The Robustness of Localization Algorithms to Signal Strength Attacks: A Comparative Study

Yingying Chen, Konstantinos Kleisouris, Xiaoyan Li, Wade Trappe, Richard P. Martin

Department of Computer Science and Wireless Information Network Laboratory
Rutgers University, 110 Frelinghuysen Rd, Piscataway, NJ 08854
{yingche,kkonst,xili,rmartin}@cs.rutgers.edu,
trappe@winlab.rutgers.edu

Abstract. In this paper, we examine several localization algorithms and evaluate their robustness to attacks where an adversary attenuates or amplifies the signal strength at one or more landmarks. We propose several performance metrics that quantify the estimator's precision and error, including Hölder metrics, which quantify the variability in position space for a given variability in signal strength space. We then conduct a trace-driven evaluation of several point-based and area-based algorithms, where we measured their performance as we applied attacks on real data from two different buildings. We found the median error degraded gracefully, with a linear response as a function of the attack strength. We also found that area-based algorithms experienced a decrease and a spatial-shift in the returned area under attack, implying that precision increases though bias is introduced for these schemes. We observed both strong experimental and theoretic evidence that all the algorithms have similar average responses to signal strength attacks.

1 Introduction

Secure localization is important for distributed sensor systems because the position of sensor nodes is a critical input for many sensor network tasks, such as tracking, monitoring and geometric-based routing. However, assuring the validity of localization results is not straight-forward because these algorithms rely on physical measurements that can be affected by non-cryptographic attacks. Although there has been recent research on securing localization, to date there has been no study on the robustness of localization algorithms to physical attacks. In this paper, we investigate the susceptibility of a wide range of signal strength localization algorithms to attacks on the Received Signal Strength (RSS). RSS is an attractive basis for localization because all commodity radio technologies, such as 802.11, 802.15.4, and Bluetooth provide it, and thus the same algorithms can be applied across different platforms. Also, using RSS allows the localization system to reuse the existing communication infrastructure, rather than requiring the additional cost needed to deploy specialized localization infrastructure, such as ceiling-based ultrasound, GPS, or infrared methods.

In this work, we investigate the response of several localization algorithms to unanticipated power losses and gains, i.e. attenuation and amplification attacks. In these attacks, the attacker modifies the RSS of a sensor node or landmark, for example, by placing an absorbing or reflecting material around the node. Specifically, we investigate point-based and area-based RF fingerprinting algorithms, whereby a database of collected RF fingerprints are measured at several landmarks for an initial set of locations. In order to evaluate the robustness of these algorithms, we provide a generalized characterization of the localization problem, and then present several performance metrics suitable for quantifying performance. We present a new family of metrics, which we call Hölder metrics, for quantifying the susceptibility of localization algorithms to perturbations in signal strength readings. We use worst-case and average-case versions of the Hölder metric, which describe the maximum and average variability as a function of changes in the RSS. We then experimentally evaluate the performance of a wide variety of localization algorithms after applying attenuation and amplification attacks to real data measured from two different office buildings.

Using experimentally observed localization performance, we found that the error for a wide variety of algorithms scaled with surprising similarity under attack. The single exception was the Bayesian Networks algorithm, which degraded slower than the others in response to attacks against a single landmark. In addition to our experimental observations, we found a similar average-case response of the algorithms using our Hölder metrics. However, we observed that methods which returned an average of likely positions had less variability and are thus less susceptible than other methods.

We also observed that all algorithms degraded gracefully, experiencing linear scaling in localization error as a function of the amount of loss or gain (in dB) an attack introduced. This observation applied to various statistical descriptions of the error, leading us to conclude that no algorithm “collapses" in response to an attack. This is important because it means that, for all the algorithms we examined, there is no tipping point at which an attacker can cause gross errors. In particular, we found the mean error of most of the algorithms for both buildings scaled between 1.3-1.8 ft/dB when all the landmarks were attenuated simultaneously, and 0.5-0.8 ft/dB when attenuating a single landmark. We also showed experimentally that RSS can be easily attenuated by 15 dB, and that, as a general rule of thumb, very simple signal strength attacks can lead to localization errors of 20-30 ft.

Finally, we conducted a detailed evaluation of area-based algorithms as this family of algorithms return a set of potential locations for the transmitter. Thus, it is possible that these algorithms might return a set with a larger area in response to an attack and could have less precision (or more uncertainty) under attack. However, we found all three of our area-based algorithms shifted the returned areas rather than increased returned area. Further, one of the algorithms, the Area Based Probability (ABP) scheme, significantly shrank the size of the returned area in response to very large changes in signal strength.

The rest of this paper is organized as follows. We first discuss related work in Section 2. Next, in Section 3 we give an overview of the algorithms used in our study and discuss how signal strength attacks can be performed. In Section 4, we provide a formal model of the localization problem as well as introduce the metrics that we use in this paper.

We then examine the performance of the algorithms through an experimental study in Section 5, and discuss the Hölder metrics for these algorithms in Section 6. Finally, we conclude in Section 7.

2 Related Work

In general, localization algorithms can be categorized as: range-based vs. range-free, scene matching, and aggregate or singular. The range-based algorithms involve distance estimation to landmarks using the measurement of various physical properties like RSS [1], Time Of Arrival (TOA) [2] and Time Difference Of Arrival (TDOA) [3]. Rather than use precise physical property measurements, range-free algorithms use coarser metrics like connectivity [4] or hop-counts [5] to landmarks to place bounds on candidate positions. In scene matching approaches, a radio map of the environment is constructed, either by measuring actual samples, using signal propagation models, or some combination of the two. A node then measures a set of radio properties (often just the RSS of a set of landmarks), the *fingerprint*, and attempts to match these to known location(s) on the radio map. These approaches are almost always used in indoor environments because signal propagation is extensively affected by reflection, diffraction and scattering, and thus ranging or simple distance bounds cannot be effectively employed. Matching fingerprints to locations can be cast in statistical terms [6,7], as a machine-learning classifier problem [8], or as a clustering problem [9]. Finally, a third dimension of classification extends to aggregate or singular algorithms. Aggregate approaches use collections of many nodes in the network in order to localize (often by flooding), while localization of a node in singular methods only requires it to communicate to a few landmarks. For example, algorithms using optimization [10] or multidimensional scaling [4] require many estimates between nodes.

Recently, it has been recognized that there are many non-cryptographic attacks that can affect localization performance. For example, wormhole attacks tunnel through a faster channel to shorten the observed distance between two nodes [11]. Compromised nodes may delay response messages to disrupt distance estimation [12] and compromised landmarks may even broadcast completely invalid information [13]. Physical barriers can directly distort the physical property used by localization. [12] provided a thorough survey of potential attacks to various localization algorithms based on their underlying physical properties.

Secure localization algorithms have been proposed to address these attacks. [14] uses a distance bounding protocol [15,16] to upperbound the distance between two nodes. Location estimation (via multilateration) with distances from the bounding protocol can be verified against these bounds and any inconsistency will then indicate attack. [17] uses hidden and mobile base stations to localize and verify location estimate. Since such base station locations are hard for attackers to infer, it is hard to launch an attack, thereby providing extra security. [18] uses both directional antenna and distance bounding to achieve security. Compared to all these methods, which employ location verification and discard location estimate that indicates under attack, [13] and [12] try to eliminate the effect of attack and still provide good localization. [12] makes use of the data redundancy and robust statistical methods to achieve reliable localization in the presence of attacks. [13] proposes to detect attacks based on data inconsistency from received beacons and to use a greedy search or voting algorithm to eliminate the malicious beacon information.

In our work, we focus only on fingerprinting algorithms that use RSS, and provide an investigation into the feasibility of signal strength attacks as well as the susceptibility of fingerprinting algorithms to such attacks. Of previous work, only [12] proposed a possible solution to the fingerprint-based localization, but the susceptibility of different fingerprinting methods was not completely investigated.

3 Algorithms and Signal Strength Attacks

In this paper we are only concerned with localization algorithms that employ signal strength measurements. There are several ways to classify localization schemes that use signal strength: range-based schemes, which explicitly involve the calculation of distances to landmarks; and RF fingerprinting schemes whereby a radio map is constructed using prior measurements, and a device is localized by referencing this radio map. For this study, we focus on indoor localization schemes, and therefore we restrict our attention to RF fingerprinting methods, which have had more success for indoor environments. RF fingerprinting methods can be further broken down into two main categories: point-based methods, and area-based methods.

Point-based methods return an estimated point as a localization result. A primary example of a point-based method is the RADAR scheme [9]. Variations of RADAR, such as Averaged RADAR and Gridded RADAR have been proposed in [19]. On the other hand, area-based algorithms return a *most likely* area in which the true location resides. Two examples of area-based localization algorithms are the Area Based Probability (ABP) method [19] and the Bayesian Networks method [20]. One of the major advantages of area-based methods compared to point-based methods is that they return a region, which has an increased chance of capturing the transmitter's true location.

For this paper, we have selected a representative set of algorithms from each class of RF fingerprinting schemes for conducting our analysis. The algorithms we have selected are presented in Table 1. Although there are a variety of other fingerprinting localization algorithms that may be studied, our results are general and can be applied to other point-based and area-based methods. More details for these algorithms can be found in [9, 19, 20].

To attack signal-strength based localization systems, an adversary must attenuate or amplify the RSS readings. This can be done by applying the attack at the transmitting device, e.g. simply placing foil around the 802.11 card; or by directing the attack at the landmarks. For example, we may steer the lobes and nulls of an antenna to target select landmarks. A broad variety of attenuation attacks can be performed by introducing materials between the landmarks and sensors [12]. We measured the effect of different

Table 1. Algorithms under study

Algorithm	**Abbreviation**	**Description**
Area-Based		
Simple Point Matching	SPM	Maximum likelihood matching of the RSS to an area using thresholds.
Area Based Probability	ABP-α	Bayes rule matching of the RSS to an area probabilistically bounded by the confidence level α%.
Bayesian Network	BN	Returns the most likely area using a Bayesian network approach.
Point-Based		
RADAR	R1	Returns the closest record in the Euclidean distance of signal space.
Averaged RADAR	R2	Returns the average of the top 2 closest records in the signal map.
Gridded RADAR	GR	Applies RADAR using an interpolated grid signal map.
Highest Probability	P1	Applies maximum likelihood estimation to the received signal.
Averaged Highest Probability	P2	Returns the average of the top 2 likelihoods.
Gridded Highest Probability	GP	Applies likelihoods to an interpolated grid signal map.

materials on the RF propagation when inserted between the landmarks and the sensors. Figure 1 shows the experimental results. These materials are easy to access and attacks utilizing these kind of materials can be simply performed with low cost. Based upon the results in Figure 1, we see that there is a linear relationship between the unattacked signal strength and the attacked signal strength in dB for various materials. The linear relationship suggests that there is an easy way for an adversary to control the effect of his/her attack on the observed signal strength.

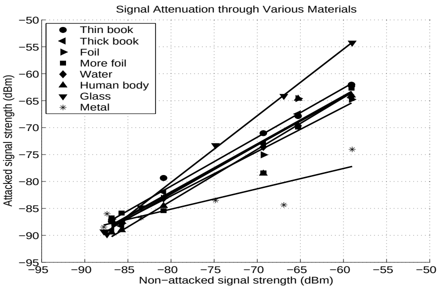


Fig. 1. Signal attenuation when going through a barrier

In the rest of this paper, we will use the linear attenuation model to describe the effect of an attack on the RSS readings at one or more landmarks. The resulting attacked readings are then used to study the consequent effects on localization for the algorithms surveyed above. In particular, in this study, we apply our attacks to individual landmarks, which might correspond to placing a barrier directly in front of a landmark, as well as to the entire set of landmarks, which corresponds to placing a barrier around the transmitting device. Similar arguments can be made for amplification attacks, whereby barriers are removed between the source and receivers. Although there are many different and more complex signal strength attack methods that can be used, we believe their effects will not vary much from the linear signal strength attack model we use in this paper, and note that such sophisticated attacks could involve much higher cost to perform.

4 Measuring Attack Susceptibility

The aim of a localization attack is to perturb a set of signal strength readings in order to have an effect on the localization output. When selecting a localization algorithm, it is desirable to have a set of metrics by which we can quantify how susceptible a localization algorithm is to varying levels of attack by an adversary. In this section, we shall provide a formal specification for an attack, and present several measurement tools for quantifying the effectiveness of an attack.

4.1 A Generalized Localization Model

In order to begin, we need to specify a model that captures a variety of RF-fingerprinting localization algorithms. Let us suppose that we have a domain D in two-dimensions, such as an office building, over which we wish to localize transmitters. Within D, a set of n landmarks have been deployed to assist in localization. A wireless device that transmits with a fixed power in an isotropic manner will cause a vector of n signal strength readings to be measured by the n landmarks. In practice, these n signal strength readings are averaged over a sufficiently large time window to remove statistical variability. Therefore, corresponding to each location in D, there is an n-dimensional vector of signal readings $\mathbf{s} = (s_1, s_2, \cdots, s_n)$ that resides in a range R.

This relationship between positions in D and signal strength vectors defines a fingerprint function $F : D \rightarrow R$ that takes our real world position (x, y) and maps it to a signal strength reading $\mathbf{s}$. F has some important properties. First, in practice, F is not completely specified, but rather a finite set of positions (x_j, y_j) is used for measuring a corresponding set of signal strength vectors $\mathbf{s}_j$. Additionally, the function F is generally one-to-one, but is not onto. This means that the inverse of F is a function G that is not well-defined: There are holes in the n-dimensional space in which R resides for which there is no well-defined inverse.

It is precisely the inverse function G, though, that allows us to perform localization. In general, we will have a signal strength reading $\mathbf{s}$ for which there is no explicit inverse (e.g. perhaps due to noise variability). Instead of using G, which has a domain restricted to R, we consider various pseudo-inverses G_{alg} of F for which the domain of G_{alg} is the complete n-dimensional space. Here, the notation G_{alg} indicates that there may be different *algorithmic* choices for the pseudo-inverse. For example, we shall denote G_R to be the RADAR localization algorithm. In general, the function G_{alg} maps an n-dimensional signal strength vector to a region in D. For point-based localization algorithms, the image of G_{alg} is a single point corresponding to the localization result. On the other hand, for area-based methods, the localization algorithm G_{alg} produces a set of likely positions.

An attack on the localization algorithm is a perturbation to the correct n-dimensional signal strength vector $\mathbf{s}$ to produce a corrupted n-dimensional vector $\tilde{\mathbf{s}}$. Corresponding to the uncorrupted signal strength vector $\mathbf{s}$ is a correct localization result $\mathbf{p} = G_{alg}(\mathbf{s})$, while the corrupted signal strength vector produces an attacked localization result $\tilde{\mathbf{p}} = G_{alg}(\tilde{\mathbf{s}})$. Here, $\mathbf{p}$ and $\tilde{\mathbf{p}}$ are set-valued and may either be a single point or a region in D.

4.2 Attack Susceptibility Metrics

We wish to quantify the effect that an attack has on localization by relating the effect of a change in a signal strength reading $\mathbf{s}$ to the resulting change in the localization result $\mathbf{p}$. We shall use $\mathbf{p}_0$ to denote the correct location of a transmitter, $\mathbf{p}$ to denote the estimated location (set) when there is no attack being performed, and $\tilde{\mathbf{p}}$ to denote the position (set) returned by the estimator after an attack has affected the signal strength. There are several performance metrics that we will use:

Estimator Distance Error: An attack will cause the magnitude of $\mathbf{p}_0 - \tilde{\mathbf{p}}$ to increase. For a particular localization algorithm G_{alg} we are interested in the statistical characterization of $||\mathbf{p}_0 - \tilde{\mathbf{p}}||$ over all possible locations in the building. The characterization of $||\mathbf{p}_0 - \tilde{\mathbf{p}}||$ depends on whether a point-based method or an area-based method is used, and can be described via its mean and distributional behavior. For a point-based method, we may measure the cumulative distribution (cdf) of the error $||\mathbf{p}_0 - \tilde{\mathbf{p}}||$ over the entire building. For area-based methods, we replace $\tilde{\mathbf{p}}$, which is a set, with its median (along the x and y dimensions separately). Thus, for area-based metrics, we calculate the CDF of the distance between the median of the estimated locations $\tilde{\mathbf{p}}_{med}$ and the true location, i.e. $||\mathbf{p}_0 - \tilde{\mathbf{p}}_{med}||$.

The CDF provides a complete statistical specification of the distance errors. It is often more desirable to look at the average behavior of the error. For point-based methods, the average distance error is simply $E[||\mathbf{p}_0 - \tilde{\mathbf{p}}||]$, which is just the average of $||\mathbf{p}_0 - \tilde{\mathbf{p}}||$ over all locations. Area-based methods allow for more options in defining the average distance error. First, for a particular value of $\mathbf{p}_0$, $\tilde{\mathbf{p}}$ is a set of points. For each $\mathbf{p}_0$, we get a collection of error values $||\mathbf{p}_0 - \mathbf{q}||$, as $\mathbf{q}$ varies over points in $\tilde{\mathbf{p}}$. For each $\mathbf{p}_0$, we may extract the minimum, 25th percentile, median, 75th percentile, and maximum. These quartile values of $||\mathbf{p}_0 - \mathbf{q}||$ are then averaged over the different positions $\mathbf{p}_0$.

Estimator Precision: An area-based localization algorithm returns a set $\mathbf{p}$. For localization, precision refers to the size of the returned estimated area. This metric quantifies the average value of the area of the localized set $\mathbf{p}$ over different signal strength readings $\mathbf{s}$. Generally speaking, the smaller the size of the returned area, the more precise the estimation is. When an attack is conducted, it is possible that the precision of the answer $\tilde{\mathbf{p}}$ is affected.

Precision vs. Perturbation Distance: The perturbation distance is the quantity $||\mathbf{p}_{med} - \tilde{\mathbf{p}}_{med}||$. The precision vs. perturbation distance metric depicts the functional dependency between precision and increased perturbation distance.

Hölder Metrics: In addition to error performance, we are interested in how dramatically the returned results can be perturbed by an attack. Thus, we wish to relate the magnitude of the perturbation $||\mathbf{s} - \tilde{\mathbf{s}}||$ to its effect on the localization result, which is measured by $||G_{alg}(\mathbf{s}) - G_{alg}(\tilde{\mathbf{s}})||$. In order to quantify the effect that a change in the signal strength space has on the position space, we borrow a measure from functional analysis [21], called the Hölder parameter (also known as the Lipschitz parameter) for G_{alg}. The Hölder parameter H_{alg} is defined via

$$H_{alg} = \max_{\mathbf{s},\mathbf{v}} \frac{||G_{alg}(\mathbf{s}) - G_{alg}(\mathbf{v})||}{||\mathbf{s} - \mathbf{v}||}.$$

For continuous G_{alg}, the Hölder parameter measures the maximum (or worst-case) ratio of variability in position space for a given variability in signal strength space. Since the traditional Hölder parameter describes the worst-case effect an attack might have, it is natural to also provide an average-case measurement of an attack, and therefore we introduce the average-case Hölder parameter

$$\overline{H}_{alg} = \text{avg}_{\mathbf{s},\mathbf{v}} \frac{\|G_{alg}(\mathbf{s}) - G_{alg}(\mathbf{v})\|}{\|\mathbf{s} - \mathbf{v}\|}.$$

These parameters are only defined for continuous functions G_{alg}, and many localization algorithms are not continuous. For example, if we look at G_R for RADAR, the result of varying a signal strength reading is that it will yield a *stair-step* behavior in position space, i.e. small changes will map to the same output and then suddenly, as we continue changing the signal strength vector, there will be a change to a new position estimate (we have switched over to a new Voronoi cell in signal space). In reality, this behavior does not concern us too much, as we are merely concerned with whether adjacent Voronoi cells map to close positions. We will revisit this issue in Section 6. Finally, we emphasize that Hölder metrics measure the perturbability of the returned results, and do not directly measure error.

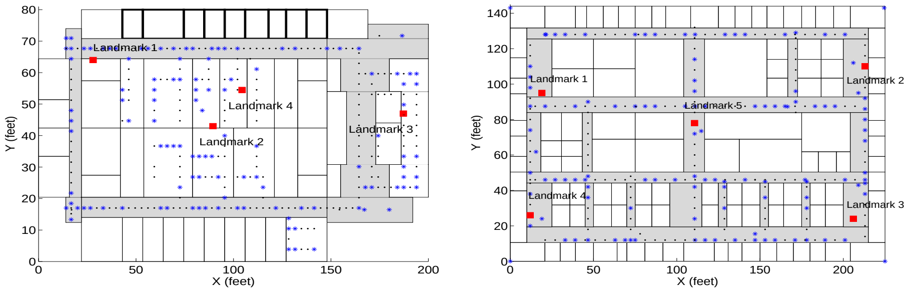
5 Experimental Results

In this section we present our experimental results. We first describe our experimental method. Next, we examine the impact of attacks on the RSS to localization error when attacking all landmarks simultaneously as well as single-landmark attacks. We then quantify the algorithms' linear responses to RSS changes. Finally, we present a precision study that investigates the impact of attacks on the returned areas for area-based algorithms.

5.1 Experimental Setup

Figure 2 shows our experimental set up. The floor map on the left, (a) is the 3rd floor of the CoRE building at Rutgers, which houses the computer science department and has an area of 200x80ft (16000 ft^2). The other floor shown in (b) is an industrial research laboratory (we call the Industrial Lab), which has an area of 225x144ft (32400 ft^2). The stars are the training points, the small dots are testing points, and the larger squares are the landmarks, which are 802.11 access points. Notice that the 4 CoRE landmarks are more co-linear than the 5 landmarks in the Industrial Lab.

For both attenuation and amplification attacks, we ran the algorithms but modified the RSS of the testing points. We altered the RSS by +/-5 dB to +/-25 dB, in increments of 5 dB. We experimented with different ways to handle signals that would fall below the detectable threshold of -92 dBm for our cards. We found that substituting the minimal signal (-92 dBm) produced about the same localization results and did not require changing the algorithms to special case missing data.



(a) CoRE

(b) Industrial Lab

Fig. 2. Deployment of landmarks and training locations on the experimental floors

We experimented different training set sizes, including 35, 115, 225, 253 and 286 points. Although there are some small differences, we found that the behavior of the algorithms matches previous results and varied little after using 115 training points, and we thus used a training set size of 115 for this study.

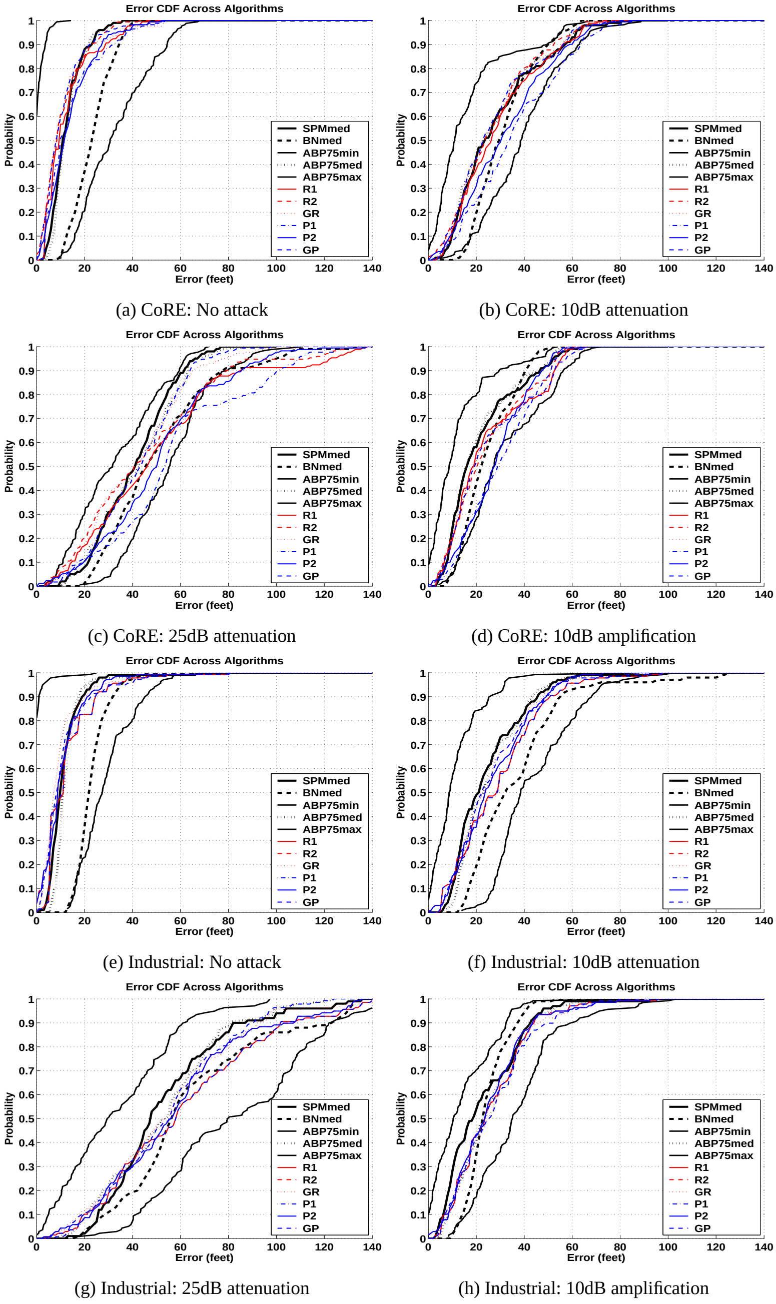
5.2 Localization Error Analysis

In this section, we analyze the estimator distance error through the statistical characterization of $\|\mathbf{p}_0 - \tilde{\mathbf{p}}\|$ by presenting the error CDFs of all the algorithms as a function of attenuation and amplification attacks. The CDF provides a complete statistical specification of the distance errors.

Figure 3(a) shows the normal performance of the algorithms for the CoRE building and (e) shows the results for the Industrial Lab. For the area-based algorithms, the median tile error is presented, as well as the minimum and maximum tile errors for ABP-75. As in previous work, the algorithms all obtain similar performance, with the exception of BN which slightly under-performs the other algorithms.

Figures 3(b) and 3(c) show the error CDFs under simultaneous landmark attenuation attacks of 10 and 25 dB for CoRE, respectively, while Figure 3(f) and 3(g) show the similar results in the industrial lab. First, bulk of the curves shift to the right by roughly equal amounts: no algorithm is qualitatively more robust than the others. Comparing the two buildings, the results show that the industrial lab errors are slightly higher for attacks at equal dB, but again, qualitatively the impact of the building environment is not very significant.

Figures 3(d) and 3(h) show the error CDFs for the CoRE and Industrial Lab under a 10 dB amplification attack. The results are qualitatively symmetric with respect to the outcome of the 10 dB attenuation attack. We found that, in general, comparing amplifications to attenuations of equal dB, the errors were qualitatively the same.



(a) CoRE: No attack

(b) CoRE: 10dB attenuation

(c) CoRE: 25dB attenuation

(d) CoRE: 10dB amplification

(e) Industrial: No attack

(f) Industrial: 10dB attenuation

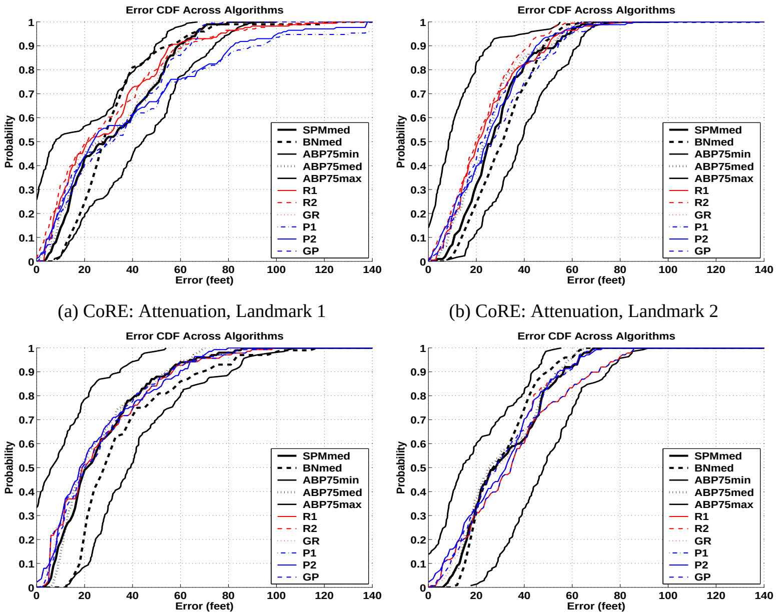
(g) Industrial: 25dB attenuation

(h) Industrial: 10dB amplification

Fig. 3. Error CDF across localization algorithms when attacks are performed on all the landmarks.

An interesting feature is that the minimum error for APB-75 also shifts to the right by roughly the same amount as the other curves. Figures 3(a) and 3(e) show that, in the non-attacked case, the minimum tile error for ABP-75 is quite small, meaning that the localized node is almost always within or very close to the returned area. However, under attacks, the closest part of the returned area moves away from the true location at the same rate as the median tile. We observed similar effects for the SPM and BN algorithms.

Next, we examine attacks against a single landmark. We found attacks against certain landmarks had a much higher impact than against others in the CoRE building. Figure 4(a) and 4(b) show the difference in the error CDF by comparing attacks of landmarks 1 and 2. Figure 2(a) shows that landmark 1 is at the southern end of the building, while landmark 2 is in the center and is close to landmark 4. The tail of the curves in Figure 4(a) are much worse than for 4(b), showing that when landmark 1 is attacked significantly more high errors are returned. we observed a similar effect for amplification attacks.



(a) CoRE: Attenuation, Landmark 1

(b) CoRE: Attenuation, Landmark 2

(c) Industrial: Attenuation, Landmark 4

(d) Industrial: Attenuation, Landmark 5

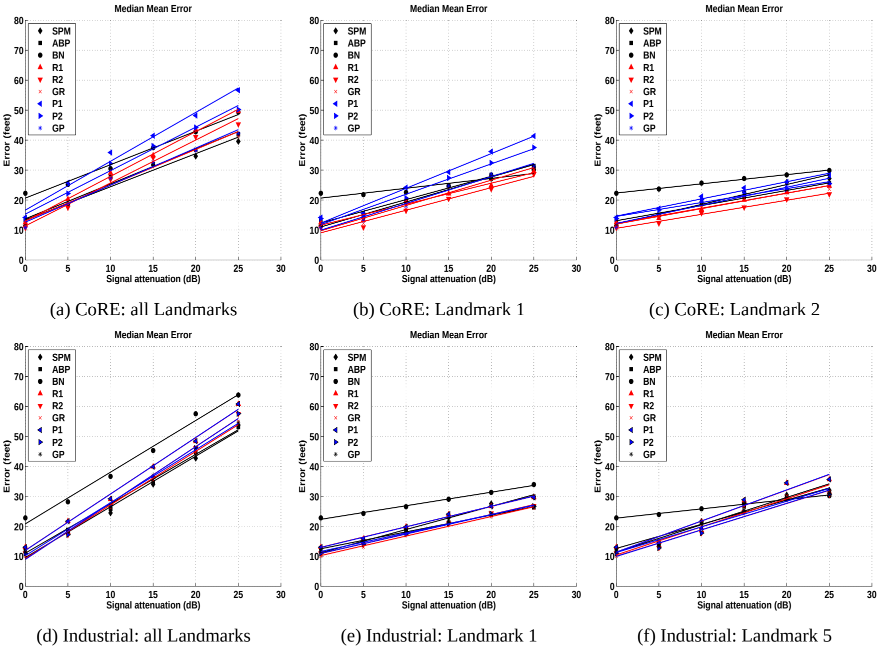
Fig. 4. Error CDF across localization algorithms when attacks are performed on an individual landmark. The attack is 25dB of signal attenuation.

The Industrial Lab results in Figures 4(c) and (d) show much less sensitivity to landmark placement compared to the CoRE building. Figure 2(b) shows that landmark 5 is centrally located and we initially suspected this would result in attack sensitivity. However, the error CDFs show that the remaining 4 landmarks provide sufficient coverage: as landmark 5 is attacked, the error CDFs are not much different from attacking landmark 4.

5.3 Linear Response

In this section, we show that the average distance error, $E[||\mathbf{p}_0 - \tilde{\mathbf{p}}||]$, of all the algorithms scales in a linear way to attacks: the localization error changes linearly with respect to the amount of signal strength change in dB (recall it is a log-scaled change in power).

Figure 5 plots the median error vs. RSS attenuation for simultaneous landmark attacks in Figure 5(a) and 5(d), and for individual landmarks in the other figures. Points are measured data, and the lines are linear least-squares fits. The most important feature is that, in all cases, the median responses of all the algorithms fits a line extremely well, with an average R^2-statistic of 0.98 for both the CoRE and Industrial Lab, and a worse-case R^2 of 0.94 for both buildings. Comparing the slopes across all the algorithms, we found a mean change in positioning error vs. signal attenuation of 1.55 ft/dB under simultaneous attacks with a minimum of 1.3 ft/dB and maximum of 1.8 ft/dB. For the single landmark attack, the slope was substantially less, 0.64 ft/dB, although BN degrades consistently less than the other algorithms at 0.44 ft/dB. The linear fit results are quite important as it means that no algorithm has a cliff where the average positioning error suffers a catastrophic failure under attack. Instead, it remains proportional to the severity of the attack.



(a) CoRE: all Landmarks

(b) CoRE: Landmark 1

(c) CoRE: Landmark 2

(d) Industrial: all Landmarks

(e) Industrial: Landmark 1

(f) Industrial: Landmark 5

Fig. 5. Median mean error across localization algorithms under attenuation attack

While the median error characterizes the overall response to attacks, it does not address whether an attacker can cause a few, large errors. We examined the response of the maximum error as a function of the strength of the attack, i.e. how the 100th percentile error scales as a function of the change in dB. We note that this characterization is not the same as, nor is directly related to, the Hölder metrics. Those metrics define the rates of change between physical and signal space within the localization function itself, while here we characterize the change in the estimator error to the change in signal, i.e. $\|\mathbf{p}_0 - \tilde{\mathbf{p}}\| / \|\mathbf{s} - \mathbf{v}\|$.

Figure 6 plots the worst-case error for each algorithm as a function of signal dB for the CoRE building. The figure shows that almost all the responses are again linear, with least-squares fits of R^2 values of 0.84 or higher, though SPM does not have a linear response. The second important point is the algorithms' responses vary, falling into three groups. BN, R1 and R2 are quite poor, with the worse case error scaling at about 4 ft/dB. P1 and P2, are in a second class, scaling at close to 3 ft/dB. The gridded algorithms, GP and GR, as well as ABP-75 fair better, scaling at 2 ft/dB or less. Finally, SPM is in a class by itself, with a poor linear fit (R^2 of 0.61) and the maximum error topping out at about 85 ft after 15 dB of attack.

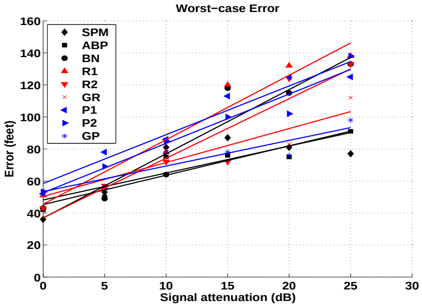
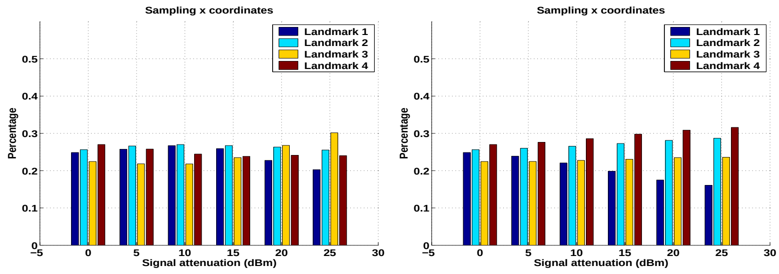


Fig. 6. Maximum error as a function of attack strength for CoRE

Examining the error CDFs and the maximum errors, we can see that most of the localizations move fairly slowly in response to an attack, at about 1.5 ft/dB. However, for some of the algorithms, particularly BN, R1 and R2, the top part of the error CDF moves faster, at about 4 ft/dB. What this means is that, for a select few points, an attacker can cause more substantial errors of over 100 ft. However, at most places in the building, an attack can only cause errors with much less magnitude.

Figure 5 show that BN is more robust compared to other algorithms for individual landmark attacks. Recall BN uses a Monte-Carlo sampling technique (Gibbs sampling) to compute the full joint-probability distribution for not just the position coordinates, but also for every node in the Bayesian network. Under a single landmark attack we found the network reduces the contribution of network nodes directly affected by the attacked landmark to the full joint-probability distribution while increasing other landmarks' contributions. In effect, the network "discounts" the attacked landmark's contribution to the overall joint-density because the attacked data from that landmark is highly unlikely given the training data.

To show this effect we developed our own Gibbs sampler so that we could observe the relative contributions of each node in the Bayesian network to the final answer.

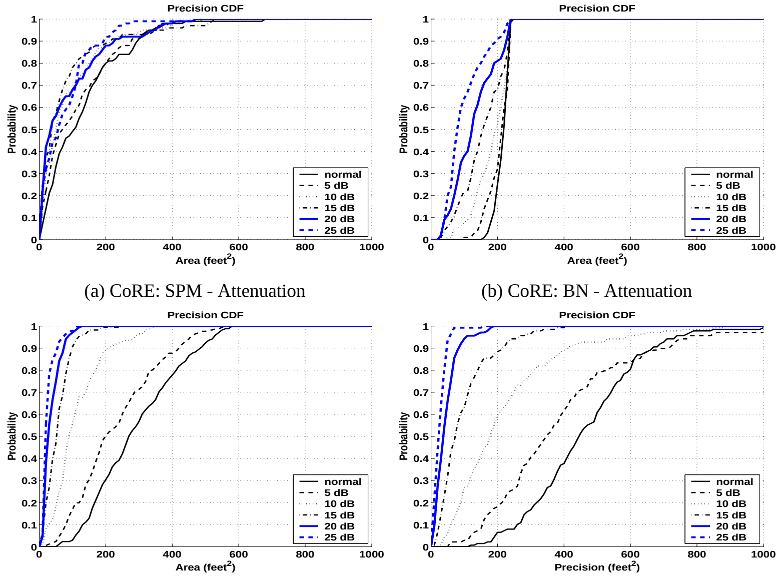


(a) CoRE: all landmarks

(b) CoRE: Landmark 1

Fig. 7. Contribution of each Landmark during sampling in the BN algorithm under attenuation attacks.

Figure 7 shows the percentage contribution for each landmark to overall joint-density. For instance, in CoRE, the contribution of each landmark starts almost uniformly. When Landmark 1 under attack, the contribution of Landmark 1 goes from 0.25 down to 0.15.



(a) CoRE: SPM - Attenuation

(b) CoRE: BN - Attenuation

(c) CoRE: ABP - Attenuation

(d) Industrial: ABP - Amplification

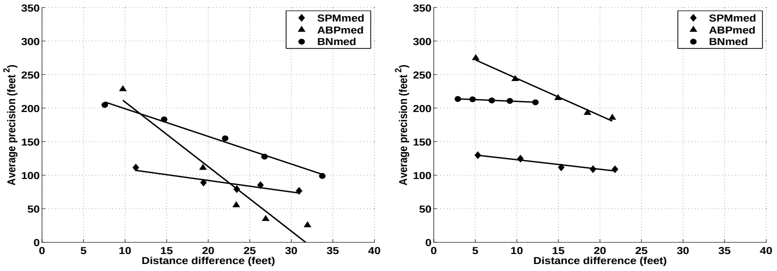
Fig. 8. Analysis of precision CDF across area-based algorithms. The attack is performed on all the landmarks.

5.4 Precision Study

In this section, we examine the area-based algorithms' precision in response to attacks. Figure 8 shows the CDF of the precision (i.e. size of the returned area) for different area-based algorithms under attack for all the landmarks in CoRE and Industrial Lab.

We found the algorithms did not become less precise in response to attacks, but rather, the algorithms tended to shift and shrink the returned areas. Figure 8(a) shows a small average shrinkage for SPM in the CoRE building, and likewise, 8(b) shows a similar effect for BN.

ABP-75 had the most dramatic effect. Figures 8(c) and 8(d) show the precision versus the attack strength for both buildings. The shrinkages are quite substantial. We found that, under attack, the probability densities of the tiles shrank to small values that were located on a few tiles– reflecting the fact that an attack causes there not to be a likely position to localize a node. We also found that this effect held for amplification attacks, as is shown in Figure 8(d). The shrinking precision behavior may be useful for attack detection, although a full characterization of how this effect occurs remains for future work.



(a) CoRE: all landmarks

(b) CoRE: Landmark 1

Fig. 9. Precision vs. perturbation distance under attenuation attack

Examining this effect further, Figure 9 presents the precision vs. the attack strength, with a least squares line fit. Figure 9(a) shows the effect when attacking all landmarks on the CoRE building. Figure 9(b) shows a downward trend, but much weaker, when attacking one landmark. We observed similar results for the Industrial Lab. We see mostly linear changes in precision in response to attacks, although with great differences between the algorithms. The figures show that the decrease in precision as function of dB is particularly strong for ABP-75.

6 Discussion about Hölder Metrics

In the previous section we examined the experimental results, and looked at the performance of several localization algorithms in terms of error and precision. We now focus on the performance of these localization algorithms in terms of the Hölder metrics. The Hölder metrics measure the variability of the *returned* answer in response to changes in the signal strength vectors.

We first discuss the practical aspects of measuring H and $\overline{H}$ for different algorithms. In Section 4, the Hölder parameters are defined by calculating the maximum and average over the entire n-dimensional signal strength space. In practice, it is necessary to perform a sampling technique to measure H and $\overline{H}$. Additionally, as noted earlier, the definition of H and $\overline{H}$ are only suitable for (Hölder) continuous functions G_{alg}. In reality, several localization algorithms, such as RADAR, are not continuous and involve the tessellation of the signal strength space into Voronoi cells V_j, and thus only a discrete set of localization results are produced (image of V_j under G_{alg}). Hence, for any $\mathbf{s} \in V_j$ we have $G_R(\mathbf{s}) = (x_j, y_j)$. Unfortunately, for neighboring Voronoi cells, we may take $\mathbf{s} \in V_j$ and $\mathbf{v} \in V_i$ such that they are arbitrarily close (i.e. $\|\mathbf{s} - \mathbf{v}\| \to 0$), while $\|G_R(\mathbf{s}) - G_R(\mathbf{v})\| \neq 0$. In such a case, the formal calculation of H and $\overline{H}$ is not possible. However, for our purposes, we are only interested in measuring the notion of adjacency of Voronoi cells in signal space yielding *close* localization results. Thus, our calculation of H and $\overline{H}$ is only performed over the centroids of the various Voronoi cells for localization algorithms that tessellate of signal strength space.

Table 2. Analysis of (worst-case) H and (average-case) $\overline{H}$

Algorithms	**CoRE:** H	**LAB:** H	**CoRE:** $\overline{H}$	**LAB:** $\overline{H}$
Area-Based				
SPM	23.7646	11.0659	1.8856	2.3548
ABP-75	20.0347	23.0652	1.8548	2.3424
BN	31.7324	14.9168	2.0595	2.5873
Point-Based				
R1	36.2400	20.7846	1.9750	2.3677
R2	19.8586	8.7313	1.9138	2.3058
GR	35.9880	20.6886	1.9691	2.3628
P1	20.8832	20.7846	1.9793	2.3683
P2	19.8586	8.7313	1.9178	2.3058
GP	21.8303	20.6886	1.9649	2.2882

The Hölder parameters for the different localization algorithms are presented in Table 2. Examining these results, there are several important observations that can be made. First, if we examine the results for $\overline{H}$ we see that, for each building, all of the algorithms have very similar $\overline{H}$ values. Hence, we may conclude that the average variability of the returned localization result to a change in the signal strength vector is roughly the same for all algorithms. This is an important result as it means, regardless of which RF fingerprinting localization system we deploy, the average susceptibility of the returned results to an attack is essentially identical.

However, if we examine the results for H, which reflects the worst-case susceptibility, then we see that there are some differences across the algorithms. First, comparing H and $\overline{H}$ for both point-based and area-based algorithms, we see that the worst-case variability can be much larger than the average variability. Additionally, the point-based methods appear to cluster. Notably, RADAR (R1) and Gridded Radar (GR) have similar performance across both CoRE and the Industrial Lab, while averaged RADAR (R2) and averaged Highest Probability (P2) have similar performance across both buildings. A very interesting phenomena is observed by looking at the algorithms that returned an average of likely locations (R2 and P2). Across both buildings these algorithms exhibited less variability compared to other algorithms. This is to be expected as averaging is a smoothing operation, which reduces variations in a function. This observation suggests that R2 and P2 are more robust from a worst-case point-of-view than other point-based algorithms.

7 Conclusion

In this paper, we analyzed the robustness of RF-fingerprinting localization algorithms to attacks that target signal strength measurements. We first examined the feasibility of conducting amplification and attenuation attacks, and observed a linear dependency between non-attacked signal strength and attacked signal strength readings for different barriers placed between the transmitter and a landmark receiver. We provided a set of performance metrics for quantifying the effectiveness of an attenuation/amplification attack. Our metrics included localization error, the precision of area-based algorithms, and a new family of metrics, called Hölder metrics, that quantify the variability of the returned answer versus change in the signal strength vectors. We conducted a trace-driven evaluation of several point-based and area-based localization algorithms where the linear attack model was applied to data measured in two different office buildings. We found that the localization error scaled similarly for all algorithms under attack. Further, we found that, when attacked, area-based algorithms did not experience a degradation in precision although they experienced degradation in accuracy. We then examined the variability of the localization results under attack by measuring the Hölder metrics. We found that all algorithms had similar average variability, but those methods returned the average of a set of most likely positions exhibited less variability. This result suggests that the average susceptibility of the returned results to an attack is essentially identical across point-based and area-based algorithms, though it might be desirable to employ either area-based methods or point-based methods that perform averaging in order to lessen the worst-case effect of a potential attack.

References

1. Hightower, J., Vakili, C., Borriello, G., Want, R.: (Design and calibration of the spoton ad-hoc location sensing system) (unpublished).
2. Enge, P., Misra, P.: Global Positioning System: Signals, Measurements and Performance. Ganga-Jamuna Pr (2001)
3. Priyantha, N., Chakraborty, A., Balakrishnan, H.: The cricket location-support system. In: Proceedings of the ACM International Conference on Mobile Computing and Networking (MobiCom). (2000)
4. Shang, Y., Ruml, W., Zhang, Y., Fromherz, M.P.J.: Localization from mere connectivity. In: Proceedings of the Fourth ACM International Symposium on Mobile Ad-Hoc Networking and Computing (MobiHoc). (2003)
5. Niculescu, D., Nath, B.: Ad hoc positioning system (APS). In: Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM). (2001) 2926–2931
6. Youssef, M., Agrawal, A., Shankar, A.U.: WLAN location determination via clustering and probability distributions. In: Proceedings of IEEE PerCom'03, Fort Worth, TX (2003)
7. Roos, T., Myllymaki, P., H.Tirri: A Statistical Modeling Approach to Location Estimation. IEEE Transactions on Mobile Computing **1**(1) (2002)
8. Battiti, R., Brunato, M., Villani, A.: Statistical Learning Theory for Location Fingerprinting in Wireless LANs. Technical Report DIT-02-086, University of Trento, Informatica e Telecomunicazioni (2002)

9. Bahl, P., Padmanabhan, V.N.: Radar: An in-building rf-based user location and tracking system. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). (2000)
10. Doherty1, L., Pister, K.S.J., ElGhaoui, L.: Convex position estimation in wireless sensor networks. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). (2001)
11. Hu, Y., Perrig, A., Johnson, D.: Packet leashes: a defense against wormhole attacks in wireless networks. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). (2003)
12. Li, Z., Trappe, W., Zhang, Y., Nath, B.: Robust statistical methods for securing wireless localization in sensor networks. In: Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks (IPSN 2005). (2005)
13. Liu, D., Ning, P., Du, W.: Attack-resistant location estimation in sensor networks. In: Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks (IPSN 2005). (2005)
14. Capkun, S., Hubaux, J.P.: Secure positioning of wireless devices with application to sensor networks. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). (2005)
15. Brands, S., Chaum, D.: Distance-bounding protocols (1994)
16. Sastry, N., Shankar, U., Wagner, D.: Secure verification of location claims. In: Proceedings of the 2003 ACM workshop on wireless security. (2003) 1–10
17. Capkun, S., Hubaux, J.: (Securing localization with hidden and mobile base stations) to appear in Proceedings of IEEE Infocom 2006.
18. Lazos, L., Poovendran, R., Capkun, S.: Rope: robust position estimation in wireless sensor networks. In: Proceedings of the Fourth International Symposium on Information Processing in Sensor Networks (IPSN 2005). (2005) 324–331
19. Elnahrawy, E., Li, X., Martin, R.P.: 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). (2004)
20. Madigan, D., Elnahrawy, E., Martin, R., Ju, W., Krishnan, P., Krishnakumar, A.S.: Bayesian indoor positioning systems. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM). (2005) 324–331
21. Lang, S.: Real and Functional Analysis. Springer (1993)