This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2009 proceedings.

Determining the Number of Attackers and Localizing Multiple Adversaries in Wireless Spoofing Attacks

Jie Yang*, Yingying Chen*, Wade Trappe†, Jerry Cheng†
*Dept. of ECE, Stevens Institute of Technology
Castle Point on Hudson, Hoboken, NJ 07030
{jyang, yingying.chen}@stevens.edu
† WINLAB and Dept. of Statistics, Rutgers University
110 Frelinghuysen Rd, Piscataway, NJ 08854
trappe@winlab.rutgers.edu, *jcheng@stat.rutgers.edu*

***Abstract*—Wireless spoofing attacks are easy to launch and can significantly impact the performance of networks. Although the identity of a node can be verified through cryptographic authentication, conventional security approaches are not always desirable because of their overhead requirements. In this paper, we propose to use location information, a physical property associated with each node, hard to falsify, and not reliant on cryptography, as the basis for (1) detecting spoofing attacks; (2) determining the number of attackers when multiple adversaries masquerading as a same node identity; and (3) localizing multiple adversaries. We formulate the problem of determining the number of attackers as a multi-class detection problem. We first propose two cluster-based mechanisms to determine the number of attackers. We then develop SILENCE that employs the minimum distance testing of RSS values in addition to cluster analysis and can achieve better accuracy than other methods under study that merely use cluster analysis alone. We further developed an integrated detection and localization system that can localize the positions of multiple attackers. We evaluated our techniques through two testbeds using both an 802.11 (WiFi) network and an 802.15.4 (ZigBee) network in two real office buildings. Our experimental results show that SILENCE can achieve over 90% Hit Rate and Precision when determining the number of attackers. Additionally, our localization results using a representative set of algorithms provide strong evidence of high accuracy of localizing multiple adversaries.**

I. Introduction

Due to the openness of the wireless transmission medium, adversaries can monitor any transmission. Further, adversaries can easily purchase low-cost wireless devices and use these commonly-available platforms to launch a variety of attacks with little effort. Among various types of attacks, identity-based spoofing attacks are especially easy to launch and can cause significant damage to network performance. For instance, in an 802.11 network, it is easy for an attacker to gather useful MAC address information during passive monitoring and then modify its MAC address by simply issuing an *ifconfig* command to masquerade as another device. In spite of existing 802.11 security techniques including Wired Equivalent Privacy (WEP), WiFi Protected Access (WPA), or 802.11i (WPA2), such methodology can only protect data frames - an attacker can still spoof management or control frames to cause significant impact on networks.

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, under a malicious attack, multiple adversaries may masquerade as the same identity and collaborate to launch a denial-of-service attack quickly. Therefore, it is important to (1) detect the presence of spoofing attacks, (2) determine the number of attackers, and (3) localize multiple adversaries and eliminate them.

Most existing approaches to address potential spoofing attacks employ cryptographic schemes [5], [6], which require key management overhead. In this work, we propose to use location information, a physical property associated with each node that is hard to falsify and not reliant on cryptography as the basis for detecting spoofing attacks. Since we are concerned with attackers who have different locations than legitimate wireless nodes, utilizing location information to address spoofing attacks has the unique power to not only identify the presence of these attacks but also localize adversaries. An added advantage of employing location to detect spoofing attacks is that it will not require any additional cost or modification to the wireless devices themselves.

The works that are closely related to us are [3], [7], [8]. [3] proposed the use of matching rules of signalprints for spoofing detection, [7] modeled the RSS readings using a Gaussian mixture model and [8] used RSS and K-means cluster analysis to detect 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 localize multiple adversaries after attack detection. Although [8] 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 our work are: (1) GADE: a Generalized Attack Detection ModEl that can both detect spoofing attacks as well as determine the number of adversaries using cluster analysis methods based on 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. We then formulate the problem of determining the number of attackers as a multi-class detection problem. We proposed two cluster

analysis methods, namely Silhouette Plot and System Evolution. We further developed a mechanism called SILENCE for testing SILhouette Plot and System EvolutioN with Minimum DistanCE of clusters, which performs the minimum distance testing in addition to the cluster analysis to improve the accuracy of determining the number of attackers. The key advantage of our approach is that our detector can determine the number of attackers, which is a challenging problem that has not been addressed by prior work.

Moreover, we developed an integrated system, IDOL, which utilizes the results of the number of attackers returned by SILENCE to further localize multiple adversaries. As we demonstrated through our experiments using both an 802.11 networks as well as an 802.15.4 network in two real office building environments, GADE is highly effective in spoofing detection - SILENCE can accurately determine the number of attackers with over 90% hit rates and precision. Further, using a set of representative localization algorithms, we show that IDOL can achieve similar localization accuracy when localizing adversaries to that of under normal conditions. One key observation is that IDOL can handle attackers using different transmission power levels, thereby providing strong evidence of the effectiveness of localizing multiple adversaries when there are multiple attackers in the network.

The rest of the paper is organized as follows. We place our work in the context of related research in Section II. We present GADE, our generalized attack detection model in Section III. We formulate the problem of determining the number of attackers using multi-class detection and propose our cluster-analysis based mechanisms in Section IV. In Section V, we present IDOL, the integrated detection and localization system. Finally, we conclude our work in Section VI.

II. Related Work

The traditional approach to prevent spoofing attacks is to use cryptographic-based authentication [5], [6]. As it is not always desirable to use authentication due to limited resources on nodes and infrastructural overhead involved, recently new approaches utilizing physical properties associated with wireless transmission have been proposed. [4] has introduced a security layer that is separate from conventional network authentication methods using forge-resistant relationships based on the packet traffic. [9] utilizes properties of the wireless channel to support security objectives. The works that are most closely related to us are [3], [7]: one proposed the use of matching rules of signalprints for spoofing detection, and the other modeled the RSS readings using a Gaussian mixture model. 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 don't have the ability to localize the positions of the adversaries after attack detection.

Turning to studying localization techniques, in spite of its several meter-level accuracy, using RSS [10]–[13] 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

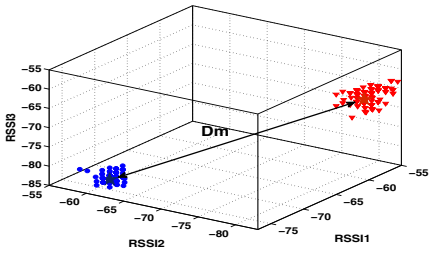


Fig. 1. Illustration of RSS readings from two physical locations.

of various physical properties such as RSS [10], [11], Time Of Arrival (TOA) [14] and Time Difference Of Arrival (TDOA). Range-free algorithms [15] use coarser metrics to place bounds on candidate positions. Another method of classification describes the strategy used to map a node to a location. Lateration approaches [14], use distances to landmarks, while angulation uses the angles from landmarks. Scene matching strategies [10] use a function that maps observed radio properties to locations on a pre-constructed signal map or database.

In this work, we focus on static nodes, which are common for spoofing scenarios [7]. We will address spoofing detection in mobile environments in future work. Our work differs from the previous work in that we use the location information to assist in attack detection instead of relying on cryptographic-based approaches. Further, our work is novel because none of the exiting work can determine the number of attackers when there are multiple adversaries masquerading as the same identity. Moreover, our approach can accurately localize multiple adversaries even when the attackers varying their transmission power levels to trick the system of their true locations.

III. GADE: Generalized Attack Detection ModEl

In this section, we present a Generalized Attack Detection ModEl (GADE), which consists of two phases: *attack detection*, which detects the presence of an attack, and *number determination*, which determines the number of adversaries and will be presented in the next section.

A. Cluster Analysis for Attack Detection

We formulate spoofing detection as a statistical significance testing problem, where the null hypothesis is $\mathcal{H}_0$: normal (no attack). In significance testing, a test statistic $\mathbf{T}$ is used to evaluate whether the observed data belongs to the null-hypothesis or not. For a detailed description of using significance testing for spoofing detection, please refer to our previous work [8].

The challenge in spoofing detection is to devise strategies that use the uniqueness of location, but not using location directly as the attackers' positions are unknown. We propose to use received signal strength (RSS), a property closely correlated with location in physical space and is readily available in the existing wireless networks. We define the RSS value vector as $\mathbf{s} = \{s_1, s_2, ...s_n\}$ where n is the number of landmarks/access points (APs) that are monitoring the RSS of the wireless nodes and know their locations. Since a RSS vector presents strong spatial relations, each RSS reading vector

corresponds to a point in a n-dimensional signal space [16], and 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 should form different clusters in signal space. We assume that when a spoofing attack is conducted, the victim node is also present in the same management domain of the network. Figure 1 shows RSS reading vectors of three landmarks from two different physical locations. Clearly, two clusters are formed. Based on this observation, we propose to use cluster analysis as our theoretic foundation to perform spoofing detection.

The Partitioning Around Medoids (PAM) Method [17] is a popular iterative descent clustering algorithm. The PAM method arbitrarily chooses K sample points as the initial medoids if we partition the data set into K clusters. It then subsequently swaps new sample points as new medoids to reduce the cost of the objective function, which is the sum of the dissimilarities of all the sample points to their nearest medoid:

$$J_{min} = \sum_{j=1}^{K} \sum_{\mathbf{s_n} \in C_j} \|\mathbf{s_n} - \mathbf{M_j}\|^2, \quad (1)$$

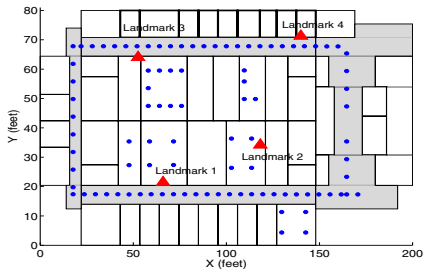
where $\mathbf{s_n}$ is a RSS vector representing the nth sample point and M_j is the sample point that is chosen as the medoid for the jth cluster C_j in signal space. Compared to the popular K-means method [8], the PAM method is more robust in the presence of noise and outliers. Thus, the PAM method is more suitable in determining clusters from RSS streams, which can be unreliable and varying with time due to random noise and environmental bias [18].

In our attack detection phase, we partition the RSS vectors from the same node identity into 2 clusters (i.e.$K = 2$) no matter how many attackers are using this identity, since our objective is to detect the presence of attacks. We then choose the distance between two medoids D_m as the test statistic T in our significance testing for spoofing detection, $D_m = ||\mathbf{M_i} - \mathbf{M_j}||$, where M_i and M_j are the medoids of two clusters. We note that different from most existing work, e.g., [7], our spoofing detectors do not assume RSS readings follow any distributions (e.g. Gaussian distribution). Under normal conditions, the test statistic D_m should be small since there is basically only one cluster from a single physical location. However, under a spoofing attack, there is more than one node at different physical locations claiming the same node identity. As a result, more than one clusters will be formed in the signal space and D_m will be large as the medoids are derived from the different RSS clusters associated with different locations in physical space.

B. Experimental Methodology

We next present our experimental methodology. We conducted experiments in two office buildings: one is the Wireless Information Network Laboratory (WINLAB) using an 802.11 (WiFi) network and the other is the Computer Science Department at Rutgers University using an 802.15.4 (ZigBee) network as presented in Figure 2. The size of these two floors are 219x169ft and 200x80ft respectively. Figure 2 (a) shows 5 landmarks in red stars in the 802.11 networks, whereas there

(a) 802.11 network



(b) 802.15.4 network

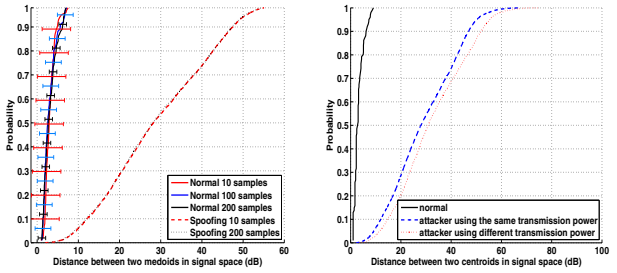
Fig. 2. Landmark setups and testing locations in two networks within two office buildings.

are 4 landmarks deployed as red triangles in the 802.15.4 network shown in Figure 2 (b).

The small dots in the floor maps are the locations used for testing. There are 101 locations for the 802.11 network and 94 locations for the 802.15.4 network. At each location, 300 packet-level RSS samples are collected. Further, to evaluate the robustness of our approach in handling attacks using different transmission power levels, we collected packets at varying transmission power levels from $30mW$ ($15dBm$) to $1mW$ ($0dBm$) for the 802.11 network. We randomly chose point combinations on the floor and treated one point as the position of the original node, and the rest as the positions of the spoofing nodes. Then, we ran tests through all the possible combinations of testing points for cases of 2, 3, and 4 attackers masquerading as a single node identity. In addition, we built an integrated system to both detect attacks as well as localize the positions of adversaries. We use the leave-one-out method in localization algorithms, which means we choose one location as the testing node whereas the rest of the locations as training data. The experimental results will be presented in the following sections respectively.

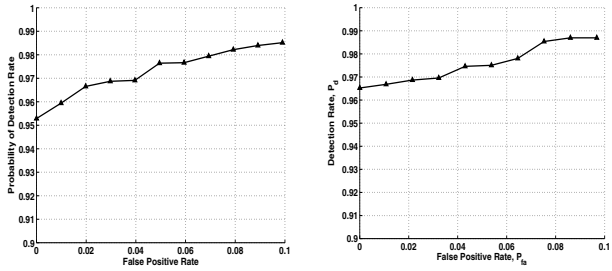
C. Evaluation of attack detection

Impact of Threshold and Sampling Number: The thresholds of test statistics define the critical region for the significance testing. Appropriately setting a threshold τ enables the attack detector to be robust to false detections. Figure 3 shows the Cumulative Distribution Function (CDF) of D_m in signal space under both normal conditions as well as with spoofing attacks. We observed that the curve of D_m shifted greatly to the right under spoofing attacks. Thus, when $D_m > \tau$, we can declare the presence of a spoofing attack. The short lines across the CDF lines are the averaged variances of D_m under different sampling numbers. We observed that the CDF curves of different sampling numbers are almost



(a) Sampling number study (b) Different transmission power

Fig. 3. 802.11 network: Cumulative Distribution Function (CDF) of distance between medoids D_m in signal space.



(a) 802.11 network (b) 802.15.4 network

Fig. 4. Receiver Operating Characteristic (ROC) curves when using PAM method to perform attack detection.

mixed together, which indicate that for a given threshold τ similar detection rate will be achieved under different sampling numbers. However, the averaged variance decreases with the increasing number of samples - the short-term RSS samples is not as stable as the long-term RSS samples. The more stable the D_m is, the more robust the detection mechanism can be. Therefore, there is a trade off between the number of RSS samples needed to perform spoofing detection and the time the system can declare the presence of an attack. For this study we use 200 RSS samples, which has a variance of $0.84dB$.

Handling Different Transmission Power Levels: If a spoofing attacker sends packets at a different transmission power level from the original node, based on our cluster analysis there will be two distinct RSS clusters in signal space (i.e., D_m will be large). We varied transmission power for an attacker from $30mW$ ($15dBm$) to $1mW$ ($0dBm$). We found that in all cases D_m is larger than normal conditions. Figure 3 (b) presents an example of the Cumulative Distribution Function (CDF) of the D_m for the 802.11 network when the spoofing attacker used transmission power of 10dB to send packets, whereas the original node used 15dB transmission power level. We observed that the curve of D_m under the different transmission power level shifts to the right indicating larger D_m values. Thus, spoofing attacks launched by using different transmission power levels will be detected effectively in GADE.

Performance of Detection: To evaluate the effectiveness of using cluster analysis for attack detection, Figure 4 presents the ROC curves of using D_m as a test statistic to perform attack detection for both the 802.11 and the 802.15.4 networks. The results are encouraging, showing that for false positive rates less than 10%, the detection rate are above 98% when the threshold τ is around $10dB$. Even when the false positive rate goes to zero, the detection rate is still more than 95% for both networks. Further, we obtained similar results to [8] when studying the effects on the detection rate by varying the distance between the spoofing node and the original node. This indicates that GADE is highly effective in detecting the presence of an attack.

IV. Determining the Number of Attackers

After detecting the presence of a attack, the next phase is to determine the number of attackers, using the same node identity to launch spoofing attacks, so that we can further localize the multiple adversaries and eliminate them. We first describe how to measure the accuracy when determining the number attackers. We then present two mechanisms, *Silhouette Plot* and *System Evolution*. Next, we describe the $SILENCE$ mechanism that we developed, which performs an evaluation based on minimum cluster distance on top of the cluster analysis to improve the accuracy of determining the number of attackers.

A. Effectiveness

Inaccurate estimation of the number of attackers will cause failure in localizing 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 multi-class detection problem and is similar to determining how many clusters exist in the RSS readings. If C is the set of all classes, i.e., all possible combination of number of attackers. For instance, $C = \{1, 2, 3, 4\}$. For a class of specific number of attackers c_i, e.g., $c_i = 3$, we define P_i as the positive class of c_i and all other classes (i.e., all other number of attackers) as negative class N_i:

$$P_i = c_i, \tag{2}$$

$$N_i = \bigcup_{j \neq i} c_j \in C. \tag{3}$$

Further, we are interested in the statistical characterization of the percentage that the number of attackers can be accurately determined over all possible testing attempts with mixed number of attackers. Associated with a specific number of attackers, i, we define the Hit Rate HR_i as $HR_i = \frac{N_{true}}{P_i}$ where N_{true} is the true positive detection of class c_i. Let N_{false} be the false detection of the class c_i out of the negative class N_i that do not have i number of attackers. We then define the false positive rate FP_i for a specific number of attackers of class c_i as $FP_i = \frac{N_{false}}{N_i}$. Then the Precision is defined as:

$$Precision_i = \frac{N_{true}}{N_{true} + N_{false}}. \tag{4}$$

F-measure: F-measure is originated from information retrieval and measures the accuracy of a test by considering both the Hit Rate and the Precision [19]:

$$F - measure_i = \frac{2}{\frac{1}{Precision_i} + \frac{1}{HitRate_i}}. \tag{5}$$

Multi-class ROC graph: We further use the multi-class ROC graph to measure the effectiveness of our mechanisms. Particularly, we use two methods [20]: $class -$

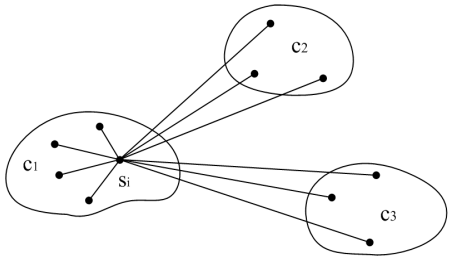


Fig. 5. Illustration of the construction of Silhouettes, $K = 3, j = 1$.

$reference\ \ based$ and $benefit - error\ \ based$. The class-reference based formulation produces C different ROC curves when handling C classes based on P_i and N_i. Further, in the C-class detection problem the traditional 2x2 confusion matrix, including True Positives, False Positives, False Negatives, and True Negatives, becomes an CxC matrix, which contains the C benefits (true positives) and $C^2 - C$ possible errors (false positives). The benefit-error based method is based on the CxC matrix. For example, when $C = 3$ with possible number of attackers of $\{2, 3, 4\}$, the benefits are 3 and the possible errors are 6.

B. Silhouette Plot

1) Attacker Number Determination: A Silhouette Plot is a graphical representation of a cluster [21]. To determine the number of attackers, we construct Silhouettes in the following way: the RSS sample points $\mathbf{S} = \{\mathbf{s_1}, ..., \mathbf{s_N}\}$ (with N as the total number of samples) are the data set and we let $C = (c_1, ..., c_K)$ be its clustering into K clusters. Let $d(\mathbf{s}_k, \mathbf{s}_l)$ be the distance between $\mathbf{s}_k$ and $\mathbf{s}_l$. Let $c_j = \{\mathbf{s}_1^j, ..., \mathbf{s}_{mj}^j\}$ be the j-th cluster, $j = 1, ..., K$, where $m_j = |c_j|$. The average distance a_i^j between the i-th RSS vector in the cluster c_j and the other RSS vectors in the same cluster is thus given by:

$$a_i^j = \frac{1}{m_j - 1} \sum_{\substack{k=1 \\ k \neq i}}^{m_j} d(\mathbf{s}_i^j, \mathbf{s}_k^j), \quad i = 1, ..., m_j. \tag{6}$$

Further, the minimum average distance between the i-th RSS vector in the cluster c_j and all the RSS vectors clustered in the clusters c_k, $k = 1, ..., K, k \neq j$ is given by:

$$b_i^j = \min_{\substack{n=1,...K \\ n \neq j}} \{\frac{1}{m_n} \sum_{k=1}^{m_n} d(\mathbf{s}_i^j, \mathbf{s}_k^n)\}, \quad i = 1, ..., m_j. \tag{7}$$

Then the silhouette width of the i-th RSS vector in the cluster c_j is defined as:

$$w_i^j = \frac{b_i^j - a_i^j}{\max\{a_i^j, b_i^j\}}. \tag{8}$$

From the Equation (8), it follows that $-1 \leq w_i^j \leq 1$. We can now define the silhouette of the cluster c_j:

$$W_j = \frac{1}{m_j} \sum_{i=1}^{m_j} w_i^j. \tag{9}$$

Hence, the global Silhouette index for partition p that partitions the data set into K clusters is given by:

$$W(K)_p = \frac{1}{K} \sum_{j=1}^{K} w_j. \tag{10}$$

Number of Attackers	2	3	4
802.11 network, Hit Rate	99.59%	89.81%	80.52%
802.11 network, Precision	91.85%	87.29%	99.33%
802.11 network, F-measure	95.56%	88.53%	88.94%
802.15.4 network, Hit Rate	99.46%	91.05%	83.77%
802.15.4 network, Precision	93.22%	85.71%	99.67%
802.15.4 network, F-measure	96.24%	88.30%	91.03%

TABLE I
SILHOUETTE PLOT: HIT RATE, PRECISION, AND F-MEASURE OF DETERMINING THE NUMBER OF ATTACKERS

Finally, we define Silhouette Coefficient SC to determine the number of attackers:

$$SC = \max_K W(K)_p. \tag{11}$$

SC is used for the selection of a "best" value of the cluster number K (i.e., the optimal number of attackers) by choosing the K to make $W(K)$ as high as possible across all partitions. Since the objective of constructing silhouettes is to obtain SC, we note that there are no adjustable parameters in this detection scheme.

2) Experimental Evaluation: Table I presents experimental values of Hit Rate, Precision, and F-measure when the attacker number $i = \{2, 3, 4\}$ for both the 802.11 and the 802.15.4 networks. We observed that the performance of Silhouette Plot in both networks are qualitatively the same. We found that when the number of attackers equals to 2, i.e., 2 attackers masquerading the same identity in the network, the Silhouette Plot achieves both the highest Hit Rate, above 99%, and the highest F-measure value, over 95%. Further, the case of 4 attackers achieves the highest Precision above 99%, which indicates that the detection of the number of attackers is more accurate, however, the Hit Rate decreases to about 80%. Moreover, the Precision of the case of 3 attackers is lower than the cases of 2 and 4 attackers. This is because the cases of 2 attackers and 4 attackers are likely to be mistakenly determined as the case of 3 attackers. In general, our observation indicates that the Hit Rate decreases as the number of attackers increases. However, when the number of attackers increases, the adversaries also increase the probability to expose themselves. In the rest of our study we will only present the results up to 4 attackers that masquerade the same node identity simultaneously.

C. System Evolution

1) Attacker Number Determination: The System Evolution is a new method to analyze cluster structures and estimate the number of clusters [22]. The System Evolution method uses the *twin-cluster* model, which are the two closest clusters (e.g. clusters a and b) among K potential clusters of a data set. The twin-cluster model is used for energy calculation. The Partition Energy $E_p(K)$ denotes the border distance between the twin clusters, whereas the Merging Energy $E_m(K)$ is calculated as the average distance between elements in the border region of the twin clusters. The border region includes a number of sample points chosen from clusters a and b that are closer to its twin cluster than any other points within its own cluster. For instance, if cluster a contains total M_a sample points, in the twin-cluster model, a will be partitioned into $D_a = \frac{\sqrt{M_a}}{2}$ parts. Then the number of sample points in the border region

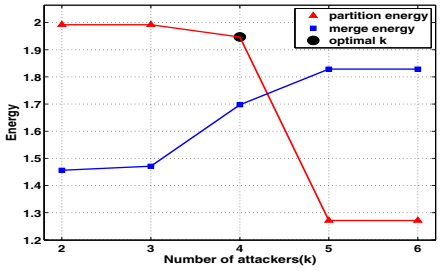


Fig. 6. System Evolution: detection of 4 adversaries masquerading the same node identity.

is defined as $n_a = \frac{M_a}{D_a}$. The same rule is carried out for its twin cluster b. Thus we compute the Partition Energy $E_p(K)$ as:

$$E_p(K) = \frac{1}{n_a + n_b}\{\sum_{i=1}^{n_a} \min_{j=1,..n_b} D(a_i, b_j) + \sum_{j=1}^{n_b} \min_{i=1,..n_a} D(a_i, b_j)\}, \quad (12)$$

and the Merging Energy $E_m(K)$ as:

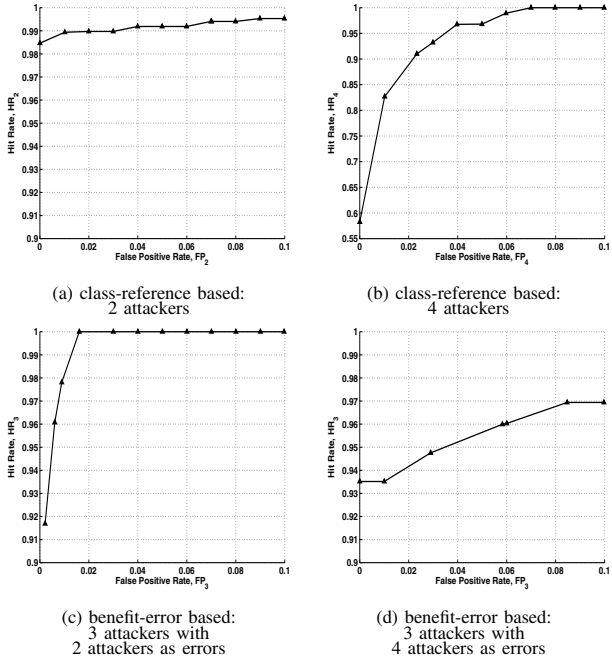
$$E_m(K) = \frac{1}{\binom{n_a+n_b}{2}} \sum_{i=1}^{(n_a+n_b-1)} \sum_{j=i+1}^{(n_a+n_b)} D(\mathbf{s}_i, \mathbf{s}_j), \quad (13)$$

where $D(a_i, b_j)$ is the Euclidean/Pearson distance between the elements a_i and b_j in clusters a and b respectively. And $\mathbf{s}_i, \mathbf{s}_j \in \{a_i\} \bigcup \{b_j\}$, which are the elements in the border region of the twin clusters.

The basic idea behind using the System Evolution method to determine the number of attackers is that all the rest of clusters are separated if the twin clusters are separable. Starting from the initial state with $K = 2$, the algorithm works with PAM by changing the number of clusters in a data set through the partitioning process $E_p(K) > E_m(K)$ and the merging process $E_m(K) \geq E_p(K)$ alternatively. The algorithm stops when it reaches a equilibrium state $K_{optimal}$, at which the optimal number of clusters is found in the data set: $K_{optimal} = K$, if $E_p(K) > E_m(K)$ and $E_p(K+1) \leq E_m(K+1)$.

Figure 6 presents an example of using the System Evolution method to determine the number of attackers in the 802.11 network. It shows the energy calculation vs. the number of clusters. The $K_{optimal}$ is obtained when $K = 4$ with $E_p(4) > E_m(4)$ and $E_p(5) < E_m(5)$ indicating that there are 4 adversaries in the network using the same identity to perform spoofing attacks.

2) Experimental Evaluation: In this section, we show our study of System Evolution using multi-class ROC graphs. We perform threshold τ' testing on $E_p(K) - E_m(K)$. We can then obtain the number of attackers $K_{optimal}$ based on: $K_{optimal} = K$, if $E_p(K) - E_m(K) > \tau'$ and $E_p(K+1) - E_m(K+1) \leq \tau'$. Figure 7 presents the multi-class ROC graphs using both class-reference based method (i.e., the cases of 2 and 4 attackers) and benefit-error based method (i.e., the case of 3 attackers) by varying the threshold τ'. Because of the overall higher Hit Rate under the 802.15.4 network, we only present the results of the 802.11 network in Figure 7. By using the class-reference



(a) class-reference based: 2 attackers

(b) class-reference based: 4 attackers

(c) benefit-error based: 3 attackers with 2 attackers as errors

(d) benefit-error based: 3 attackers with 4 attackers as errors

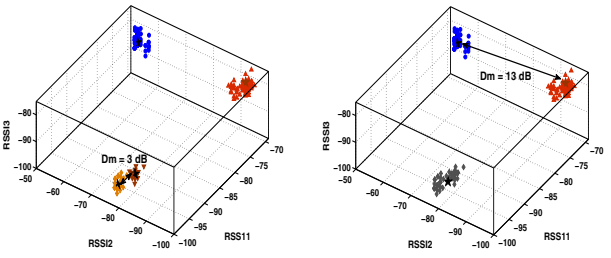
Fig. 7. System Evolution in 802.11 network: Multi-class Receiver Operating Characteristic (ROC) Graphs of Hit Rate vs. False Positive.

based method, in Figure 7 (a) and (b), we observed better performance of Hit Rate under the case of 2 attackers than the case of 4 attackers when the False Positive Rate decreases. Turning to examine the ROC graphs of the case of 3 attackers by using the benefit-error based method as shown in Figure 7 (c) and (d), we found that bounded by less than 10% False Positive Rate, the Hit Rate is lower when treating 4 attackers as errors than treating 2 attackers as errors. This indicates that the probability of misclassifying 3 attackers as 4 attackers is higher than that of misclassifying 3 attackers as 2 attackers.

D. The SILENCE Mechanism

The advantage of Silhouette Plot is that it is suitable for estimating the best partition. Whereas the System Evolution method performs well under difficult cases such as when there exists slightly overlapping between clusters and there are smaller clusters near larger clusters [22]. However, we observed that for both Silhouette Plot and System Evolution methods, the Hit Rate decreases as the number of attackers increases, although the Precision increases. This is because the clustering algorithms can not tell the difference between real RSS clusters formed by attackers at different positions and fake RSS clusters caused by outliers and variations of the signal strength. Figure 8 illustrates such a situation where there are 3 attackers masquerading the same identity. Silhouette Plot returns the number of attackers $K_{sp} = 4$ as shown in Figure 8 (a). We found that the minimum distance between two clusters in Silhouette Plot is very small because two clusters are actually from a single physical location. Further, Figure 8 (b) shows that System Evolution returns the number of attackers $K_{se} = 3$, the correct number of attackers, and the minimum distance between two clusters is large indicating that the clusters are from different physical locations.

Based on this observation, we developed $SILENCE$, test-



(a) Silhouette Plot (b) System Evolution

Fig. 8. Illustration of the minimum cluster distance using cluster analysis methods under the case of 3 attackers

Number of Attackers	2	3	4
802.11 network, Hit Rate	99.67%	98.21%	90.06%
802.11 network, Precision	98.86%	91.42%	99.72%
802.11 network, F-measure	99.27%	94.69%	94.64%
802.15.4 network, Hit Rate	99.93%	96.04%	87.80%
802.15.4 network, Precision	96.99%	89.04%	99.96%
802.15.4 network, F-measure	98.44%	92.41%	93.49%

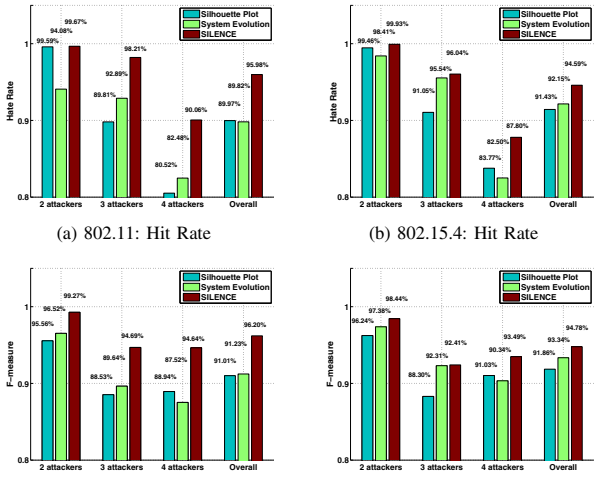
TABLE II
SILENCE: HIT RATE, PRECISION, AND F-MEASURE OF DETERMINING THE NUMBER OF ATTACKERS

ing SILhouette Plot and System EvolutioN with minimum distanCE of cluster, which evaluates the minimum distance between clusters on top of the pure cluster analysis to improve the accuracy of determining the number of attackers. The number of attackers K in SILENCE is thus determined by:

$$K = \begin{cases} K_{sp} & \text{if } K_{sp} = K_{se}; \\ K_{sp} & \text{if } \min(D_m^{obs})_{K_{sp}} > \min(D_m^{obs})_{K_{se}}; \\ K_{se} & \text{if } \min(D_m^{obs})_{K_{sp}} < \min(D_m^{obs})_{K_{se}}, \end{cases} \quad (14)$$

where D_m^{obs} is the observed value of D_m between two clusters. SILENCE takes the advantage of both Silhouette Plot and System Evolution and further makes the judgment by checking the minimum distance between the returned clusters to make sure the clusters are produced by attackers instead of RSS variations and outliers. Hence, when applying SILENCE to the case shown in Figure 8, SILENCE returns $K = 3$ as the number of attackers, which is the true positive in this example.

1) Experimental Evaluation: The effectiveness of using SILENCE to determine the number of attackers is presented in Table II. And Figure 9 presents the comparison of Hit Rate and F-measure of SILENCE to those of Silhouette Plot and System Evolution methods. The key observation is that there is a significant increase of Hit Rate for all the cases of the number of attackers under study. In particular, for the 802.11 network, the Hit Rate has increased from $89\% \sim 92\%$ in Silhouette Plot and System Evolution to 98% using SILENCE for the case of 3 attackers and from $80\% \sim 82\%$ to 90% for the 4 attackers case. Whereas for the 802.15.4 network, the Hit Rate has increased from around $91\% \sim 95\%$ to 96% in SILENCE for the case of 3 attackers and from 84% to 88% for the 4 attackers case. Further, We observed that SILENCE has better performance over all the 2, 3 and 4 attackers in terms of F-measure. The overall improvement of F-measure is from 91% to 96% for 802.11 network, and from $92\% \sim 93\%$ to 95% for 802.15.4 network. Further, comparing with Silhouette Plot and System Evolution, the computational cost of SILENCE does not increase much. We experienced that



(a) 802.11: Hit Rate (b) 802.15.4: Hit Rate

(d) 802.11: F-measure (d) 802.15.4: F-measure

Fig. 9. Hit Rate and F-measure comparison of SILENCE to methods using cluster analysis alone such as Silhouette and System Evolution.

SILENCE can determine the number of attackers within one second for each experimental run. These results demonstrate that SILENCE, a mechanism that combines minimum distance testing and cluster analysis together to perform multi-class attacker detection, is more effective than using techniques based on cluster analysis alone.

V. IDOL: INTEGRATED DETECTION AND LOCALIZATION FRAMEWORK

In this section we present our integrated system that can both detect spoofing attacks, determine the number of attackers, and localize multiple adversaries. The experimental results are presented to evaluate the effectiveness of our approach, especially when attackers using different transmission power levels.

A. Framework

The traditional localization approaches are based on averaged RSS from each node identity inputs to estimate the position of a node. However, 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. The traditional method of averaging RSS readings cannot differentiate RSS readings from different locations and thus is not feasible for localizing adversaries.

Different from traditional localization approaches, our integrated detection and localization system utilizes the RSS medoids returned from SILENCE as inputs to localization algorithms to estimate the positions of adversaries. The return positions from our system includes the location estimate of the original node and the attackers in the physical space.

Handling adversaries using different transmission power levels: An adversary may vary the transmission power levels when performing spoofing attacks so that the localization system cannot estimate its location accurately. We examine the pass loss equation that models the received power as a

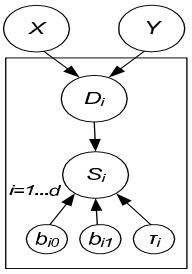


Fig. 10. Bayesian graphical model in our study.

function of the distance to the landmark:

$$P(d)[dBm] = P(d_0)[dBm] - 10\gamma \log_{10}\left(\frac{d}{d_0}\right), \qquad (15)$$

where $P(d_0)$ represents the transmitting power of a node at the reference distance d_0, d is the distance between the transmitting node and the landmark, and γ is the path loss exponent. Further, we can express the difference of the received power between two landmarks, i and j, as:

$$P(d_i) - P(d_j) = 10\gamma_i \log_{10}\left(\frac{d_i}{d_0}\right) - 10\gamma_j \log_{10}\left(\frac{d_j}{d_0}\right). \qquad (16)$$

Based on Equation (16), we found that the difference of the corresponding received power between two different landmarks is independent of the transmission power levels. Thus, when an adversary residing at a physical location varies its transmission power to perform a spoofing attack, the difference of the RSS readings between two different landmarks from the adversary is a constant since the RSS readings are obtained from a single physical location. We can then utilize the difference of the medoids vectors in signal space obtained from SILENCE to localize adversaries.

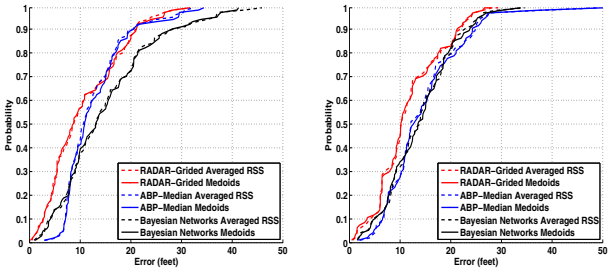
B. Algorithms

In order to evaluate the generality of IDOL for localizing adversaries, we have chosen a set of representative localization algorithms ranging from nearest neighbor matching in signal space (RADAR [10]), to probability-based (Area-Based Probability [11]), and to multilateration (Bayesian Networks [23]).

RADAR-Gridded: The RADAR-Gridded algorithm is a scene-matching localization algorithm extended from [10]. 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.

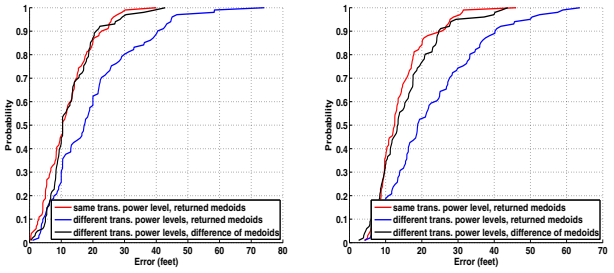
Area Based Probability (ABP): ABP also utilizes an interpolated signal map [11]. Further, the experimental area is divided into a regular grid of equal sized tiles. ABP assumes the distribution of RSS for each landmark follows a Gaussian distribution with mean as the expected value of RSS reading vector $\mathbf{s}$. ABP then computes the probability of the wireless device being at each tile L_i, with $i = 1...L$, on the floor using Bayes' rule:

$$P(L_i|\mathbf{s}) = \frac{P(\mathbf{s}|L_i) \times P(L_i)}{P(\mathbf{s})} \qquad (17)$$



(a) 802.11 network (b) 802.15.4 network

Fig. 11. Comparison of localization errors between using medoids from cluster analysis and using averaged RSS.



(a) RADAR-Grided (b) ABP-Median

Fig. 12. Localization errors when adversaries using different transmission power levels.

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

Bayesian Networks (BN): BN localization is a multilateration algorithm that encodes the signal-to-distance propagation model into the Bayesian Graphical Model for localization [23]. Figure 10 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 τ_i: $s_i \sim N(b_{0i} + b_{1i} \log D_i, \tau_i)$. Through Markov Chain Monte Carlo (MCMC) simulation, BN returns the sampling distribution of the possible location of X and Y as the localization result.

C. Experimental Evaluation

Figure 11 presents the localization error CDF when using the returned RSS medoids from SILENCE and the averaged RSS respectively for RADAR-Gridded, ABP, and Bayesian Networks in two networks. We observed similar localization performance when using the returned RSS medoids to the traditional approaches using averaged RSS. Further, Figure 12 presents the CDF of localization error of RADAR-Gridded and ABP when adversaries using different transmission power levels. To evaluate the performance of our approach by using the difference of returned medoids, three cases are presented in Figure 12: (1) Adversaries used the same transmission

power levels as the original node and the returned medoids are used; (2) Adversaries changed their transmission power level from $15dB$ to $10dB$ and the returned medoids are used; and (3) Adversaries changed their transmission power level from $15dB$ to $10dB$ and the difference of returned medoids are used. The key observation from Figure 12 is that the performance of using the difference of returned medoids in handling adversaries using different transmission power levels is comparable to the results when adversaries used the same transmission power levels as the original node. Further, the localization performance is much worse than the traditional approaches if the difference of returned medoids is not used when localizing adversaries using different transmission power levels, shown as the case 2 above. In particular, when using our approach we can achieve the median error of 13 feet for both RADAR-Grided and ABP in case 3, a $40\% \sim 50\%$ performance improvement, comparing to the median errors of 20 feet and 19 feet for RADAR-Gridded and ABP respectively in case 2. Thus, IDOL is highly effective in localizing multiple adversaries with or without changing their transmission power levels.

VI. Conclusion

In this work we proposed to use location information, a physical property associated with each node, hard to falsify, and not reliant on cryptography, as the basis for detecting spoofing attacks in wireless networks. Our spoofing detectors do not assume the input data follow any distributions. Our approach can both detect 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. Determining the number of adversaries is a particularly challenging problem. We developed 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 .

We conducted experiments on two testbeds through both an 802.11network (WiFi) and an 802.15.4 (ZigBee) network in two real office building environments. We found that our detection mechanisms are highly effective in both detecting the presence of attacks with detection rates over 98% and determining the number of adversaries, achieving over 90% hit rates and precision simultaneously when using SILENCE. Further, based on the number of attackers determined by our mechanisms, our integrated detection and localization system can localize any number of adversaries even when attackers using different transmission power levels. The performance of localizing adversaries achieves similar results as those under normal conditions, thereby, providing strong evidence of the effectiveness of our approach in detecting spoofing attacks, determining the number of attackers and localizing adversaries.

References

[1] J. Bellardo and S. Savage, "802.11 denial-of-service attacks: Real vulnerabilities and practical solutions," in *Proceedings of the USENIX Security Symposium*, 2003, pp. 15 – 28.

[2] F. Ferreri, M. Bernaschi, and L. Valcamonici, "Access points vulnerabilities to dos attacks in 802.11 networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2004.

[3] D. Faria and D. Cheriton, "Detecting identity-based attacks in wireless networks using signalprints," in *Proceedings of the ACM Workshop on Wireless Security (WiSe)*, September 2006.

[4] Q. Li and W. Trappe, "Relationship-based detection of spoofing-related 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.

[5] B. Wu, J. Wu, E. Fernandez, and S. Magliveras, "Secure and efficient key management in mobile ad hoc networks," in *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (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," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, April 2008.

[8] Y. Chen, W. Trappe, and R. P. Martin, "Detecting and localizing wirelss spoofing attacks," in *Proceedings of the Fourth Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, May 2007.

[9] 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.

[10] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, March 2000, pp. 775–784.

[11] E. Elnahrawy, X. Li, and R. P. Martin, "The limits of localization using signal strength: A comparative study," in *Proceedings of the First IEEE International Conference on Sensor and Ad hoc Communcations and Networks (SECON 2004)*, Oct. 2004, pp. 406–414.

[12] Y. Chen, J. Francisco, W. Trappe, and R. P. Martin, "A practical approach to landmark deployment for indoor localization," in *Proceedings of the Third Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*, September 2006.

[13] J. Yang and Y. Chen, "A theoretical analysis of wireless localization using RF-based fingerprint matching," in *Proceedings of the Fourth International Workshop on System Management Techniques, Processes, and Services (SMTPS)*, April 2008.

[14] P. Enge and P. Misra, *Global Positioning System: Signals, Measurements and Performance*. Ganga-Jamuna Pr, 2001.

[15] T. He, C. Huang, B. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes in large scale sensor networks," in *Proceedings of the Ninth Annual ACM International Conference on Mobile Computing and Networking (MobiCom'03)*, 2003.

[16] Y. Chen, K. Kleisouris, X. Li, W. Trappe, and R. P. Martin, "The robustness of localization algorithms to signal strength attacks: a comparative study," in *Proceedings of the International Conference on Distributed Computing in Sensor Systems (DCOSS)*, June 2006, pp. 546–563.

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

[18] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic, "Models and solutions for radio irregularity in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, pp. 221–262, 2006.

[19] C. van Rijsbergen, *Information Retrieval, Second Edition*. Butterworths, 1979.

[20] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters, Elsevier*, vol. 27, pp. 861–874, 2006.

[21] P. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, no. 1, pp. 53–65, November 1987.

[22] K. Wang, "Estimating the number of clusters via system evolution for cluster analysis of gene expression data," Computer Science Department, Xidian University, P.R.China, Technical Report NO. 2007-258, 2007.

[23] D. Madigan, E. Elnahrawy, R. Martin, W. Ju, P. Krishnan, and A. S. Krishnakumar, "Bayesian indoor positioning systems," in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, March 2005, pp. 324–331.