Analysis on Perfect Location Spoofing Attacks Using Beamforming

Ting Wang and Yaling Yang
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Email: {wangting, yyang8}@vt.edu

***Abstract*—Location spoofing attacks pose serious threats to the location based wireless network mechanisms. Most existing literature focuses on detecting location spoofing attacks or design of robust localization algorithms. However, our study shows that, in many circumstances, perfect location spoofing (PLS) can stay undetected even if robust localization algorithms or detection mechanisms are used. In this paper, we present theoretical analysis on the feasibility of beamforming-based PLS attacks and how it is affected by the anchor deployment. We formulate PLS as a nonlinear feasibility problem based on smart antenna array pattern synthesis. Due to the intractable nature of this feasibility problem, we solve it using semidefinite relaxation (SDR) in conjunction with a heuristic local search algorithm. Simulation results show the effectiveness of our analytical approach and provide insightful advices for defence against PLS attacks.**

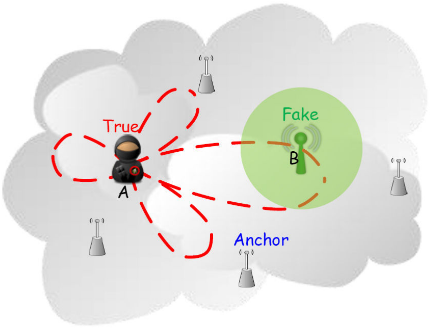


Fig. 1. A Perfect Location spoofing attack.

I. Introduction

Radio localization systems have been integrated into many wireless network solutions. For example, location-based access control (LBAC) determines the users' privileges of accessing critical information by taking the users' physical locations into account [1], [2]. Identity spoofing detection mechanisms use location information to differentiate malicious nodes from legitimate nodes [3]. In these applications, the correctness of the location results provided by the localization systems is critical. Attacks on localization systems can cause errors in location estimation and consequently break these location based mechanisms of wireless networks.

Generally, there are two categories of attacks on localization systems. The first category is location concealing, where the adversary does not have a specific target fake location. The goal of location concealing is simply to distort the measurements of the localization system so that the true location of the adversary cannot be identified. The other category is location spoofing, where an attacker masquerades as being at another target location by falsifying the measurements of the localization system. (An illustration is given by Figure 1.) Between the two categories, the latter is more of a threat to the security of wireless networks in the sense that locationally masqueraded attackers can take illegitimate advantage of the network resources and launch further attacks to the network. For instance, in applications of LBAC, if the attacker can masquerade to be at a position where high access privileges are given, he / she may illegitimately access confidential resources. In tracing of adversary in wireless networks, attackers with capability of masquerading locations may plant the crime on innocent wireless nodes and disturb the judgement of the security mechanism. Similarly, the identity attack detection schemes that rely on localization may also fail due to this kind of location spoofing attacks [3]–[5]. In this paper, we focus on investigation of location spoofing attack to received signal strength (RSS) based localization systems.

The existence of potential location spoofing attacks has been identified for quite a few years. In [6], it is experimentally shown that by attenuating or amplifying the RSS readings at the anchors, the localization system may conclude in false location estimation. Bauer et al. show that attackers with directional antennas [7] have the ability to bias the location estimation to a direction of their choice in addition to introducing significant localization errors. All these works indicate that location spoofing is possible. A few robust localization algorithms have been proposed to countermeasure location spoofing attacks [6], [8], [9], in which statistical analysis methods are used to detect and eliminate some of the biased RSS measurements.

Unfortunately, despite all the existing efforts, we have discovered that robust RSS localization schemes are all limited in their effectiveness no matter what statistical methods they use. In many circumstances, there exist location spoofing attacks that can stay undetected under all of these robust RSS based localization algorithms. We call such attack as "*perfect location spoofing (PLS) attack*". To fully understand the limitations of robust localization algorithms and the level of threat of PLS attacks, in this paper, we provide a thorough theoretical analysis on the capability of using PLS attacks to evade robust RSS localization algorithms. Our analysis assumes that the attacker uses beamforming to control the directional gains of his radio and hence falsifies the RSS readings at the anchors. To the best of our knowledge, our analysis is the first to provide answers to the following critical and fundamental questions: *Is it possible for an attacker to launch location spoofing attacks against any robust RSS localization algorithm? Can an attacker launch a PLS attack to a specific location no matter where he/she is? What can be done to reduce the possibility of PLS attacks?*

To answer these questions, we first formulate the PLS problem as a nonlinear feasibility problem based on smart antenna pattern synthesis. Due to the intractable nature of this feasibility problem, we derive close upper and lower bounds of the solution to the feasibility problem through semidefinite relaxation and local search method. Using these upper and lower bounds, we further investigate how the feasibility of PLS is related to anchor deployments and antenna capability. Based on the investigation results, we provide guidelines for localization anchor deployment strategies that can significantly reduce the threat of PLS attacks. To our best knowledge, our work is the first study that presents mathematical formulation and theoretical analysis on the feasibility of PLS attacks, how they are affected by the anchor deployment and how PLS attacks can be solved.

The rest of the paper is organized as follows. Section II introduces the basic knowledge about RSS based localization. The attack model is described in Section III. Section IV gives an overview of our analysis. Section V provides the formulation of the PLS problem. Section VI introduces our approach for solving this problem. Experimental results are reported in Section VII and defence against PLS attacks is discussed in Section VIII. Finally, Section IX concludes the paper.

II. RSS BASED LOCALIZATION

In this section, we give a brief overview of RSS-based localization systems which are adopted by most commercial applications. In an RSS based localization system, there are multiple signal receivers placed at specific locations which measure the RSS of wireless nodes and report the measurements to the system. These signal receivers are referred as "*anchors*". Localization algorithms are then used to compute location estimates based on anchor measurements. These algorithms fall into two categories: fingerprint based approaches [10], [11] and propagation model based approaches [12], [13]. Fingerprint localization collects RSS measurements of wireless nodes at sample locations and stores these measurements in database in the off-line phase. During the on-line phase, the actual RSS measurements of the target wireless node are compared with the stored database and through pattern matching, a location estimation can be returned. In propagation model based RSS localization schemes, the distances from the wireless node to the anchors are estimated using large scale path loss model. With these estimated distances, the localization result is obtained geometrically by trilateration.

For both types of approaches, robust data processing algorithms [6], [8], [9] can be used to detect and eliminate some abnormal and biased anchor measurements introduced by location spoofing attacks. However, no matter what robust algorithms are used for these two types of approaches, we will show in the next section that some location spoofing attacks can still be undetectable.

III. ATTACK MODEL

The location spoofing attack model is illustrated in Figure 1, where point "A" is the true location of the attacker and point "B" is the attacker's fake location. Note that no matter which localization algorithm is used in the system, as long as the attacker ensures that the RSS readings at all the anchors are the same as what the RSS readings should be for a wireless node at the fake location, there is no way for any robust localization algorithm to detect the location spoofing attack. We call this type of location spoofing attack as "*perfect location spoofing (PLS)*". Under PLS attack, the location estimations produced by all localization algorithms are the same fake location within a reasonable small error range that are determined by the noise level and the precision of the algorithms.

For conventional omnidirectional radios, realizing PLS attack is not possible because the path loss vector from point "A" to the anchors are different from the path loss vector from point "B". Thus, wireless nodes with omnidirectional antennas at the two points will produce different RSS reading vectors at the anchors. An attacker at point A, however, can use smart antenna's beamforming capability to solve this problem and make the RSS readings at anchors the same as the readings should be when he is at position B. To do this, the attacker first obtains the anchor locations by wardriving [14] (moving around to locate anchors in the neighboring area), spying or other underhand means. Then, the attacker tunes the diverse directional gains produced by beamforming to compensate the difference between the two path loss vectors from A and B, and hence, falsify the RSS reading vectors at the anchors. To estimate the path loss, the attacker can leverage signal propagation models and field measurement methods. If the falsified RSS reading vector is within the range of typical noised RSS reading vectors produced by a legitimate wireless node at the fake location, we say that a PLS attack is created.

In the remainder of this paper, we assume that the attacker is equipped with a circular smart antenna array which consists of N_{ant} isotropic elements placed over a circle with radius R. The i^{th} antenna element is located with the phase angle ϕ_i. The beamforming pattern of this circular smart antenna array is expressed as [15]:

$$G(\theta) = \sum_{i=1}^{N_{ant}} w_i \exp[j\frac{2\pi}{\lambda}R\cos(\theta - \phi_i)], \qquad (1)$$

where λ is the signal wavelength, θ represents the direction and $\mathbf{w} = [w_1, w_2, \cdots, w_{N_{ant}}]^H$ is the complex weight vector which can be tuned to change the radiation pattern.

We choose circular antenna array as an example to illustrate the formulation of our analysis because it can produce flexible asymmetric beamforming patterns and easily deflect a beam through 2π. However, our analysis is not limited to circular

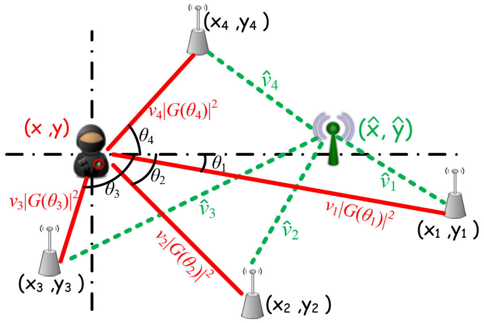


Fig. 2. Compensating path loss differences using beamforming.

antenna array. It is straightforward to plug antenna models with other geometric forms into the analysis by replacing equation (1) with their corresponding beamforming functions.

IV. Problem Overview

To understand how we analyze the conditions that an attacker can launch a PLS attack, let us look at a simple example illustrated in Figure 2, where an attacker at location (x, y) wants to fake his/her location at $(\hat{x}, \hat{y})$. There are K anchors in the neighboring area and their locations are denoted by $(x_k, y_k),\ k = 1, 2, \cdots, K$. Since less than 4 anchors are not enough to uniquely localize even a legitimate node with unknown transmit power, we are only interested in the cases where $K \geq 4$. Suppose the expected path loss from the true location, (x, y), to the i^{th} anchor is v_i and $\mathbf{v} = [v_1, v_2, \cdots v_K]^T$. The expected path loss from the fake location, $(\hat{x}, \hat{y})$, to the i^{th} anchor is $\hat{v}_i$ and $\hat{\mathbf{v}} = [\hat{v}_1, \hat{v}_2, \cdots, \hat{v}_K]^T$. Denote the direction angle of the i^{th} anchor with respect to the true location (x, y) as θ_i and $\Theta = [\theta_1, \theta_2, \cdots \theta_K]^T$. To realize a PLS attack, all the anchors' RSS readings of the attacker should be the same as the RSS readings of a legitimate user at the fake location $(\hat{x}, \hat{y})$. Hence, the attacker must tune his/her beamforming pattern defined in (1) to satisfy the following constrains,

$$v_k|G(\theta_k)|^2 \approx \hat{v}_k, \forall k = 1, 2, \cdots, K. \tag{2}$$

Essentially, (2) shows that the requirement of PLS for the attacker is that the beamforming gains in the directions of the anchors must compensate the difference between the path loss vectors from the attacker's true location and from the fake location. If a $\mathbf{w}$ that satisfies all the constraints in (2) can be obtained, we conclude that a PLS attack is feasible. Otherwise, it is infeasible to falsify all the RSS readings at the anchors by beamforming and the abnormal RSS readings can potentially be detected by some robust localization algorithms. It is important to note that we use an approximate sign in (2) instead of an equal sign. This is because signal path loss usually has random variations. Hence, as long as the attacker's RSS readings are close enough to the desired fake values, a localization system cannot tell whether the discrepancy is caused by attack or natural variations in path loss.

In spite of the straightforward concept of formula (2), we will show in Section V that the formal mathematical formulation of the PLS feasibility problem is NP-hard in general. In order to solve this problem, we leverage semidefinite relaxation (SDR) technique and a local search algorithm to get approximate answers. Simulation results show that our approximation method can closely approximate the intractable feasibility problem and is effective in analyzing the feasibility of PLS attacks under different situations. The analysis will answer the following questions:

- How the feasibility of PLS attacks is related to the anchor density and the hardware capability of the attacker's smart antenna?
- What kind of anchor deployment is good for guard against location spoofing attacks?
- Given an anchor deployment and a particular fake location, where the attacker could be hiding to launch a PLS attack?

The answers to the above questions provide insightful guidance for deployment of spoof resistant localization system and tracing of location spoofing adversaries.

V. Problem Formulation

In this section, we formulate the feasibility of PLS in the form of a non-linear programming problem.

According to the log-distance path loss model, the expected path loss vectors from the attacker's true location (x, y) and his/her target fake location $(\hat{x}, \hat{y})$ to the anchors at locations $(x_k, y_k),\ k = 1, 2, \cdots, K$ can be expressed as the following equations respectively,

$$\begin{aligned} &\mathbf{v} = [v_1, v_2, \cdots v_K]^T, \\ &v_k = PL_0[(x_k - x)^2 + (y_k - y)^2]^{\alpha/2}, \\ &\hat{\mathbf{v}} = [\hat{v}_1, \hat{v}_2, \cdots, \hat{v}_K]^T, \\ &\hat{v}_k = PL_0[(x_k - \hat{x})^2 + (y_k - \hat{y})^2]^{\alpha/2}, \end{aligned} \tag{3}$$

where PL_0 is the path loss at the reference distance $d_0 =$ 1m and α is the path loss exponent. As illustrated in Figure 2, the angle vector $\Theta = [\theta_1, \theta_2, \cdots \theta_K]^T$ is defined by the relationships:

$$\begin{aligned} &\cos\theta_k = \tfrac{x_k - x}{\sqrt{(x_k - x)^2 + (y_k - y)^2}}, \\ &\sin\theta_k = \tfrac{y_k - y}{\sqrt{(x_k - x)^2 + (y_k - y)^2}}, \\ &k = 1, 2, \cdots, K. \end{aligned} \tag{4}$$

From (1), the beamforming directional gain in the direction of the k^{th} anchor can be written as

$$|G(\theta_k)|^2 = |\mathbf{w}^H \mathbf{h}_k|^2, \tag{5}$$

where

$$\mathbf{h}_k = \begin{bmatrix} \exp[j\frac{2\pi}{\lambda}R\cos(\theta_k - \phi_1)] \\ \exp[j\frac{2\pi}{\lambda}R\cos(\theta_k - \phi_2)] \\ \vdots \\ \exp[j\frac{2\pi}{\lambda}R\cos(\theta_k - \phi_{N_{ant}})] \end{bmatrix}. \tag{6}$$

In Section IV, we have briefly described the requirement for a PLS using (2). We use an approximation sign in formula (2) because $v_k|G(\theta_k)|^2$ does not need to be exactly equal to $\hat{v}_k$ for realizing a PLS due to path loss variations in nature environment. Following shadow fading model, the path loss is a Gaussian distributed random variable with a standard deviation in dB. As long as the difference between $v_k|G(\theta_k)|^2$ and $\hat{v}_k$ is within the typical deviation level, a localization system is not able to differentiate the falsified RSS measurements from noised RSS measurements. Thus we use the standard deviation of the Gaussian noise, $\delta(\text{dB}) > 0$, as the threshold that defines the closeness requirement for the approximation in (2) to hold. Based on this threshold, we convert (2) into

$$|10\log_{10}(v_k|G(\theta_k)|^2) - 10\log_{10}(\hat{v}_k)| \leq \delta(\text{dB}). \tag{7}$$

For simplicity, note that (7) is equivalent to

$$\frac{1}{\delta} \leq \frac{v_k}{\hat{v}_k}|G(\theta_k)|^2 \leq \delta. \tag{8}$$

By plugging (5) into (8), we get

$$\frac{1}{\delta} \leq \frac{v_k}{\hat{v}_k}|\mathbf{w}^H\mathbf{h}_k|^2 \leq \delta. \tag{9}$$

Letting

$$\mathbf{f}_k = (\frac{v_k}{\hat{v}_k})^{\frac{1}{2}}\mathbf{h_k}, \tag{10}$$

(9) becomes:

$$\frac{1}{\delta} \leq |\mathbf{w}^H\mathbf{f}_k|^2 \leq \delta. \tag{11}$$

The feasibility of PLS now can be modeled as:

$$\begin{array}{ll} \text{find any} & \mathbf{w}^H \\ \text{s.t.} & |\mathbf{w}^H\mathbf{f}_k|^2 \leq \delta \\ & |\mathbf{w}^H\mathbf{f}_k|^2 \geq \frac{1}{\delta} \\ & k = 1, 2, \cdots, K. \end{array} \tag{12}$$

The formulation in (12) is a nonlinear feasibility problem, in which the answer is "yes" or "no". However, this feasibility problem is difficult to answer as we have the following claim.

***Claim** 1:* The feasibility problem (12) is NP-hard.

Proof: To prove the complexity of problem (12), let us consider its complementary problem:

$$\begin{array}{ll} \max\limits_{\mathbf{w}} & t = \min\limits_{k}\{|\mathbf{w}^H\mathbf{f}_k|^2\}_{k=1}^{K} \\ \text{s.t.} & |\mathbf{w}^H\mathbf{f}_k|^2 \leq \delta,\ k = 1, 2, \cdots, K. \end{array} \tag{13}$$

Given the solution t^* of (13), problem (12) is feasible if and only if $t^* \geq \frac{1}{\delta}$ and it is infeasible if and only if $t^* < \frac{1}{\delta}$. Conversely, suppose problem (12) can be solved in polynomial time. Then for any given value of $\eta \in [0, \delta]$, we can also solve the following problem in polynomial time.

$$\begin{array}{ll} \text{find} & \mathbf{w}^H \\ \text{s.t.} & |\mathbf{w}^H\mathbf{f}_k|^2 \leq \delta \\ & |\mathbf{w}^H\mathbf{f}_k|^2 \geq \eta \\ & k = 1, 2, \cdots, K. \end{array} \tag{14}$$

Since the value of η lies in the interval $[0, \delta]$, we can use bisection method to find the turning point η', so that when $\eta > \eta'$, problem (14) is infeasible and for $\eta < \eta'$, problem (14) is feasible. This turning point η' is the maximum value of η which makes problem (14) feasible, so it is exactly the solution to the max-min problem (13). Because bisection method is a polynomial-time algorithm, the reduction from problem (13) to problem (12) is also polynomial. Thus, problem (12) and problem (13) are bidirectionally polynomial-time reducible. Hence, if we can show (13) is NP-hard, we prove that (12) is NP-hard too.

Now consider the case that the solution to (13) is obtained when $t^* = |\mathbf{w}^H\mathbf{f}_{k^*}|^2$, and (13) becomes:

$$\begin{array}{ll} \max\limits_{\mathbf{w}} & |\mathbf{w}^H\mathbf{f}_{k^*}|^2 \\ \text{s.t.} & |\mathbf{w}^H\mathbf{f}_k|^2 \leq \delta \\ & |\mathbf{w}^H\mathbf{f}_k|^2 \geq |\mathbf{w}^H\mathbf{f}_{k^*}|^2 \\ & k = 1, 2, \cdots, K. \end{array} \tag{15}$$

By moving $|\mathbf{w}^H\mathbf{f}_k|^2$ to the other side of the inequality, (13) is finally recast as

$$\begin{array}{ll} \max\limits_{\mathbf{w}} & \mathbf{w}^H(\mathbf{f}_{k^*}\mathbf{f}_{k^*}^H)\mathbf{w} \\ \text{s.t.} & \mathbf{w}^H(\mathbf{f}_k\mathbf{f}_k^H)\mathbf{w} \leq \delta \\ & \mathbf{w}^H(\mathbf{f}_{k^*}\mathbf{f}_{k^*}^H - \mathbf{f}_k\mathbf{f}_k^H)\mathbf{w} \leq 0 \\ & k = 1, 2, \cdots, K. \end{array} \tag{16}$$

In (16), $\mathbf{f}_k\mathbf{f}_k^H$ is a Hermitian positive semi-definite matrix and $\mathbf{w}^H(\mathbf{f}_{k^*}\mathbf{f}_{k^*}^H - \mathbf{f}_k\mathbf{f}_k^H)\mathbf{w}$ is an indefinite matrix. Thus, (16) is a non-convex quadratically constrained quadratic programming (QCQP) problem, which is NP-hard in general [16]. Thus, (13) is NP-hard in general and so is (12). ■

VI. Solving PLS Problem

Since the feasibility problem of PLS defined in (12) is, in general, NP-hard, we cannot analyze the properties of PLS by directly solving it. Therefore, in this section, we first provide the derivation of a relaxed problem, the solution to which will provide us an upper bound for the feasibility answers to the PLS problem (meaning if the relaxed problem is infeasible, (12) is definitely infeasible). Then we introduce a heuristic algorithm which, in most of the feasible situations of problem (12), can actually find a feasible solution through local search around carefully selected starting points. The local-search-based heuristical algorithm essentially serves as our lower bound on the PLS problem (12) (meaning that if local search can find a feasible solution, we know problem (12) is definitely feasible). Our experiment in Section VII will show that these two bounds are actually very tight and hence can be used as great approximation tools for analyzing the original PLS problem in (12).

A. Relaxation

To get the relaxed problem, first, we add an objective function to problem (12), so that when multiple solutions exist, just the one with minimum objective value is returned. Meanwhile, by letting $\mathbf{Q}_k = \mathbf{f}_k\mathbf{f}_k{}^H$, we reformulate the PLS problem as:

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & obj = \sum\limits_{k=1}^{K}(\mathbf{w}^H\mathbf{Q}_k\mathbf{w} - 1)^2 \\ \text{s.t.} & \mathbf{w}^H\mathbf{Q}_k\mathbf{w} \leq \delta \\ & \mathbf{w}^H\mathbf{Q}_k\mathbf{w} \geq \frac{1}{\delta} \\ & k = 1, 2, \cdots, K. \end{array} \tag{17}$$

The physical meaning of the objective function in (17) is the squared error of the approximation in (7) and it achieves zero when $|\mathbf{w}^H\mathbf{f}_k|^2 = \frac{v_k|G(\theta_k)|^2}{\hat{v}_k} = 1,\ \forall k \in \{1,\cdots,K\}$. Physically, a weighting vector $\mathbf{w}$ that solves (17) produces a beamforming pattern which is closest to the ideal pattern. Meanwhile, if (17) can be solved, then the PLS problem in (12) can also be solved. This is because if a solution can be found for (17), this solution must satisfy (17)'s constraints. Since (12) and (17) have the same constraints, (17)'s solution is a solution to (12). If (17) does not have a feasible solution, we know that (12) also does not have any solution.

Since $\mathbf{w}^H\mathbf{Q}_k\mathbf{w} = \text{trace}(\mathbf{w}^H\mathbf{Q}_k\mathbf{w}) = \text{trace}(\mathbf{w}\mathbf{w}^H\mathbf{Q}_k)$, we can recast (17) by assuming $\mathbf{X} = \mathbf{w}\mathbf{w}^H$ and get:

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & obj = \sum\limits_{k=1}^{K}(\text{trace}(\mathbf{X}\mathbf{Q}_k)-1)^2 \\ \text{s.t.} & \text{trace}(\mathbf{X}\mathbf{Q}_k) \le \delta \\ & \text{trace}(\mathbf{X}\mathbf{Q}_k) \ge \frac{1}{\delta} \\ & k=1,2,\cdots,K \\ & \mathbf{X} \succeq 0 \\ & \text{rank}(\mathbf{X}) = 1. \end{array} \tag{18}$$

By $\mathbf{X} \succeq 0$, we mean that $\mathbf{X}$ is a Hermitian positive semidefinite matrix.

Note that since problem (12) is NP-hard in general, so is problem (18). Hence, in the following, we will seek a heuristic solution by analyzing a relaxed version of (18). Our relaxation is based on the observation that problem (18) is very similar to a semidefinite programming problem except that the last constraint "rank($\mathbf{X}$) $= 1$" is non-convex. It is known that a semidefinite programming problem is solvable within polynomial time. Hence, we relax problem (18) by ignoring the rank constraint and get the following SDR problem.

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \sum\limits_{k=1}^{K}(\text{trace}(\mathbf{X}\mathbf{Q}_k)-1)^2 \\ \text{s.t.} & \text{trace}(\mathbf{X}\mathbf{Q}_k) \le \delta \\ & \text{trace}(\mathbf{X}\mathbf{Q}_k) \ge \frac{1}{\delta} \\ & k=1,2,\cdots,K \\ & \mathbf{X} \succeq 0 \end{array} \tag{19}$$

The SDR problem can be solved efficiently and its optimal solution provides a lower bound for the objective value in (17). If the SDR problem has no solution, we are safe to conclude that (12) has no feasible solution. This is because the relaxation makes the feasible region of (12), which is the same as the feasible region of (17), a subset of the SDR problem's feasible region.

In addition, the SDR problem not only provides us a way to weed out infeasible situations, it also can provide us with clues to search for feasible solutions to the PLS problem. Note that due to the relaxation, the optimal solution $\mathbf{X}_{opt}$ to the SDR problem may violate the "rank($\mathbf{X}$) $= 1$" constraint in (18). Since (18) and (12) essentially have the same constraints, this also means that $\mathbf{X}_{opt}$ is not feasible for problem (12) in such a case. However, note that $\mathbf{X}_{opt}$ does have the nice property that it satisfies the other constraints in (18) which means that it could be close to the feasible region of the original PLS problem. Thus, there may exist feasible solutions to the PLS problem around $\mathbf{X}_{opt}$. Based on this observation, we propose an effective local search algorithm to search for feasible solution to (12).

B. Heuristic Algorithm

While the solution $\mathbf{X}_{opt}$ of (19) gives us one starting point of local search for feasible solution to problem (12), this single starting point may not be enough. This is because the feasible space of $\mathbf{X}$ that satisfies the constraints of (19) is much larger than the feasible space that satisfies (18)'s constraints. (satisfying (18)'s constraints means satisfying (12)'s constraints). Hence, $\mathbf{X}_{opt}$, which is the optimal solution of (19), may locate far from the feasible space of (18) so that local search around $\mathbf{X}_{opt}$ cannot find a point that satisfies (18)'s constraints. To solve this problem, we essentially need to revise (19) to give us more starting points for local search. All these starting points must satisfy all the constraints of (19) except the rank constraint.

Based on this idea, we use a heuristic algorithm to approximately solve the PLS problem in (12) as follows. Our approach is a combination of a local search algorithm and a feasible region partitioning algorithm. First, starting from $\mathbf{X}_{opt}$, the local search algorithm tries to find a feasible solution to the PLS problem. In case that $\mathbf{X}_{opt}$ is far beyond the feasible region of (12) so that local search around $\mathbf{X}_{opt}$ fails, the partitioning algorithm segments the feasible region of (19) and generates additional starting points for the local search algorithm. These additional starting points enable local search to span more of the feasible region of the SDR problem (19), so that the chance of finding a feasible solution to (12) enhances. In the following, we describe the details of our partitioning algorithm and local search process.

1) Partitioning the feasible region: Algorithm 1 shows the procedure of partitioning the feasible region of the SDR problem. A partitioned SDR sub-problem is given by the following form:

$$\begin{array}{ll} \min\limits_{\mathbf{w}} & \sum\limits_{k=1}^{K}(\text{trace}(\mathbf{Z}\mathbf{Q}_k)-1)^2 \\ \text{s.t.} & \text{trace}(\mathbf{Z}\mathbf{Q}_k) \le b_k \\ & \text{trace}(\mathbf{Z}\mathbf{Q}_k) \ge a_k \\ & k=1,2,\cdots,K \\ & \mathbf{Z} \succeq 0. \end{array} \tag{20}$$

where a_k and b_k represent the new bounds defined by the partitioned feasible region pieces. For each k, $[a_k, b_k]$ is chosen to be either $[r(0), r(1)]$ or $[r(1), r(2)]$, with $r(0) = \frac{1}{\delta}$; $r(1) = (\delta + \frac{1}{\delta})/2$; $r(2) = \delta$. Thus, in total 2^K SDR sub-problems can be established. Note that problem (20)'s solution $\mathbf{Z}_{opt}$ is the same as $\mathbf{X}_{opt}$ in the following properties: Both $\mathbf{Z}_{opt}$ and $\mathbf{X}_{opt}$ are Hermitian and semidefinite matrix. Both $\mathbf{Z}_{opt}$ and $\mathbf{X}_{opt}$ satisfy the first two constraints of (18) but may violate the (18)'s last rank constraint.

This partitioning process is initiated when local search around $\mathbf{X}_{opt}$ cannot find a feasible $\mathbf{w}_l$. The SDR sub-problems are tested one by one. For each SDR sub-problem, after getting the solution $\mathbf{Z}_{opt}$, we use a local search algorithm (described in Section VI-B2) to search for a feasible $\mathbf{w}_l$ that satisfies

Algorithm 1 Partitioning of SDR feasible region.

```
Input: $\delta$
Output: $\mathbf{w}$
$r(0) = \frac{1}{\delta},\ r(1) = \frac{\delta + \frac{1}{\delta}}{2},\ r(2) = \delta$
for $j = 1 \to 2^K$ do
    $\mathbf{p} \leftarrow j_{(2)}$; {convert $j$ from decimal to binary and store the binary bits into vector $\mathbf{p}$}
    $a_k = r([\mathbf{p}]_k),\ b_k = r([\mathbf{p}]_k + 1),\ k = 1, 2, \cdots, K$;
    Solve problem (20);
    if (20) is feasible then
        Call Algorithm 2 with input $\mathbf{Z}_{opt}$ and get $\mathbf{w}$.
        if $\mathbf{w} \neq$ NULL then
            Return $\mathbf{w}$;
        end if
    end if
end for
```

Algorithm 2 Randomization based local search.

```
Input: $\mathbf{X},\ \mathbf{Q}_k,\ \delta,\ N_s$
Output: $\mathbf{w}$
Initialize $\mathbf{w} =$ NULL;
Generate $\mathbf{V}$ such that $\mathbf{V}\mathbf{V}^H = \mathbf{X}$;
for $i = 1 \to N_s$ do
    Generate a $N_{ant}$ by 1 vector $\mathbf{e}_l$ on the unit sphere of $\mathbb{C}^{N_{ant}}$;
    Test the feasibility of (12) with $\mathbf{w}_l = \mathbf{V}\mathbf{e}_l$;
    if feasible then
        Return $\mathbf{w} = \mathbf{w}_l$;
    end if
    Generate a $N_{ant}$ by 1 vector $\mathbf{u}_l$ with
    $Re\{[\mathbf{u}_l]_i\} \sim \mathcal{N}(0,1)$ and $Im\{[\mathbf{u}_l]_i\} \sim \mathcal{N}(0,1)$
    $i = 1, 2, \cdots, N_{ant}$;
    Test the feasibility of (12) with $\mathbf{w}_l = \mathbf{V}\mathbf{u}_l$;
    if feasible then
        Return $\mathbf{w} = \mathbf{w}_l$;
    end if
end for
```

all constraints in (18). If the current SDR sub-problem is infeasible, or the local search algorithm fails in finding a feasible $\mathbf{w}_l$, the next SDR sub-problem is tested. This process stops when either a feasible $\mathbf{w}_l$ is obtained, or all the 2^K SDR sub-problems have been tested.

2) Local search: The steps of the local search procedure are illustrated in Algorithm 2. In the local search algorithm, candidate solution vectors, $\mathbf{w}_l$, for (12) are randomly generated based on a given starting point $\mathbf{X}$. Note that $\mathbf{X}$ is a Hermitian and semidefinite matrix. Hence, we can use Cholesky decomposition to decompose $\mathbf{X}$ into the product of a lower triangular matrix $\mathbf{V}$ and its conjugate transpose, denoted as: $\mathbf{V}\mathbf{V}^H = \mathbf{X}$. A candidate vector is created by multiplying $\mathbf{V}$ by a randomly generated vector.

We adopt two effective methods for random vector generation which have been used by [17]. In the first method, we let $\mathbf{w}_l = \mathbf{V}\mathbf{e}_l$, where $\mathbf{e}_l$ is on the unit sphere of the N_{ant}-dimensional space of complex numbers, with each of its element having a phase uniformly distributed on $[0, 2\pi)$. In the second method, we let $\mathbf{w}_l = \mathbf{V}\mathbf{u}_l$, and $\mathbf{u}_l$ has both the real and the imaginary parts of each element following independent standard Gaussian distribution.

Suppose the randomization is repeated N_s times. Each time we put $\mathbf{w}_l$, generated by the two methods, into the PLS problem (12) and check if $\mathbf{w}_l$ satisfies all the constraints. After N_s randomization runs, if none of the $\mathbf{w}_l$ satisfies (12)'s constraints, we claim that the local search fails.

3) Discussions of the heuristic algorithm: At worst, our heuristic algorithm might have to test all the 2^K SDR sub-problems before providing a heuristic solution to the PLS problem (12). Nevertheless, according to our simulations, checking all the 2^K SDR sub-problem can be finished within several minutes with $K \leq 10$. Meanwhile, when $K \geq N_{ant}$, in our simulations, the SDR problem in (19) is almost always infeasible and the PLS problem (12) is answered as infeasible instantly. Hence, the only challenging part for the heuristic algorithm is when N_{ant} is fairly large ($N_{ant} > 10$), which is unlikely to happen due to the physical limitation on the number of antenna elements of an antenna array in practical antenna engineering. In addition, while large N_{ant} may theoretically yield feasible solutions for K that is larger than 10, in practice, such feasible solutions are usually not usable in practical scenarios. This is because these solutions are very sensitive to the exact fine-tuning of the antenna pattern such that normal fuzziness in the radiation pattern due to reflection and diffraction of the radio signal usually invalidate their feasibility.

Following our SDR based heuristic algorithm, the PLS feasibility problem (12) could end up in three different cases. In the first case, the SDR problem in (19) is infeasible, which indicates that the PLS problem (12) is also infeasible. In the second case, the SDR problem is feasible and a feasible $\mathbf{w}$ to the PLS problem is obtained using our heuristic algorithm. Then it is sufficient that PLS in this case is feasible. In the third case, the SDR problem is feasible, but our heuristic algorithm returns no feasible $\mathbf{w}$ for (12). The feasibility of the PLS problem is unknown. Essentially, these three cases show that the feasibility answer to the PLS problem lies in between the feasibility answer to the SDR problem and the feasibility answer of our heuristic solution. Hence, the solution to the SDR problem and our heuristic solution can be seen as an upper bound and a lower bound for the PLS problem in (12), respectively. Fortunately, our experiments show that the third case appears very rarely. This essentially means that the feasibility answer to the PLS problem (12) is tightly bounded by the feasibility solution of the SDR problem and our heuristic algorithm.

VII. Simulation Results

In this section, we simulate the SDR relaxed problem and the heuristic algorithm described in Section VI to analyze the feasibility of PLS under different circumstances. Through the analysis, we discuss guidelines for deployment of localization systems and possible solutions to defend against beamforming-based PLS attacks. Our simulation also confirms that our relaxed SDR problem and the local-search heuristic algorithm provide very tight bounds on the PLS problem (12). The SDR problems in the simulations are solved using cvx [18] in MATLAB environment on a desktop with Intel 2.8G Hz CPU and 3Gb memory.

In the simulations, we also add the practical consideration about possible beamforming aiming error. In practice, there might be a small aiming error when the attacker points the beamforming pattern to the anchors. Thus, we require that the beamforming pattern for PLS also satisfies the following additional constraints,

$$\frac{1}{\delta} \leq \frac{v_k}{\hat{v}_k} |G(\theta_k \pm \gamma_\theta)|^2 \leq \delta, \;\; k = 1, 2, \cdots, K. \qquad (21)$$

where γ_θ represents a small angle of aiming error. Otherwise, only a small aiming error could fail the location spoofing attack. The following simulation results are obtained by assuming $\gamma_\theta = 1°$.

A. *Fixed Spoofing Distance*

First we analyze the feasibility of PLS attacks with fixed true location (x, y) and fake location $(\hat{x}, \hat{y})$ under different anchor deployments. This part of analysis shows the capability of PLS attacks. We randomly generate anchors in a $200 \times 200 m^2$ 2-D space. They are spaced reasonably far enough from each other to mimic real system deployment since a real localization system rarely have two anchors sit very close to each other. The path loss exponent α is set to be $\alpha = 3$. The attacker's true location is (0,0) and the fake location is (30, 40).

1) PLS beamforming pattern: An example of the beamforming pattern of a PLS attack is shown in Figure 3. From the figure, we can see how the beamforming pattern compensates the differences between the path loss vector from the true location (x, y) to the anchors and the path loss vector from the fake location $(\hat{x}, \hat{y})$. Towards the directions of anchors which are closer to the fake location than to the true location, the directional gain is comparatively large. On the contrary, for the anchors which are closer to the true location, the directional gain is much smaller. In this way, the overall effect of beamforming together with the natural path loss makes the RSS reading vector at the anchors seems as if they are generated by signal transmitted from the fake location.

2) Success rate of PLS: We use Monte Carlo simulations to estimate the success rate of PLS with different combinations of the number of antenna elements (N_{ant}) and the number of anchors (K) in the 2-D space. In each simulation run, the locations of the anchors are randomly generated. We also vary the value of the noise threshold δ from 1 dB to 3 dB. For each combination of parameters, totally 200 simulation runs are launched. The number of times where the local-search heuristic algorithm can find feasible solutions (N_{heuri}) and the situations where the relaxed SDR problem has solutions(N_{relax}) are recorded in Table I. Note that the heuristic algorithm provides a lower bound on the original PLS problem at (12), while the SDR solutions provide an upper bound. Hence, the number of times that the original PLS problem has feasible solutions lies in between N_{heuri} and N_{relax}.

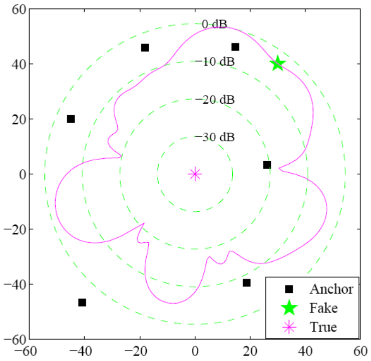


Fig. 3. A beamforming pattern of PLS.

TABLE I
NUMBER OF SUCCESSFUL CASES OUT OF 200 SIMULATION RUNS. (N_{heuri}/N_{relax})

δ	K	N_{ant}=6	N_{ant}=8	N_{ant}=10	N_{ant}=12
1dB	4	66/70	142/160	170/181	175/192
	5	4/7	78/97	93/152	106/182
	6	0/0	20/34	43/97	31/162
	7	0/0	0/5	16/64	9/95
	8	0/0	0/0	3/29	1/33
2dB	4	96/96	129/129	171/172	180/181
	5	10/11	105/107	148/148	165/170
	6	0/0	55/56	110/110	134/141
	7	0/0	15/15	81/84	97/107
	8	0/0	1/1	43/50	74/78
3dB	4	80/84	144/145	169/169	186/188
	5	15/16	117/120	148/152	170/176
	6	0/0	60/62	117/120	133/145
	7	0/0	18/20	99/100	100/107
	8	0/0	0/1	44/47	78/86

Three general trends in Table I can be observed. First, both N_{relax} and N_{heuri} increase when the number of antenna elements N_{ant} increases. The reason is that smart antenna array with more elements has more flexibility in tuning the radiation pattern, and hence is more capable in creating PLS attacks. Mathematically, more antenna elements leads to more tunable $[\mathbf{w}]_i$ and hence larger degree of freedom in the problems. Second, the success rates of the two problems tend to increase when the value of the threshold δ increases. This is because greater value of δ means looser bounds and larger feasible region of the PLS problem. Third, both N_{relax} and N_{heuri} decrease when K increases. Mathematically, adding anchors means adding extra constraints to problem (12), so that the feasible region decreases.

Another noticeable aspect of Table I is that the gap between N_{heuri} and N_{relax} is related to the noise threshold δ which indicates the natural variance in path loss. The smaller the δ, the larger the gap. Since the number of times that the

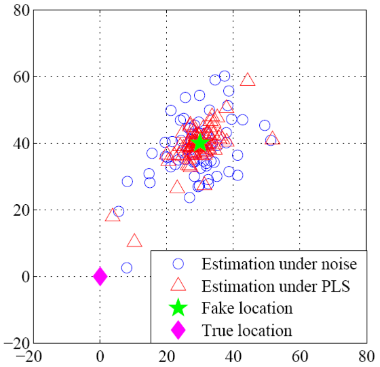


Fig. 4. Spoofed localization results v.s. noised localization results.

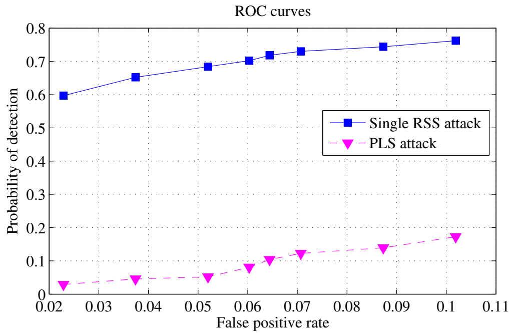


Fig. 5. ROC curves of attack detection.

PLS problems in (12) is feasible falls in between N_{heuri} and N_{relax}, smaller gap between N_{heuri} and N_{relax} indicates better approximate solutions to the PLS problem (12). Fortunately, in real environment, the natural path loss variance is much larger than 1dB (often in the range of 3-8 dB). Hence, we can expect that the gap between N_{heuri} and N_{relax} is very small in most of the real environments.

3) Location estimation under PLS attacks: Remember that when we formulate the PLS problem, we claim that the RSS readings of the attacker at anchors do not need to be exactly the same as the RSS readings of a legitimate user at the target fake location. As long as the RSS readings of the two situations are close enough, localization systems cannot tell location spoofing from normal variations in path loss. To demonstrate this point, we test least square estimation (LSE) as an example of location estimation algorithm. We compare the location estimates produced by the LSE algorithm under PLS attacks with that under normal path loss variations. For the PLS attacks, we chose the threshold as $\delta = 1$ dB, and the normal path loss variation is modeled as Gaussian noise with variance 1 dB. The output location estimation results for 200 simulation runs are shown in Figure 4. We can see that the localization results under PLS attacks are all around the fake location and within the error range of typical estimation results for legitimate users under normal path loss variations.

To get an insight into how PLS attacks are difficult to detect, we compare the detection rates under PLS attack and single RSS attack using the location spoofing attack detection algorithm introduced in [6], where the regression residuals of Linear Least Squares (LLS) are utilized to detect RSS attacks. The results after 1000 simulation runs are shown in Figure 5 and the settings for PLS attack is the same with Figure 4. The single RSS attack is simulated by adding a 30 dB bias to one of the RSS measurements. By comparing the ROC curves, we can see that the detection rates under PLS attacks are much lower and close to the false positive rates, which makes the attack detection algorithm almost invalid.

B. Fixed Anchor Deployment

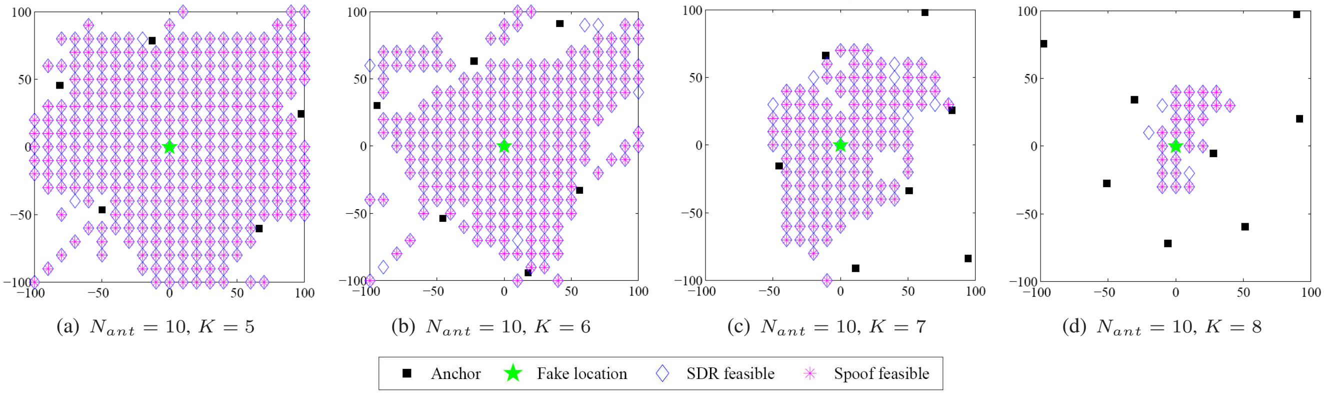
In this part, we investigate how the feasibility of PLS is affected by the anchor deployment and the relative position of the attacker's true location with respect to the anchors. In the simulations, the fake location $(\hat{x}, \hat{y})$ and the anchor locations are all fixed. We vary the true location (x, y) along a quare grid in the whole simulated 2-D space to identify the feasible locations where the attacker can launch PLS attacks.

A group of simulation results are shown in Figure 6. The parameters used in the simulations are $N_{ant} = 10$, $\delta = 2$ dB, $K = \{5, 6, 7, 8\}$. In Figure 6, the asterisk represents a true location (x, y) of the attacker where our heuristic algorithm has found feasible solutions to the PLS problem in (12). In other words, at the locations marked by asterisks, the attacker is able to successfully spoof his/her location to the fake location (shown as green pentagram in the center of the figure) using our heuristic algorithm. The blue diamonds represent true locations where the SDR problems are feasible. All the other locations on the grid without marks are locations where PLS is impossible. In the figure, most of the locations where SDR problems are feasible are also marked by asterisks (sign of feasible PLS problem). This implies that the answers to the PLS problem (12) are tightly bounded by solutions to the SDR problem and the solutions to our heuristic algorithm.

By comparing the four graphs in Figure 6, we find that when the anchor density increases, the attacker has less choices of locations where he/she can launch a PLS attack. This indicates that higher density of anchor deployment is effective in lowering the success rate of PLS attacks in the 2-D geometric space. Meanwhile, most of the asterisks are within an area which is bounded by the nearest surrounding anchors around the fake location. Thus anchor deployment with higher density not only lowers the success rate of PLS attacks, but also limits the possible hiding locations of an attacker to be close to the fake location. In other words, the attacker's feasible spoofing distance is smaller, and hence, the capability of location spoofing attacks is weakened. Thus, we conclude that high density anchor deployment is helpful to resist location spoofing attacks.

VIII. Guard Against Location Spoofing Attacks

From the simulation results, advices for spoof resistant anchor deployment in localization systems can be obtained. Generally, increasing the anchor density is good for guard against location spoofing attacks. However, uniformly increasing the anchor density could be expensive. It will be more efficient to design the anchor deployment according to the



(a) $N_{ant} = 10$, $K = 5$

(b) $N_{ant} = 10$, $K = 6$

(c) $N_{ant} = 10$, $K = 7$

(d) $N_{ant} = 10$, $K = 8$

Fig. 6. Geometrical statistic of location spoofing feasibility.

practical security requirements and specific characteristics of the application environment. For instance, in applications of LBAC, it is desirable to put more anchors around crucial areas where it is highly dangerous if an attacker can fake to be inside this area. These high-density anchors around crucial secure areas ensure that attackers are not able to spoof to the crucial locations. Another way to tackle PLS attack to RSS localization systems is to use mobile anchors. The mobility of anchors makes it extremely difficult for the attacker to locate the anchors. Hence, the attacker will not be able to figure out a desired beamforming pattern to launch PLS attack.

Our analysis is also helpful for tracing location spoofing attackers. When a PLS attack is discovered by means outside of localization systems, our analysis could help to narrow down the possible hiding locations of the attacker since PLS can only happen in certain areas. Furthermore, our analysis is also useful for finding the weakness of an anchor deployment and hence can help security personnel to effectively add extra anchors to prevent location spoofing attacks.

IX. Conclusion

In this paper, we analyzed the feasibility of PLS attacks using beamforming and investigated its relationship with the anchor deployment of localization systems. We utilized SDR technique and a heuristic local search algorithm to efficiently solve the PLS feasibility problem which is NP-hard in general. Simulation results show that our approach provides great approximation to the PLS feasibility problem. Observation on the simulation results indicates that localization system with higher anchor density performs better in resisting PLS attacks. Meanwhile, our approach is able to narrow down the possible hiding positions of PLS attackers. Our analysis provides insightful guidance for spoofing resistant anchor deployment and the proposed analytical approach can be used to trace location spoofing attackers as well as evaluate and improve the anchor deployment of localization systems.

Acknowledgment

This work is supported by National Science Foundation fund number ECCS-0802112 and the Institute for Critical Technology and Applied Science (ICTAS).

References

[1] C. A. Ardagna, M. Cremonini, E. Damiani, S. D. C. di Vimercati, and P. Samarati, "Supporting location-based conditions in access control policies," in *Proceedings of the 2006 ACM Symposium on Information, computer and communications security*, ser. ASIACCS, 2006.

[2] I. Ray and M. Kumar, "Towards a location-based mandatory access control model," *Computers & Security*, vol. 25, no. 1, pp. 36–44, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016740480500101X

[3] Y. Chen, W. Trappe, and R. P. Martin, "Detecting and localizing wireless spoofing attacks," in *Proceedings of IEEE SECON*, 2007.

[4] D. B. Faria and D. R. Cheriton, "Detecting identity-based attacks in wireless networks using singalprints," in *Proceedings of ACM Workshop on Wireless Security (WiSe)*, 2006.

[5] Y. Sheng, K. Tan, G. Chen, D. Kotz, and A. Campbell, "Detecting 802.11 MAC layer spoofing using received signal strength," in *Proceedings of IEEE INFOCOM*, 2008.

[6] Y. Chen, W. Trappe, and R. P. Martin, "Attack detection in wireless localization," in *Proceedings of IEEE INFOCOM*, 2007.

[7] K. Bauer, D. McCoy, E. Anderson, M. Breitenbach, G. Grudic, D. Grunwald, and D. Sicker, "The directional attack on wireless localization: how to spoof your location with a tin can," in *Proceedings of IEEE GLOBECOM*, 2009.

[8] J. H. Lee and R. Buehrer, "Location spoofing attack detection in wireless networks," in *Proceedings of IEEE GLOBECOM*, 2010.

[9] X. Li, Y. Chen, J. Yang, and X. Zheng, "Designing localization algorithms robust to signal strength attacks," in *Proceedings of IEEE INFOCOM*, 2011.

[10] P. Bahl and V. N. Padmanabhan, "Enhancements to the radar user location and tracking system," Microsoft Research, Tech. Rep., 2000.

[11] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *Proceedings of IEEE INFOCOM*, vol. 2, 2004, pp. 1012–1022.

[12] M. Robinson and I. Psaromiligkos, "Received signal strength based location estimation of a wireless lan client," in *Proceedings of IEEE WCNC*, vol. 4, 2005, pp. 2350–2354.

[13] S. Kim, H. Jeon, and J. Ma, "Robust localization with unknown transmission power for cognitive radio," in *Proceedings of IEEE MILCOM*, 2007, pp. 1–6.

[14] D. Han, D. G. Andersen, M. Kaminsky, K. Papagiannaki, and S. Seshan, "Access point localization using local signal strength gradient," in *Proceedings of Passive & Active Measurement (PAM)*, Seoul, South Korea, Apr. 2009.

[15] R. Vescovo, "Pattern synthesis with null constraints for circular arrays of equally spaced isotropic elements," in *IEE Proceedings of Microwaves, Antennas and Propagation*, vol. 143, no. 2, 1996, pp. 103–106.

[16] Z. Luo, W. Ma, A. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, may 2010.

[17] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, June 2006.

[18] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Apr. 2011.