.11bWLAN NICs by extracting radiometric signatures, such as frequency magnitude, phase errors, and I/Q origin offset, to defend against identity attacks. However, there is additional overhead associated with wireless channel response and radiometric signature extraction in wireless networks. Li and Trappe [4]

Introduced a security layer that used forge-resistant relationships based on the packet traffic, including MAC sequence number and traffic pattern, to detect spoofing attacks. The MAC sequence number has also been used in [13] to perform spoofing detection. Both the sequence number and the traffic pattern can be manipulated by an adversary as long as the adversary learns the traffic pattern under normal conditions. The works [3], [7], [14] using RSS to defend against spoofing attacks are most closely related to our work. Faria and Cheriton [3] proposed the use of matching rules of signal

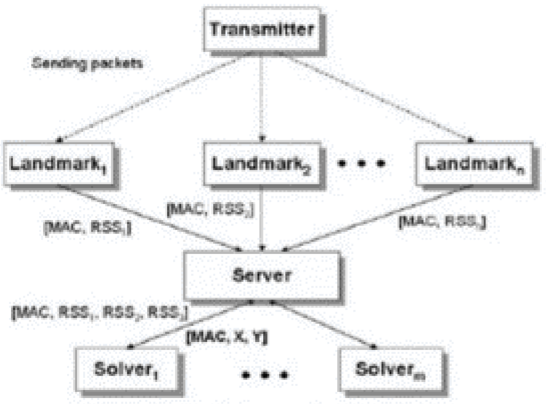

Fig1. Revelation System Architecture

prints for spoofing detection. Sheng et al. [7] modeled the RSS readings using a Gaussian mixture model. Sang and Arora [14] proposed to use the node's "spatial signature," including Received Signal Strength Indicator (RSSI) and Link Quality Indicator (LQI) to authenticate messages in wireless networks. However, none of these approaches are capable of determining the number of attackers when there are multiple adversaries collaborating to use the same identity to launch malicious attacks. Further, they do not have the ability to localize the positions of the adversaries after attack detection.

While studying revelation techniques in spite of its several meter-level accuracy, using RSS [14] [15], is an attractive approach because it can reuse the existing wireless infrastructure and is highly correlated with physical locations. Dealing with ranging methodology, range-based algorithms involve distance estimation to landmarks using the measurement of various physical properties such as RSS [14] [15], Time Of Arrival (TOA), Time Difference Of Arrival (TDOA), and direction of arrival (DoA). Whereas range-free algorithms use coarser metrics to place bounds on candidate positions. Another method of classification describes the strategy used to map a node to a location. Alteration approaches use distances to landmarks, while angulation uses the angles from landmarks. Scene matching strategies [15] use a function that maps observed radio properties to locations on a pre-constructed signal map or database. Further, Chen proposed to perform detection of attacks on wireless localization [18] and Yang proposed to use the direction of

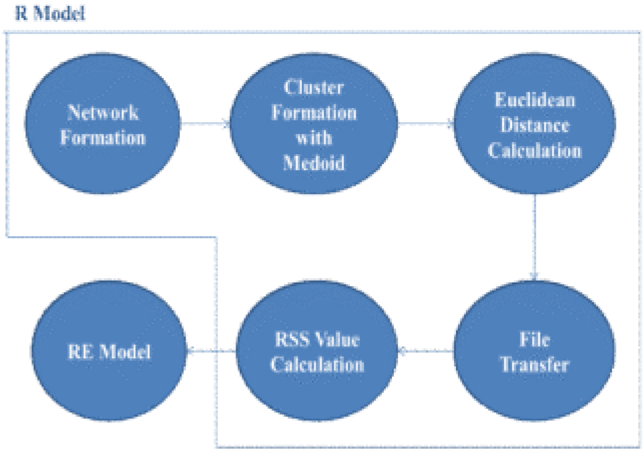

Fig2. Data Flow Diagram

This work differs from the previous studies in that here the spatial information is used to assist in attack recognition instead of relying on cryptographic-based approaches. Furthermore, this work is novel because it can determine the number of attackers when there are multiple adversaries masquerading as the same identity. In this work, we choose a group of algorithms employing RSS [16] to perform the task of revealing multiple attackers and evaluate their performance in terms of revelation accuracy. Our proposed work provides positive results in recognizing the attackers efficiently. Additionally, this approach can accurately reveal multiple adversaries even when the attackers vary their transmission power levels to trick the system of their true locations.

The overall work tends to avoid all the repercussions that tend to occur during attending to spoofing attacks in case of wireless sensor networks. Improvement in consistency as compared to the previous works is evitable in this work with the use of efficient algorithms for recognizing multiple spoofing attackers that are masquerading within the wireless networks and also revealing their positions in the network.

III. OVERVIEW OF TECHNIQUES

A. Recognition[R] Model

Recognition[R] model uses the Medoid based clustering mechanism for clustering with a node whose average dissimilarity to all the objects in the cluster is minimal. We propose the concept of received signal strength (RSS) conceived from wireless nodes to detect the spoofing attacks. n-dimensional signal space, while the RSS readings from different locations over time should form different clusters in signal space. Since under a spoofing attack, the RSS readings from the victim node and the spoofing attackers are mixed together, this observation suggests that we may conduct cluster analysis on top of RSS-based spatial correlation to find out the distance in signal space and further detect the presence of spoofing attackers in physical space.

In this work, we utilize the Partitioning Around Medoids[PAM] Method to perform K-medoid clustering analysis in RSS. The PAM method [19] is more suitable in determining clusters from RSS streams, which can be unreliable and fluctuating over time due to random noise and environmental bias [20].

Further, inaccurate estimation of the number of attackers will cause failure in revealing the multiple adversaries. As we do not know how many adversaries will use the same node identity to launch attacks, determining the number of attackers becomes a multiclass detection problem and is similar to determining how many clusters exist in the RSS readings.
We propose the use of Support Vector Machines [24] to classify the number of spoofing attackers. The advantage of using SVM is that it can combine the intermediate results (i.e., features) from different statistic methods. The SVM-based mechanism uses the training data to build a prediction model and it also takes the advantage of the combined features from two statistic methods, System Evolution and SILENCE mechanism in

performing multiclass attacker detection when multiple attackers are present in the system.

B. Revelation[RE] Model

In this section we present a real-time localization system that can be used to reveal the positions of the attackers. Revelation [RE] model can reveal the positions of the multiple attackers. Our revelation mechanism uses a representative set of algorithms that can provide positive results for recognizing and revealing multiple adversaries. In wireless spoofing attacks, the RSS stream of a node identity may be mixed with RSS readings of both the original node as well as spoofing nodes from different physical locations. Compromised localization results are a serious threat because of their impact on applications. So, it is important that the revelation mechanism we use must be able to distinguish between the readings and correctly reveal the spoofing node.

Different from the traditional localization approaches, our REvelation model utilizes the RSS measure of medoids returned Recognition model as inputs to localization algorithms to estimate the positions of adversaries. The return positions from our system include the location estimate of the original node and the attackers in the physical space. This allows revealing the position of the attackers based on the relationship between the true distance and the estimated distance for the original node and the spoofing node across localization algorithms and networks.

Fig.2. depicts the system flow of the RARE i.e. Recognition and Revelation model in order to thwart spoofing attacks in wireless networks. The [R] model recognizes the attackers and determines the number of spoofing attackers in the wireless network, and the [RE] model reveals the positions of the attackers in the wireless network.

IV. ALGORITHMS

In order to evaluate the generality of RE model for revealing adversaries, a set of representative revelation algorithms ranging from nearest neighbor matching in signal space (RADAR), to probability-based (Area-Based Probability), and to multi-alteration (Bayesian Networks) are chosen.

A. RADAR-Gridded:

The RADAR-Gridded algorithm is a scene-matching revelation algorithm. RADAR-Gridded uses an interpolated signal map, which is built from a set of averaged RSS readings with known (x, y) locations. Given an observed RSS reading with an unknown location, RADAR returns the x, y of the nearest neighbor in the signal map to the one to localize, where "nearest" is defined as the Euclidean distance of RSS points in an N-dimensional signal space, where N is the number of landmarks.

B. Area Based Probability (ABP):

ABP also utilizes an interpolated signal map. Further, the experimental area is divided into a regular grid of equal sized tiles. ABP assumes the distribution of RSS for each landmark follows a Gaussian distribution with mean as the expected value of RSS reading vectors. ABP then computes the probability of the wireless device being at each tile L_i, with $i = 1 \ldots L$, on the floor using

Bayes' rule:

$$P(L_i|s) = P(s|L_i) * p(L_i) / P(s)$$

Given that the wireless node must be at exactly one tile satisfying L_i=1 $P(L_i|s) = 1$, ABP normalizes the probability and returns the most likely tiles/grids up to its confidence α.

C. Bayesian Networks (BN):

BN localization is a multi-alteration algorithm that encodes the signal-to-distance propagation model into the Bayesian Graphical Model for localization. Figure 2 shows the basic Bayesian Network used for our study. The vertices X and Y represent location; the vertex s_i is the RSS reading from the ith landmark; and the vertex D_i represents the Euclidean distance between the location specified by X and Y and the ith landmark. The value of s_i follows a signal propagation model $s_i = b_{0i} + b_{1i} \log D_i$, where b_{0i}, b_{1i} are the parameters specific to the ith landmark.

The distance $D_i = \sqrt{} (X - x_i)^2 + (Y - y_i)^2$ in turn depends on the location (X, Y) of the measured signal and the coordinates (x_i, y_i) of the ith landmark. The network models noise and outliers by modeling the s_i as a Gaussian distribution around the above propagation model, with variance:

$$\tau_i\text{: } s_i \sim N(b_{0i} + b_{1i} \log D_i, \tau_i$$

Through Markov Chain Monte Carlo (MCMC) simulation, BN returns the sampling distribution of the possible location of X and Y as the revelation result.

V. CONCLUSION

In this work, we proposed a method for detecting spoofing attacks as well as revealing the positions of the adversaries in wireless and sensor networks. In this work, we propose 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 recognizing if there is any spoofing attack in wireless 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. 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 and can eliminate them.

Determining the number of adversaries is a particularly challenging problem. This paper uses SILENCE, a mechanism that employs the minimum distance testing in addition to cluster analysis to achieve better accuracy of determining the number of attackers than other methods under study, such as Silhouette Plot and System Evolution that use cluster analysis alone. Additionally, when the training data is available, Support Vector Machines (SVM) based mechanism [22] [23] is used to further improve the accuracy of determining the number of attackers present in the system. SVM based mechanism takes the advantage of the combined features from two statistic methods, System Evolution and SILENCE mechanism in performing multiclass attacker detection when multiple attackers are present in the system.

VI. Acknowledgment

Swapnil D., Sidharth S., and Sagar A. thank Dr. Rajasimha Makaram, the Head of the department, ISE, BMSIT, Bangalore for the facilities and valuable guidance he provided for the completion of this paper.

Special regards for Mr. Mahalakshmi S and Mr. Rajagopal who played a major role in shaping this paper the way it has turned out.

Most sincere gratitude and thanks to all friends and family for their valuable comments and suggestions for making this work a success.

References

[1] J. Bellardo and S. Savage, "802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions," *Proc.* USENIX *Security Symp.*, pp. 15-28, 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.

[3] D. Faria and D. Cheriton, "Detecting Identity-Based Attacks in Wireless Networks Using Signalprints," *Proc. ACM Workshop Wireless Security (WiSe)*, Sept. 2006.

[4] 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.

[5] B. Wu, J. Wu, E. Fernandez, and S. Magliveras, "Secure and Efficient Key Management in Mobile Ad Hoc Networks," *Proc. IEEE Int'l Parallel and Distributed Processing Symp. (IPDPS)*, 2005.

[6] 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.

[7] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 MAC Layer Spoofing Using Received Signal Strength," *Proc. IEEE INFOCOM*, Apr. 2008.

[8] J. Yang, Y. Chen, and W. Trappe, "Detecting Spoofing Attacks in Mobile Wireless Environments," *Proc. Ann. IEEE Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Networks (SECON)*, 2009.

[9] 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.

[10] M. Bohge and W. Trappe, "An Authentication Framework for Hierarchical Ad Hoc Sensor Networks," *Proc. ACM Workshop Wireless Security (WiSe)*, pp. 79-87, 2003.

[11] L. Xiao, L.J. Greenstein, N.B. Mandayam, and W. Trappe, "Fingerprints in the Ether: Using the Physical Layer for Wireless Authentication," *Proc. IEEE Int'l Conf. Comm. (ICC)*, pp. 4646-4651, June 2007.

[12] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless Device Identification with Radiometric Signatures," *Proc. 14th ACM Int'l Conf. Mobile Computing and Networking*, pp. 116-127, 2008.

[13] 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.

[14] L. Sang and A. Arora, "Spatial Signatures for Lightweight Security in Wireless Sensor Networks," *Proc. IEEE INFOCOM*, pp. 2137-2145, 2008.

[15] P. Bahl and V.N. Padmanabhan, "RADAR: An in-Building RF Based User Location and Tracking System," *Proc. IEEE INFOCOM*, 2000.

[16] E. Elnahrawy, X. Li, and R.P. Martin, "The Limits of Localization Using Signal Strength: A Comparative Study," *Proc. IEEE Int'l Conf. Sensor and Ad Hoc Comm. and Networks (SECON)*, Oct. 2004.

[17] Y. Chen, J. Francisco, W. Trappe, and R.P. Martin, "A Practical Approach to Landmark Deployment for Indoor Localization,"*Proc. IEEE Int'l Conf. Sensor and Ad Hoc Comm. and Networks (SECON)*, Sept. 2006.

[18] Y. Chen, W. Trappe, and R. Martin, "Attack Detection in Wireless Localization," *Proc. IEEE INFOCOM*, Apr. 2007.

[19] L. Kaufman and P.J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, *Wiley Series in Probability and Statistics, 1990.*

[20] G. Zhou, T. He, S. Krishnamurthy, and J.A. Stankovic, "Models and Solutions for Radio Irregularity in Wireless Sensor Networks," ACM Trans. Sensor Networks, *vol. 2,pp. 221-262, 2006*

[21] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods., *Cambridge Univ. Press, 2000.*

[22] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support VectorMachines,Software, *http://www.csie.ntu.edu.tw/cjlin/libsvm, 2001.*

[23] V. Franc and V. Hlava´c,"Multi-Class Support Vector Machine," *Proc. Int'l Conf. Pattern Recognition (ICPR), vol. 16, pp. 236-239, 2002.*

[24] C. Hsu and C. Lin, "A Comparison of Methods for Multiclass Support Vector Machines," *proc. IEEE Trans. Neural Networks, vol. 13, no. 2, pp. 415-425, Mar. 2002.*