Rogue Attribution Using Relative Signal Strength Based Location Estimation

Christine Laurendeau and Michel Barbeau
School of Computer Science, Carleton University
1125 Colonel By Drive, Ottawa, ON Canada K1S 5B6
Tel: 613-520-2600; Fax: 613-520-4334
E-mail: {claurend,barbeau}@scs.carleton.ca

Abstract

A rogue insider, in a wireless network, is an authenticated member that exploits possession of a valid identity in order to launch an attack. A typical example is the transmission of a verifiable message containing false or incomplete information. An important step, in enabling the network authorities to attribute an attack message to its originator, involves locating the physical source of the transmission. We propose a probabilistic scheme to determine the location of a transmitting rogue, with a degree of confidence, using the relative signal strength received by neighboring devices, even if the Effective Isotropic Radiated Power (EIRP) employed by the rogue is unknown. The relative received signal strength between pairs of trusted receivers are combined with a range of possible EIRP values to construct an area in Euclidian space bounded by maximum and minimum distance hyperbolas. The area contained within the intersection of multiple hyperbola pairs pinpoints the location of the rogue transmitter with a specific level of confidence.

Keywords

Wireless Security, Rogue Detection, Location Determination, Hyperbolic Localization

1 Introduction

Some of the most insidious security attacks are conducted by rogue insiders, and wireless networks are in no way immune. In a recent survey of IT security professionals [1], nearly half of the respondents reported that security breaches committed by malicious insiders engaging in corporate sabotage was a frequent occurrence. As the owner of a valid logical identity such as a MAC address or digital certificate in an open medium, the rogue insider in a wireless network can broadcast with impunity verifiable messages containing falsified information. It may also evade retribution once an attack is detected, especially if its identity is fraudulently obtained, for example through theft.

Before a rogue can be stopped, an attack must be detected. In some networks, unauthorized activity can be flagged through access control mechanisms or unusual usage patterns. However in many domains, for example vehicle safety applications where transmitted broadcast messages are digitally signed for authentication and non-repudiation, an attack may only be detected if an invalid digital certificate is used to sign the message. Attack broadcasts by a rogue insider can thus go unchallenged. In such technologies, additional mechanisms to expose attack messages are required.

Once an attack is detected, current means of attributing the attack to an insider node are based on its logical identity. This approach can be fraught with problems if the identity is forgeable. Dynamic MAC addresses are supported in some domains, for example in vehicular networks [2], to promote privacy. A rogue may easily abuse this feature to assume a new identity at will. Password-based access control and digital certificates associated with public/private key pairs constitute additional means of identification. However, in vehicular applications where messages may be signed but not encrypted, digital signatures have been deemed a critical threat to privacy for facilitating location tracking [3]. The current trend in securing this type of message in vehicular networks circumvents the use of individualized digital certificates for each node [4] [5] [6].

In a revocation scheme based solely on the originator's logical identity or other falsifiable credential, the network cannot attribute the source of an attack message directly to the rogue, nor can it prevent the attacker from using a new identity in subsequent attacks. Determining the physical location of the attack message's originator can be an important first step in apprehending the perpetrator, possibly linking its location to additional logical identities, and alerting neighboring devices to its presence in order to preemptively contain the impact of further attacks. An additional advantage to the use of a node's physical location as a unique unforgeable identity allows the network to detect and circumvent Sybil attacks, first described by Douceur [7], where one attacker may assume multiple colluding identities in order to mislead honest nodes.

We put forth a hyperbolic localization scheme for estimating the position of a transmitter using the relative Received Signal Strength (RSS) obtained by a number of trusted receivers. Since the Effective Isotropic Radiated Power (EIRP) of the transmitter is unknown and RSS values may fluctuate, a probabilistic model is used to estimate the range of distance differences between receiver pairs and the transmitter. These differences are used to construct minimum and maximum hyperbolas between two receivers where the transmitter may be located, with a certain degree of confidence. Multiple pairs of such hyperbolas may be constructed by considering different pairs of receivers. As a result, the intersecting hyperbolic area can be said to contain the transmitter with a combined degree of confidence obtained with a novel heuristic method developed for our threat model.

Section 2 outlines the existing literature in location determination. Section 3 describes our relative RSS based location estimation scheme. Section 4 evaluates the performance of our localization algorithm. Section 5 concludes the paper.

2 Related Work

The lion's share of existing research into wireless device location determination presumes the collaboration of the node being localized. Whether geometric localization is used, such as *Time of Arrival* (TOA), or the heuristic schemes commonly utilized in sensor networks [8], these techniques rely on the trusted cooperation of the node seeking to learn its position. However, in the case where the node is a rogue, any voluntarily supplied information may be falsified to enable the rogue to evade detection and retribution. As a result, we focus our efforts on the use of information inadvertently leaked by the rogue, such as the RSS associated with its transmitted message.

RSS-based localization algorithms come in two flavours: signature dependent or geometric. Signature dependent techniques rely on an existing training set of RSS signalprints established during an offline training phase. This map of known RSS measurements is subsequently consulted during the localization phase to estimate a node's location based on the similarity of the signals received at trusted base stations or access points to the signalprints found in the training set, as outlined in Bahl and Padmanabhan [9], Roos *et al.* [10] and Ladd *et al.* [11]. Such experiments have been conducted in indoor environments. Use in outdoor scenarios, for example in WiMAX/802.16 or vehicular networks, remains an open question. Geometric RSS-based localization techniques aim to estimate a node's location in Euclidian space based on the signal strength of messages received from trusted nodes within range, as proposed by Chong Liu *et al.* [12] and Bo-Chieh Liu *et al.* [13]. These schemes assume the complicity of the target node in reliably measuring the RSS of received beacon messages. As such, they cannot be used to localize a rogue node. Existing hyperbolic localization schemes are based on *Time of Arrival* (TDOA) techniques. They take an algebraic approach to estimate a device's coordinates by solving a set of non-linear hyperbolic equations, for example in Chan and Ho [14]. Signal strength fluctuations taken into account are restricted to small scale fading. As well, some of these techniques such as the ones outlined by Bo-Chieh Liu *et al.* [13] [15] assume a known distance from the transmitter to the closest receiver reference point.

Localization mechanisms that function independently of the targeted node are used in location verification and rogue detection. Location verification algorithms are predicated upon a set of trusted verifiers substantiating the alleged position of a prover within a specified area. The majority of existing schemes rely on TOA information collected by the verifiers to bound the prover's position, as outlined in Brands and Chaum [16], Sastry *et al.* [17] and Waters and Felten [18]. Xiao *et al.* [19] describe a signal strength based location verification method with an exhaustive set of possible rogue positions to identify Sybil nodes in vehicular networks. In the realm of rogue detection, Faria and Cheriton [20] outline a signature

dependent RSS-based scheme in an indoor environment to detect rogue mobile stations (MSs) in WiFi/802.11. Barbeau and Robert [21] employ RSS measurements in outdoor wireless access networks (such as WiMAX/802.16) to allow a MS to detect whether a base station (BS) advertising its availability for a handoff may be a rogue. A probabilistic RSS-based geometric model is presented where the RSS measurements obtained from neighboring BSs by a non-localized MS may be used to construct annuli whose non-empty intersection likely contains the MS. An empty intersection may indicate a RSS measurement originating from a rogue BS.

Our scheme extends the RSS-based geometric model proposed in [21] to localize a single transmitter from multiple receivers, taking into account the unknown EIRP employed by a rogue. In addition, since the compounding effect of a large EIRP range and probabilistic fluctuations in RSS measurements may result in a potentially extensive area for the rogue, we propose the use of hyperbola pairs rather than annuli for the localization. In this manner, relative rather than absolute RSS values are used to effectively reduce the area containing the rogue's probable location.

3 Localization Using Relative Signal Strength

We outline the assumptions inherent in our threat model and examine suitable propagation models for obtaining a signal source's location information from RSS values. We describe our scheme for estimating probabilistic minimum and maximum transmitter-receiver distances from RSS values and the subsequent computation of minimum and maximum bounds on the distance difference from a transmitter to a pair of receivers. Once a set of candidate transmitter areas is determined, we outline a heuristic method for combining the confidence level vested in each area to obtain the compounded probability distribution for the transmitter's location.

3.1 Threat Model

The goal of our localization scheme is to ascertain the source position of an attack message broadcast as a radio frequency (RF) signal to all receivers within its range. We assume the transmitting rogue is a mobile device and that a number of trusted receiving stations are within range and can communicate with each other over a secure channel to collect and aggregate RSS measurements. The coordinates of the receiving stations are globally known. Such a scenario is feasible in a number of wireless technologies, for example with WiMAX/802.16 [22] MSs and BSs or the On-Board Units and Road-Side Units in vehicular network architectures [23].

In order to evade revocation from the network, the rogue may combine changes in its transmission power with the judicious use of directional antennas to modify the signal gain and obfuscate its true position. As a result, no assumptions can be made regarding the EIRP employed to transmit the attack message.

3.2 Radio Propagation Models

RF signals are subject to attenuation as they propagate through the air. Large scale fading occurs when a signal encounters large terrain-based obstacles such as buildings, trees and hills. Small scale signal fading is caused by the movement of a mobile device. The cumulative effect of fading over time and distance is termed the *path loss*. A number of theoretical models for estimating path loss have been put forth in the literature for the purpose of simulating propagation environments, and these models may be classified into two categories. Empirical propagation models use probabilistic methods to predict received signal characteristics such as path loss and strength. Deterministic models are specific to a particular area and take into account the various obstacles therein. The dynamic nature of outdoor environments inherent to wireless access networks such as WiMAX/802.16 and vehicular networks lends itself better to empirical models rather than deterministic ones. As a result, we focus on the former approach.

Several empirical propagation models have been proposed for forecasting large scale path loss as a function of the distance between a transmitter and a receiver. These are of particular interest, since they lend themselves to our purpose. If the distances between the transmitter and receivers can be approximated from the path loss, which is directly proportional to the RSS values, the transmitter's location can be estimated. The Okumura model [24] predicts path loss based on transmitter and receiver antenna height, as well as the mean attenuation and the environment-based gain which can be obtained from experimental results. Miyashita *et al.* [25] observe that the Okumura model is unsuitable over

complex terrain due to the difficulty in ascertaining the required correction factors. The validity of the Hata model [26], also known as the Okumura-Hata model [25], has been demonstrated for frequencies between 150-1500 MHz, but not at the higher frequencies commonly used in newer technologies. For example, WiMAX/802.16 uses the 2-11 or 10-66 GHz bands, while DSRC vehicular networks operate in the 5.9 GHz band. The two-parameter Nakagami model [27] has been suggested as best suited for modeling channel characteristics in vehicular communications [28]. This model is dependent upon two parameters, the mean received power and a fading parameter, that are both obtained through experimental studies for a given discrete value of the distance d between a transmitter and a receiver. If d changes, so do the values of both parameters. As a result, the Nakagami model is unusable for predicting d from the measured path loss, since the parameters required to compute the path loss are dependent upon d. The log-normal shadowing model outlined by Rappaport [29] provides a simple model to measure large scale path loss from d. Since our aim is to approximate d based on a measured path loss, the log-normal shadowing model is best suited to our purpose.

3.3 Estimating Distance From Signal Strength

We outline Rappaport's large scale path loss model and describe how the minimum and maximum distance between a transmitter and a receiver can be computed from the RSS with a desired level of confidence.

3.3.1 The Log-Normal Shadowing Model

In [29], Rappaport outlines a *log-normal shadowing* model, which is a statistical path loss model for a signal received at distance d from a transmitter. This model is used to estimate the signal loss at various distances from the transmitter, based on a pre-defined reference distance d_0 close to the transmitter, a path loss exponent n dependent upon the propagation environment and the standard deviation σ for the path loss. Values for n and σ can be obtained from experimental measurements, for example in [30] and [31], where linear regression techniques are used to ascertain values of n and σ from actual path loss measurements.

In describing the log-normal shadowing model, Rappaport defines the path loss $L(d)$ of a signal at distance d as a Gaussian (Normal) distribution random variable with mean $\overline{L}(d)$ and standard deviation σ:

$$L(d) = \overline{L}(d) + X_\sigma$$

where X_σ is a Normal distribution zero-mean random variable with standard deviation σ. The mean path loss at distance d is in turn defined as:

$$\overline{L}(d) = \overline{L}(d_0) + 10n\log(\frac{d}{d_0})$$

where $\overline{L}(d_0)$ is the average path loss at the reference distance d_0, assuming free space propagation, and n is the path loss exponent. Rappaport concludes the following fact.

Fact 1. *The path loss $L(d)$ of a signal at distance d from the transmitter is defined as:*

$$L(d) = \overline{L}(d_0) + 10n\log(\frac{d}{d_0}) + X_\sigma$$

A further observation can be made about X_σ.

Fact 2. *For a selected confidence level $\mathcal{C}$, X_σ lies in the confidence interval $[-z\sigma dB, +z\sigma dB]$, where $z = \Phi^{-1}(\frac{1+\mathcal{C}}{2})$ and can be obtained from a Normal distribution table.*

From Facts 1 and 2, we can specify the probabilistic path loss more precisely.

Lemma 1. *The path loss $L(d)$ of a signal at distance d from the transmitter is defined with a confidence level $\mathcal{C}$ as:*

$$L(d) = \overline{L}(d_0) + 10n\log(\frac{d}{d_0}) \pm z\sigma$$

where $z = \Phi^{-1}(\frac{1+\mathcal{C}}{2})$.

Proof. This can be derived directly from Fact 1 and Fact 2. □

Example. Figure 1 illustrates an example of the distribution of path loss at a distance of $d = 100$ m from the transmitter. In the 2.4 GHz frequency band, the average free space path loss measured at $d_0 = 1$ m equal to 40 dB. For a path loss exponent of 2.76, we obtain the average loss at 100 m as $\overline{L}(100\text{ m}) = 95$ dB. With a standard deviation $\sigma = 5.62$, the shaded area in Figure 1 depicts 95% of the probability distribution around the average path loss. The path loss shadowing lies in the interval $[-1.96\times5.62\text{dB}, +1.96\times5.62\text{dB}] = [-11\text{dB}, +11\text{dB}]$ with probability 0.95, and so $L(100\,\text{m})$ is contained in the interval $[84\,\text{dB}, 106\,\text{dB}]$ with probability 0.95.

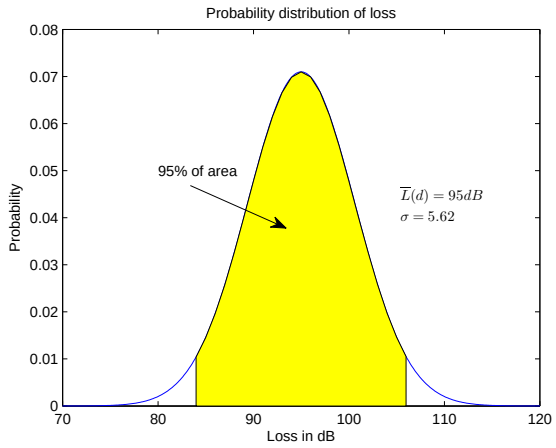

Figure 1: Example of Log-Normal Shadowing Model

3.3.2 Distance Range with Log-Normal Shadowing

In [21], Barbeau and Robert demonstrate that the minimum and maximum distances from a transmitter to a receiver can be calculated from the path loss using Rappaport's log-normal shadowing model.

Lemma 2. *For a chosen confidence level $\mathcal{C}$, the minimum and maximum distances (d^- and d^+ respectively) from a transmitter to a receiver can be computed as a function of path loss with:*

$$d^- = d_0 \times 10^{\frac{L(d)-\overline{L}(d_0)-z\sigma}{10n}}$$
$$d^+ = d_0 \times 10^{\frac{L(d)-\overline{L}(d_0)+z\sigma}{10n}}$$

Proof. The proof is based on Lemma 1 and can be found in [21]. □

In the threat model specified in Section 3.1, each receiver R_k obtains a RSS measurement RSS_k, but the path loss $L(d)$ value required to compute the distance using the equations in Lemma 2 is not readily available. We can thus replace the path loss with its equivalent, based on the transmitted power EIRP and RSS.

Fact 3. *The path loss $L(d)$ at distance d can be stated in terms of the EIRP and RSS at receiver R_k:*

$$L(d) = EIRP - RSS_k$$

A rogue may transmit at various EIRP levels or vary the signal gain with directional antennas in order to mask its location. In our solution, we address both issues by assuming an unknown value for the EIRP. We thus update Barbeau and Robert's minimum and maximum distance equations for a range of EIRP values, bounded by a minimum and a maximum EIRP, denoted as $\mathcal{P}^-$ and $\mathcal{P}^+$ respectively.

Lemma 3. *The minimum and maximum distances, d_k^- and d_k^+ respectively, between transmitter T and receiver R_k, sent within an estimated EIRP interval $[\mathcal{P}^-, \mathcal{P}^+]$, can be computed with confidence level $\mathcal{C}$ as:*

(1) $d_k^- = d_0 \times 10^{\frac{\mathcal{P}^- - RSS_k - \overline{L}(d_0) - z\sigma}{10n}}$

(2) $d_k^+ = d_0 \times 10^{\frac{\mathcal{P}^+ - RSS_k - \overline{L}(d_0) + z\sigma}{10n}}$

Alternately, we say that the probability that transmitter T is located in the area bounded by $[d_k^-, d_k^+]$ is $\mathcal{C}$:

$$Pr(d_k^- \leq T \leq d_k^+) = \mathcal{C}$$

Proof.

1. For a single EIRP value $\mathcal{P}$, Lemma 2 and Fact 3 can be combined to show that d_k^- and d_k^+ are the minimal and maximal distances respectively.
2. For a range of EIRP values $[\mathcal{P}^-, \mathcal{P}^+]$, let $\mathcal{D}_k(\mathcal{P}, \mathcal{V})$ represent the distance between a transmitter T and receiver R_k if the signal EIRP is $\mathcal{P}$ and the signal shadowing value within the shadowing interval $[-z\sigma, +z\sigma]$ is $\mathcal{V}$. Therefore, $\mathcal{D}_k(\mathcal{P}, \mathcal{V}) = d_0 \times 10^{\frac{\mathcal{P} - RSS_k - \overline{L}(d_0) + \mathcal{V}}{10n}}$.
   (i) $d_k^- = \mathcal{D}_k(\mathcal{P}^-, -z\sigma)$ is the minimum distance between T and R_k because it relies on the lower EIRP and signal shadowing bounds. We see that $\mathcal{D}_k(\mathcal{P}^-, -z\sigma)$ is less than or equal to $\mathcal{D}_k(\mathcal{P}^-, +z\sigma)$ since $-z\sigma$ is less than or equal to $+z\sigma$; it is less than or equal to $\mathcal{D}_k(\mathcal{P}^+, -z\sigma)$ since $\mathcal{P}^-$ is less than or equal to $\mathcal{P}^+$; and it is less than or equal to $\mathcal{D}_k(\mathcal{P}^+, +z\sigma)$ since $\mathcal{P}^-$ is less than or equal to $\mathcal{P}^+$ and $-z\sigma$ is less than or equal to $+z\sigma$.
   (ii) $d_k^+ = \mathcal{D}_k(\mathcal{P}^+, +z\sigma)$ is the maximum distance between T and R_k because it relies on the upper EIRP and signal shadowing bounds. We see that $\mathcal{D}_k(\mathcal{P}^+, +z\sigma)$ is greater than or equal to $\mathcal{D}_k(\mathcal{P}^+, -z\sigma)$ since $+z\sigma$ is greater than or equal to $-z\sigma$; it is greater than or equal to $\mathcal{D}_k(\mathcal{P}^-, +z\sigma)$ since $\mathcal{P}^+$ is greater than or equal to $\mathcal{P}^-$; and it is greater than or equal to $\mathcal{D}_k(\mathcal{P}^-, -z\sigma)$ as shown in (i). □

An additional observation may be gleaned from the results uncovered in Lemma 3.

Lemma 4. *For a given EIRP value $\mathcal{P}$, the minimum and maximum distances between a transmitter and a receiver are bounded solely by the signal shadowing range $[-z\sigma, +z\sigma]$ with confidence level $\mathcal{C}$.*

Proof. With a constant EIRP value $\mathcal{P} = \mathcal{P}^- = \mathcal{P}^+$, the proof can be directly inferred from Lemma 3, since $-z\sigma$ and $+z\sigma$ are the lower and upper bounds respectively of the signal shadowing range. □

3.4 Estimating Location From Distance

Minimum and maximum distances between a transmitter and a receiver have been used, for example in [21], [12] and [13], to construct a pair of rings forming an annulus within which a transmitter may be located. Multiple annuli may be computed around several receivers, and the location of the transmitter can be estimated within the annuli intersection area. However, this approach is more successful when the difference between minimum and maximum distances is not significant. If it is, the annuli may be so wide that their intersection is too large to effectively locate the transmitter, even if multiple receivers are considered.

Our approach relies on the use of the relative distance difference from a transmitter between pairs of receivers, similar to the *Time Difference of Arrival* (TDOA) technique. In TDOA, a hyperbola is constructed with two points of known coordinates at the foci. The properties of hyperbolas are such that every point on the hyperbola is at the same distance difference of the two foci. For example, if the difference in distances from a transmitter T to two receivers A and B is known, the corresponding hyperbola $\mathcal{H}_{A,B}$ can be constructed, as shown in Figure 2. The transmitter must necessarily lie on the hyperbola between A and B. If a second distance difference is known, for example between receivers B and C, a second hyperbola can be plotted, and the location of the transmitter T is discovered at the intersection of the two hyperbolas.

However in our threat model, because neither the transmitter EIRP nor its signal shadowing value are known, we cannot determine the precise distance difference between a pair of receivers. Instead, we use a TDOA-based technique combined with the estimated minimum and maximum distances between the transmitter and receivers. We can thus define a

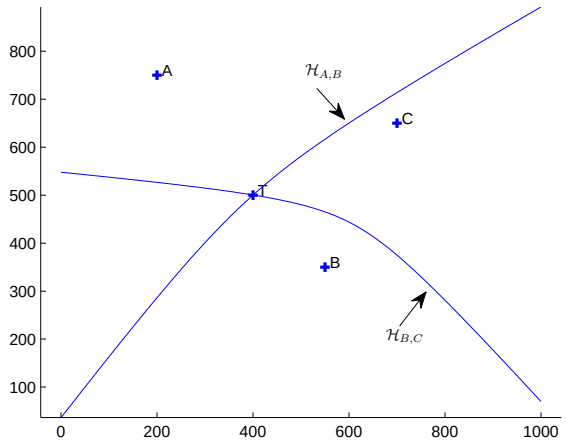

Figure 2: TDOA Example

candidate area $\mathcal{A}_{i,j}$ bounded by two hyperbolas between a pair of receivers R_i and R_j: one hyperbola at the minimum bound of the distance difference range and another at the maximum bound.

3.4.1 Computing the Distance Difference Range

We use the minimum and maximum distance equations, defined in Lemma 3, to compute the minimum and maximum bounds in the range of distance differences between a pair of receivers R_i and R_j. Intuitively, the minimum bound of this range from R_i's perspective is the difference between the distances closest to R_i yet farthest from R_j, at the minimum transmission power. This bound is typically located between the two receivers. In turn, the maximum bound of the range from R_i's perspective is the difference between the distances farthest from R_i yet closest to R_j, at the maximum EIRP, resulting in a bound that may be located beyond R_j.

Theorem 1. *Let d_i be the unknown distance between a transmitter T and receiver R_i.*

1. *The minimum bound $\Delta d^-_{i,j}$ of the distance difference range between d_i and d_j is the distance difference at the minimal EIRP ($\mathcal{P}^-$) over the full signal shadowing range $[-z\sigma, +z\sigma]$ with confidence level $\mathcal{C}$.*

$$(3)\quad \Delta d^-_{i,j} = d_0 \times 10^{\frac{\mathcal{P}^- - RSS_i - \overline{L}(d_0) - z\sigma}{10n}} - d_0 \times 10^{\frac{\mathcal{P}^- - RSS_j - \overline{L}(d_0) + z\sigma}{10n}}$$

2. *The maximum bound $\Delta d^+_{i,j}$ of the distance difference range between d_i and d_j is the distance difference at the maximal EIRP ($\mathcal{P}^+$) over the full signal shadowing range $[+z\sigma, -z\sigma]$ with confidence level $\mathcal{C}$.*

$$\Delta d^+_{i,j} = d_0 \times 10^{\frac{\mathcal{P}^+ - RSS_i - \overline{L}(d_0) + z\sigma}{10n}} - d_0 \times 10^{\frac{\mathcal{P}^+ - RSS_j - \overline{L}(d_0) - z\sigma}{10n}} \quad (4)$$

Proof. Because both R_i and R_j receive the same attack message sent with a single (albeit unknown) transmitted power, the EIRP value used to calculate the distance difference between R_i and R_j must be the same. Thus the minimum bound of the distance difference range uses $\mathcal{P}^-$, and the maximum bound uses $\mathcal{P}^+$. From Lemma 4, we know that for each EIRP value, the range of minimum and maximum distances must encompass the full signal shadowing range between $-z\sigma$ and $+z\sigma$ with confidence level $\mathcal{C}$.

Let $\mathcal{D}_k(\mathcal{P}, \mathcal{V})$ represent the distance between T and R_k if the signal EIRP is $\mathcal{P}$ and the shadowing is $\mathcal{V}$. Thus according to Lemma 3, $\mathcal{D}_k(\mathcal{P}, \mathcal{V}) = d_0 \times 10^{\frac{\mathcal{P} - RSS_k - \overline{L}(d_0) + \mathcal{V}}{10n}}$.

1. $\Delta d^-_{i,j} = \mathcal{D}_i(\mathcal{P}^-, -z\sigma) - \mathcal{D}_j(\mathcal{P}^-, +z\sigma)$ is the minimum bound of the distance difference between d_i and d_j because it relies on the lower EIRP bound and the full signal shadowing range. In other words, $\mathcal{D}_i(\mathcal{P}^-, -z\sigma) - \mathcal{D}_j(\mathcal{P}^-, +z\sigma)$ is less than or equal to $\mathcal{D}_i(\mathcal{P}^-, -z\sigma) - \mathcal{D}_j(\mathcal{P}^-, -z\sigma)$ since $-(+z\sigma)$ is less than or equal to $-(-z\sigma)$; it is less than or equal to $\mathcal{D}_i(\mathcal{P}^-, +z\sigma) - \mathcal{D}_j(\mathcal{P}^-, -z\sigma)$ since $-z\sigma$ is less than or equal to $+z\sigma$ and $-(+z\sigma)$ is less than or equal to $-(-z\sigma)$; and it is less than or equal to $\mathcal{D}_i(\mathcal{P}^-, +z\sigma) - \mathcal{D}_j(\mathcal{P}^-, +z\sigma)$ since $-z\sigma$ is less than or equal to $+z\sigma$.

2. $\Delta d^+_{i,j} = \mathcal{D}_i(\mathcal{P}^+, +z\sigma) - \mathcal{D}_j(\mathcal{P}^+, -z\sigma)$ is the maximum bound of the distance difference between d_i and d_j because it relies on the upper EIRP bound and the full signal shadowing range. In other words, $\mathcal{D}_i(\mathcal{P}^+, +z\sigma) - \mathcal{D}_j(\mathcal{P}^+, -z\sigma)$ is greater than or equal to $\mathcal{D}_i(\mathcal{P}^+, -z\sigma) - \mathcal{D}_j(\mathcal{P}^+, -z\sigma)$ since $+z\sigma$ is greater than or equal to $-z\sigma$; it is greater than or equal to $\mathcal{D}_i(\mathcal{P}^+, -z\sigma) - \mathcal{D}_j(\mathcal{P}^+, +z\sigma)$ since $+z\sigma$ is greater than or equal to $-z\sigma$ and $-(-z\sigma)$ is greater than or equal to $-(+z\sigma)$; and it is greater than or equal to $\mathcal{D}_i(\mathcal{P}^+, +z\sigma) - \mathcal{D}_j(\mathcal{P}^+, +z\sigma)$ since $-(-z\sigma)$ is greater than or equal to $-(+z\sigma)$. □

It should be noted that although $\Delta d^-_{i,j} = \mathcal{D}_i(\mathcal{P}^-, -z\sigma) - \mathcal{D}_j(\mathcal{P}^-, +z\sigma)$ is the minimum bound of the distance difference range, $\mathcal{D}_i(\mathcal{P}^-, +z\sigma) - \mathcal{D}_j(\mathcal{P}^-, -z\sigma)$ also encompasses the full shadowing range at the lower EIRP bound. The latter is used when the minimum hyperbola between the inverted order of receivers R_j and R_i is computed: $\mathcal{D}_i(\mathcal{P}^-, +z\sigma) - \mathcal{D}_j(\mathcal{P}^-, -z\sigma) = -[\mathcal{D}_j(\mathcal{P}^-, -z\sigma) - \mathcal{D}_i(\mathcal{P}^-, +z\sigma)] = -\Delta d^-_{j,i}$. Similarly, $\mathcal{D}_i(\mathcal{P}^+, -z\sigma) - \mathcal{D}_j(\mathcal{P}^+, +z\sigma) = -[\mathcal{D}_j(\mathcal{P}^+, +z\sigma) - \mathcal{D}_i(\mathcal{P}^+, -z\sigma)] = -\Delta d^+_{j,i}$.

3.4.2 Plotting the Minimum and Maximum Bound Hyperbolas

Given the definitions for the range of distance differences between a pair of receivers, we may construct the corresponding hyperbolas bounding the location of the transmitter.

Theorem 2. *Let a transmitter T be located at coordinates (x, y) and a pair of receivers R_i, R_j at coordinates (x_i, y_i) and (x_j, y_j) respectively. Let $\Delta d^-_{i,j}$ and $\Delta d^+_{i,j}$ be defined as the minimum and maximum bounds, respectively, of the distance difference range between R_i and R_j with confidence level $\mathcal{C}$. Let $\mathcal{H}^-_{i,j}$ be the hyperbola representing the minimum bound of the distance difference range between R_i and R_j, as defined by equation $\sqrt{(x-x_i)^2 + (y-y_i)^2} - \sqrt{(x-x_j)^2 + (y-y_j)^2} = \Delta d^-_{i,j}$. Let $\mathcal{H}^+_{i,j}$ be the hyperbola representing the maximum bound of the distance difference range between R_i and R_j, as defined by equation $\sqrt{(x-x_i)^2 + (y-y_i)^2} - \sqrt{(x-x_j)^2 + (y-y_j)^2} = \Delta d^+_{i,j}$.*

A transmitter T is located in the area $\mathcal{A}_{i,j}$ between the hyperbolas $\mathcal{H}^-_{i,j}$ and $\mathcal{H}^+_{i,j}$ with confidence level $\mathcal{C}$. Alternately, we say that $Pr(T \in \mathcal{A}_{i,j}) = \mathcal{C}$ and $Pr(T \notin \mathcal{A}_{i,j}) = (1 - \mathcal{C})$.

Proof. We define the distance between T and R_i as $d_i = \sqrt{(x-x_i)^2 + (y-y_i)^2}$ and the distance between T and R_j as $d_j = \sqrt{(x-x_j)^2 + (y-y_j)^2}$. If $\Delta d_{i,j} = d_i - d_j$ is defined as the distance difference between R_i and R_j, we obtain the equation for the hyperbola between R_i and R_j:

$$\sqrt{(x-x_i)^2 + (y-y_i)^2} - \sqrt{(x-x_j)^2 + (y-y_j)^2} = \Delta d_{i,j}$$

We know from Theorem 1 that $\Delta d^-_{i,j}$ is the minimum bound of the distance difference between d_i and d_j and that $\Delta d^+_{i,j}$ is the maximum bound of this difference with probability $\mathcal{C}$. We can therefore deduce that the probability of T being located in the area between $\mathcal{H}^-_{i,j}$ and $\mathcal{H}^+_{i,j}$ is $\mathcal{C}$ and the probability of T being located outside this area is $(1-\mathcal{C})$. □

An additional pair of minimum and maximum bound hyperbolas can be constructed between receivers R_i and R_j, namely the hyperbolas based on the inverted order of the receivers, R_j and R_i. Thus any pair of receivers can yield four hyperbolas to help determine the location of the transmitter. We have also noted in simulations that the maximum bound of the distance difference range between receivers is often too large for the corresponding hyperbola to be plotted to scale. However, it is still required to bound candidate hyperbolic areas for the transmitter.

3.4.3 An Example

Let us compute the candidate hyperbolic areas $\mathcal{A}_{i,j}$ and $\mathcal{A}_{j,i}$ for the location of a transmitter T with confidence level $\mathcal{C} = 0.95$, which yields the normal distribution constant $z = 1.96$. We assume a transmitter frequency of 2.4 GHz. For a reference distance $d_0 = 1$ m, we use the parameter values obtained by Liechty *et al.* [30] [32] for Line of Sight (LOS) propagation and a seven meter high antenna, where the path loss exponent is $n = 2.76$ and the standard deviation is $\sigma = 5.62$. The average path loss at d_0 is calculated with Friis' transmission equation for free space propagation [33], assuming isotropic transmitting and receiving antennas:

$$\overline{L}(d_0) = (\frac{4\pi f d_0}{c})^2 = 40 \text{ dB}$$

We assume an example layout as depicted Figure 3, where receivers R_1 and R_2 receive an attack message from a transmitter T with signal strength $RSS_1 = -79.20$ dBm and $RSS_2 = -74.27$ dBm respectively, corresponding to an actual transmitted EIRP of 30 dBm. Further, we model the EIRP range with $\mathcal{P}^- = 15$ dBm and $\mathcal{P}^+ = 40$ dBm. Equations (1) and (2) reveal that the transmitter T is located between 37 m and 1848 m from R_1 and between 24 m and 1225 m from R_2 with probability $\mathcal{C} = 0.95$. Using equations (3) and (4), we compute the minimum bound of the distance difference between d_1 and d_2 as $\Delta d^-_{1,2} = -115$ m and the maximum bound as $\Delta d^+_{1,2} = 1653$ m. Conversely, the minimum bound between d_2 and d_1 is calculated as $\Delta d^-_{2,1} = -205$ m and the maximum bound is $\Delta d^+_{2,1} = 930$ m. The minimum bound hyperbolas $\mathcal{H}^-_{1,2}$ and $\mathcal{H}^-_{2,1}$ associated with $\Delta d^-_{1,2}$ and $\Delta d^-_{2,1}$ respectively are depicted in Figure 3. The candidate areas for transmitter T include the area between $\Delta d^-_{1,2}$ and $\Delta d^+_{1,2}$, known as $\mathcal{A}_{1,2}$ and shown with dotted arrows, and the area between $\Delta d^-_{2,1}$ and $\Delta d^+_{2,1}$, named $\mathcal{A}_{2,1}$ and featured with dash-dotted arrows. T is located within $\mathcal{A}_{1,2}$ with probability 0.95 and within $\mathcal{A}_{2,1}$ with the same probability. A heuristic mechanism to combine both probabilities so that the evidence they provide is compounded is presented in Section 3.5.

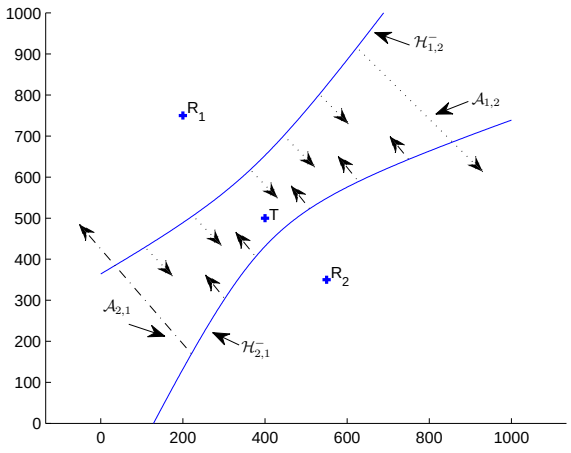

Figure 3: Minimum Hyperbolas for R_1, R_2

Minimum and maximum bound hyperbolas can be constructed between multiple pairs of receivers, forming a number of intersecting areas within which the transmitter location can be further bounded. For example, Figure 4 illustrates the minimum bound hyperbolas between receiver pairs R_1, R_2 and R_3, R_4.

3.5 Combining Location Estimation Confidence

Theorem 2 provides the means to estimate an area for the location of a transmitter, with a desired confidence level. However, given a number of receivers, hyperbolic areas can be constructed between multiple receiver pairs, and the confidence levels in all the areas can be combined together. We introduce a heuristic mechanism to compound the evidence contributed by multiple receiver pairs and probability in the intersecting areas.

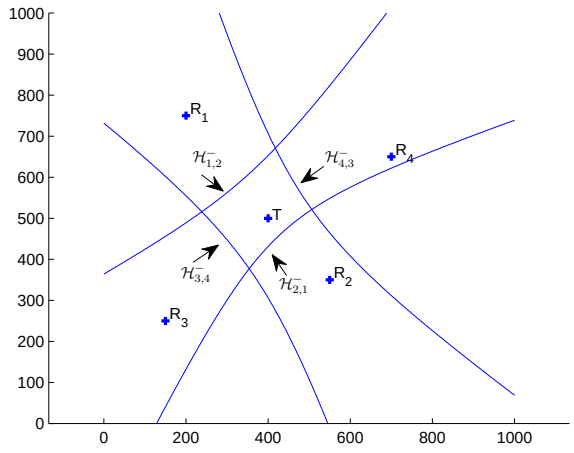

Figure 4: Minimum Hyperbolas for R_1, R_2 and R_3, R_4

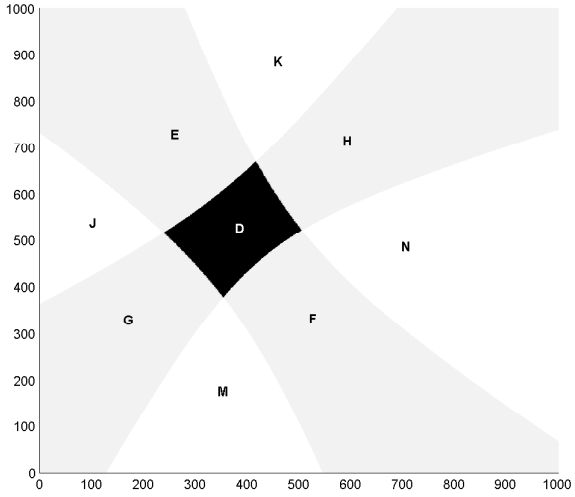

Figure 5: Example of Probability Areas

Definition 1. *Let $\mathcal{W}$ be the set of all defined hyperbolic areas computed using Theorem 2:*

$$\mathcal{W} = \{\mathcal{A}_{i,j} : \exists\, \mathcal{H}^-_{i,j}, \mathcal{H}^+_{i,j} \text{ and } \mathcal{A}_{i,j} \text{ comprises the area situated between } \mathcal{H}^-_{i,j} \text{ and } \mathcal{H}^+_{i,j}\}$$

where $n = |\mathcal{W}|$. Let ξ be the union of the Euclidian space where all hyperbolic areas in $\mathcal{W}$ overlap with each other and the space outside all the areas in $\mathcal{W}$.

We define disjunctive partitions S^k of ξ such that for all $0 \leq k \leq n$, the partition S^k contains the sub-areas of ξ situated at once within the intersection of k hyperbolic areas and within the intersection of the complements of the remaining $n-k$ hyperbolic areas in $\mathcal{W}$. More formally:

$$S^k = \left\{s_m : s_m = (\cap^k \mathcal{A}_{i,j}) \cap (\cap^{n-k} \overline{\mathcal{A}}_{i,j})\right\}$$

where $\cap^k \mathcal{A}_{i,j}$ is the intersection of any k hyperbolic areas and $\cap^{n-k} \overline{\mathcal{A}}_{i,j}$ is the intersection of the complement of the remaining $n-k$ hyperbolic areas.

Example. Figure 5 depicts the partitioning of the Euclidian space covered by the $n = 4$ hyperbolic areas $(\mathcal{A}_{1,2}, \mathcal{A}_{2,1}, \mathcal{A}_{3,4}, \mathcal{A}_{4,3})$ shown in Figure 4. To define partition G, for example, we compute the intersection of $k = 3$ hyperbolic areas $(\mathcal{A}_{1,2}, \mathcal{A}_{2,1}, \mathcal{A}_{4,3})$, which yields sub-area $D \cup G$. The complement of the remaining $n - k = 1$ hyperbolic area $\mathcal{A}_{3,4}$ consists of sub-area $J \cup G \cup M$. The intersection of both sub-areas defines partition G, since $(D \cup G) \cap (J \cup G \cup M) = G$. So we see that $G = \mathcal{A}_{1,2} \cap \mathcal{A}_{2,1} \cap \mathcal{A}_{4,3} \cap \overline{\mathcal{A}}_{3,4}$. In Figure 5, $S^4 = \{D\}$, $S^3 = \{E, F, G, H\}$, $S^2 = \{J, K, M, N\}$ and $S^1 = S^0 = \{\}$.

Having determined the number of hyperbolic areas in which each partition is situated, we assign a probability to each partition on this basis. The intuitive idea is that the more hyperbolic areas a partition belongs to, the more likely it is that the transmitter located in that partition, i.e. $Pr(T \in S^j) < Pr(T \in S^k)$ for all $0 \leq j < k \leq n$.

One method to compute the probability in a given partition involves multiplying together the probabilities assigned to the partition by every hyperbolic area. So if $Pr(T \in \mathcal{A}_{i,j}) = C$, as demonstrated in Theorem 2, then:

$$Pr((T \in \cap^k \mathcal{A}_{i,j}) \wedge (T \in \cap^{n-k} \overline{\mathcal{A}}_{i,j})) = C^k \times (1 - \mathcal{C})^{n-k}$$

However, this method fails to effectively compound the hyperbolic area confidence levels, as evidence from an increasing number of areas supporting the confidence in a given partition actually *decreases* the overall probability in that partition. Rather than supporting the existing evidence for a given intersecting partition, the simple multiplication method re-distributes the associated probability outside the partition. In the example in Figure 3, two hyperbolic areas inform us that the probability of the transmitter located in the intersection is $0.95^2 = 0.90$. But in Figure 4, the evidence provided by four hyperbolic areas locates the transmitter in a smaller area with probability $0.95^4 = 0.81$. In effect, additional evidence supporting the original finding has reduced the probability of the transmit-

ter's location in the intersecting area, leading to the counter-intuitive conclusion that the transmitter is more likely to be found in the intersection of two hyperbolic areas than in the intersection of four areas.

Other methods for combining evidence have been suggested in the literature, most notably the Dempster rule of combination [34], which was subsequently expanded in Dempster-Shafer's theory of evidence [35]. This combination method is best applied to models where the evidence is either incomplete or conflicting in the opinion of multiple observers. In those cases, Dempster's rule of combination compounds the probability assigned to uncertainty with the known evidence to support the latter. In our threat model, however, every pair of receivers assigns probability to the entire Euclidian space: the hyperbolic area is asserted with confidence level $\mathcal{C}$ and its complement with confidence level $(1-\mathcal{C})$. No area is assigned an unknown probability. In this case, Dempster's rule of combination reduces to the simple multiplication of probabilities previously outlined.

We propose an alternate, heuristic method for *compounding* evidence in such a manner that supporting evidence increases the likelihood of a transmitter located in a given area. The intuitive idea behind our heuristic method is based on the democratic principle. Given n hyperbolic areas, if all n agree that the transmitter is located in their common intersection with probability $\mathcal{C}$, then the corresponding partition consisting of this intersection is assigned probability $\mathcal{C}$. Of the remaining probability $(1-\mathcal{C})$, the proportion assigned to the intersection of $n-1$ hyperbolic areas is $\mathcal{C}$, for $\mathcal{C}\times(1-\mathcal{C})$. Of the remaining $(1-\mathcal{C})\times(1-\mathcal{C})$, $\mathcal{C}$ is assigned to the intersection of $n-2$ hyperbolic areas, and so on until the intersection of zero hyperbolic areas is assigned the remainder $(1-\mathcal{C})^n$.

Theorem 3. *Let ξ be the Euclidian space covered by n hyperbolic areas and their complements. Let the compounded probability $Pr^\diamond$ that a transmitter is located in a partition of ξ lying within k hyperbolic areas, for $k>0$, be the probability that it is situated within k hyperbolic areas, given the supporting evidence that it is within $k-1$ of these areas, combined with the probability that it is outside the remaining $n-k$ hyperbolic areas. In the case where the partition is outside of all hyperbolic areas, i.e. for $k=0$, the compounded probability reduces to the simple multiplication of probabilities. Thus:*

$$Pr^\diamond(T\in S^k)=\begin{cases}\mathcal{C}\times(1-\mathcal{C})^{n-k}, & \text{for } k>0\\ (1-\mathcal{C})^n, & \text{for } k=0\end{cases}$$

Proof.

1. (i) For $k>0$:

$$\begin{aligned}Pr^\diamond(T\in S^k) &= Pr(T\in\cap^k\mathcal{A}_{i,j}\mid T\in\cap^{k-1}\mathcal{A}_{i,j})\\ &\quad\times Pr(T\in\cap^{n-k}\ \overline{\mathcal{A}}_{i,j})\\ &= \frac{Pr(T\in[\cap^k\mathcal{A}_{i,j}]\ \cap\ [\cap^{k-1}\mathcal{A}_{i,j}])}{Pr(T\in\cap^{k-1}\mathcal{A}_{i,j})}\\ &\quad\times Pr(T\in\cap^{n-k}\ \overline{\mathcal{A}}_{i,j})\\ &\quad\text{by the definition of}\\ &\quad\text{conditional probabilities}\\ &= \frac{Pr(T\in\cap^k\mathcal{A}_{i,j})}{Pr(T\in\cap^{k-1}\mathcal{A}_{i,j})}\\ &\quad\times Pr(T\in\cap^{n-k}\ \overline{\mathcal{A}}_{i,j})\\ &\quad\text{since }\cap^k\mathcal{A}_{i,j}\subseteq\cap^{k-1}\mathcal{A}_{i,j}\\ &= \frac{[Pr(T\in\mathcal{A}_{i,j})]^k}{[Pr(T\in\mathcal{A}_{i,j})]^{k-1}}\\ &\quad\times[Pr(T\in\overline{\mathcal{A}}_{i,j})]^{n-k}\\ &= \frac{\mathcal{C}^k}{\mathcal{C}^{k-1}}\times(1-\mathcal{C})^{n-k}\\ &\quad\text{by Theorem 2}\\ &= \mathcal{C}\times(1-\mathcal{C})^{n-k}\end{aligned}$$

(ii) For $k=0$:

$$\begin{aligned}Pr^\diamond(T\in S^0) &= Pr(T\in\cap^0\mathcal{A}_{i,j})\\ &\quad\times Pr(T\in\cap^n\ \overline{\mathcal{A}}_{i,j})\\ &= Pr(T\in\xi)\times[Pr(T\in\overline{\mathcal{A}}_{i,j})]^n\\ &= 1\times(1-\mathcal{C})^n=(1-\mathcal{C})^n\\ &\quad\text{by Theorem 2}\end{aligned}$$

2. $Pr^\diamond(T\in S^k)$ is a probability distribution, since

$\sum_{k=0}^{n} Pr^{\diamond}(T \in S^k) = 1$.

$$\begin{aligned}\sum_{k=0}^{n} Pr^{\diamond}(T \in S^k) &= (1-\mathcal{C})^n \\ &\quad + \sum_{k=1}^{n}[\mathcal{C} \times (1-\mathcal{C})^{n-k}] \\ &= (1-\mathcal{C})^n + \mathcal{C} \times \sum_{k=0}^{n-1}(1-\mathcal{C})^k \\ &= (1-\mathcal{C})^n \\ &\quad + \mathcal{C} \times \left(\frac{1-(1-\mathcal{C})^n}{1-(1-\mathcal{C})}\right) \\ &= (1-\mathcal{C})^n + 1 - (1-\mathcal{C})^n \\ &= 1\end{aligned}$$

□

Example. The compounded probabilities for the example illustrated in Figures 4 and 5 are shown in Figure 6. The probability associated with the partition in $S^4 = \{D\}$ is 0.95, in $S^3 = \{E, F, G, H\}$ is $0.05 \times 0.95 = 0.0475$ and in $S^2 = \{J, K, M, N\}$ is $0.05^2 \times 0.95 = 0.002375$. If any partitions were associated with S^1, their probability would be $0.05^3 \times 0.95$, and the remainder would be assigned to S^0 with 0.05^4.

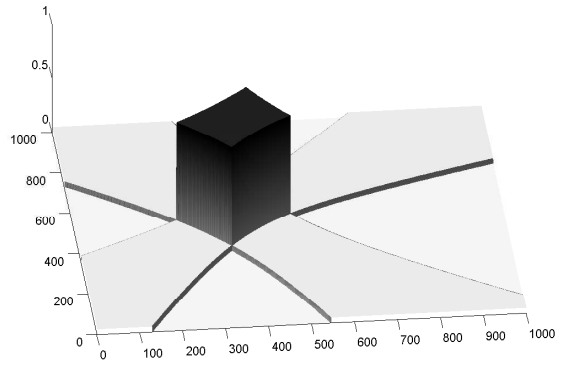

Figure 6: Compounded Probability Areas

4 Performance Evaluation

In this section, we outline the results obtained by simulating our localization algorithm using various transmitter locations with two scenarios involving fixed receivers. We also discuss the possible applications of our algorithm for access and vehicular networks.

4.1 Configuration

Results are presented for two separate scenarios, one with two receivers and the other with four receivers, located on a 1000×1000 meter grid. The transmitter location is simulated at each 100 meter interval in the grid from zero meters to 1000 meters on both the X-axis and Y-axis, as shown in Figure 7. The transmitter localization algorithm yields results for each of the possible 121 transmitter locations, given four individual confidence levels: $\mathcal{C} = 0.95$, 0.90, 0.85 and 0.80. Minimum and maximum bound hyperbolas are constructed between each pair of receivers. The simulation assumes a frequency of 2.4 GHz, as well as the values for the reference distance, path loss exponent and shadowing standard deviation determined for this frequency in an outdoor environment by Liechty *et al.* [30] [32], where d_0 equals one meter, n equals 2.76 and σ equals 5.62. A transmitter EIRP of 30 dBm is assumed for computing simulated RSS values at each receiver. For each execution, a random amount of signal shadowing is added to the RSS values along a Normal distribution, with mean zero and Liechty's shadowing standard deviation. The EIRP range is determined dynamically by taking the closest receiver to the transmitter location, i.e. the receiver with the lowest RSS, as a reference point. The EIRP range is set to the intersection of the EIRP ranges required for each remaining receiver to reach the reference point.

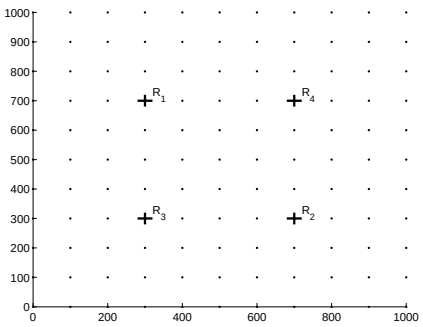

Figure 7: Four-Receiver Scenario Simulation Grid

4.2 Results

The performance of the transmitter localization algorithm is evaluated along two metrics: the success rate in correctly localizing the transmitter within a bounded area, and the minimization of the size of this area. Consequently, both metrics were gathered for each execution of the localization algorithm.

The success rate reflects the percentage of executions for which the intersection of all hyperbolic areas, i.e. the maximal probability candidate area, contains the actual transmitter location. The area size represents the percentage of the 1000×1000 grid covered by the candidate area. Optimal results are obtained when the success rate is maximized and the area size is minimized.

Figure 8 illustrates the success rate for the four-receiver scenario, given $\mathcal{C} = 0.95$. Since metrics were gathered solely for the grid points shown in Figure 7, values for the intermediate points were intrapolated linearly between the values for the computed grid points. The success rate is highest in the cross shape between the receivers, which we term the *aggregate range*. This range constitutes the zone in which the transmitter is located between at least one pair of receivers, enabling the receivers to aggregate and support each other's findings. The lowest success rates are achieved in the corners of the grid, outside the aggregate range, since these zones are not situated between any pair of receivers. Low success rates are also obtained when the transmitter is located precisely at the receiver locations. These special cases are eliminated from our subsequent analysis, because a zero meter distance is less than the reference distance of d_0 (one meter) used in the path loss model. Figure 8 displays the success rates averaged over the 1000 executions of the localization algorithm. With a confidence of 90%, the success rate associated with each grid point within the aggregate range lies in a confidence interval of $\pm 3\%$ of the mean for that point. The non-aggregate range points are situated in an interval of $\pm 4\%$ from the grid point mean, with confidence 90%. Consequently, not only is the success rate of the localization algorithm lower in the non-aggregate range, the results are also less reliable due to their greater variance. This bears out the intuition that a greater receiver coverage poses a significant advantage to the localization of a rogue device.

Figure 9 depicts the success rate as a function of the distance from the grid's midpoint, located at coordinates (500, 500), for the four-receiver scenario. This success rate is shown for each of the four confidence levels tested, with a 90% confidence interval depicted with each point for $\mathcal{C} = 0.95$. Intuitively, the farther the transmitter from the midpoint, the lower the expected success rate, and Figure 9 confirms this hypothesis for all confidence levels. At the midpoint, where the distance is zero, the highest success rate is achieved. However, because the higher success rates occur between the receivers in a

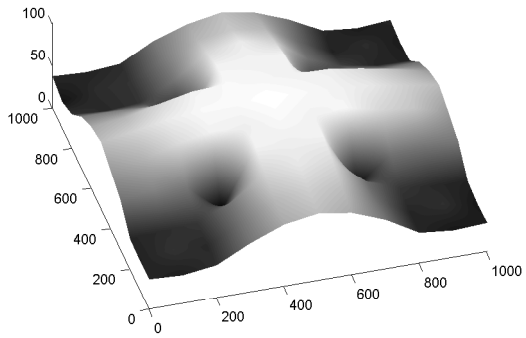

Figure 8: Success Rate for C = 0.95

cross shape rather than concentric circles, Figure 9 does not show a completely linear decrease in success rate. For example, noticeable dips in the graph occur at distances 361 m, 424 m, 500 m and 566 m. These correspond to the instances where the transmitter is located in the non-aggregate range at the corners of the grid, and thus within the range of only one receiver. Figure 10 captures the same data as Figure 9, but with the non-aggregate range points excluded. A more linear success rate is achieved as the transmitter location moves away from the midpoint.

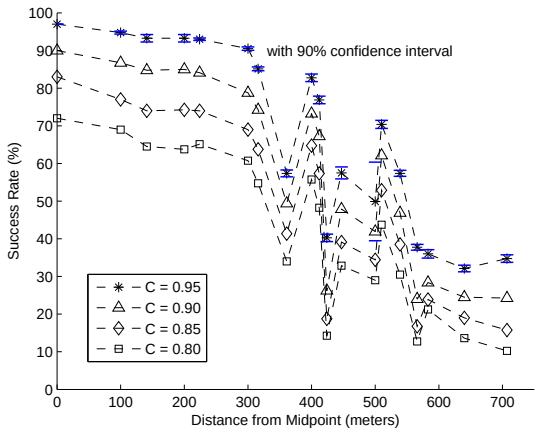

Figure 9: Success Rate for Distances from Midpoint

While the highest success rates are associated with higher confidence levels, so are the area sizes. Figure 11 shows the area size as a percentage of the total area depicted in the 1000×1000 meter grid for the simulations using the two-receiver and four-receiver scenarios. In each scenario, the area size decreases with the confidence level. This reflects the reduced value of the normalization constant z

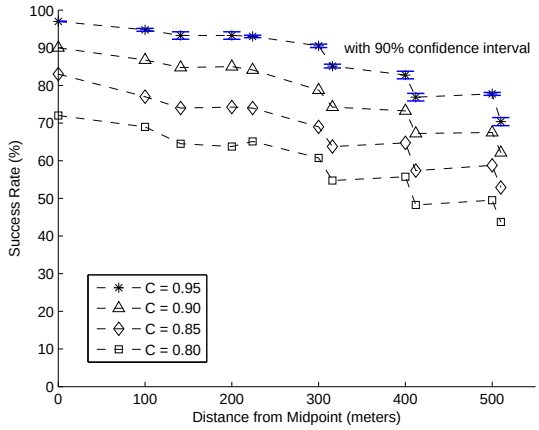

Figure 10: Success Rate for Aggregate Range Points

for lower confidence levels. The shadowing interval is therefore smaller, resulting in reduced hyperbolic areas whose intersections are correspondingly smaller. However, the candidate areas for the two-receiver scenario are significantly larger than those for the four-receiver scenario, reaching 62% of the total area in some instances. Clearly, this type of result is of very little use in locating an attacker. The four-receiver scenario yields more promising results, where even the 0.95 confidence level produces a candidate area on average below 25% of the total size of the grid.

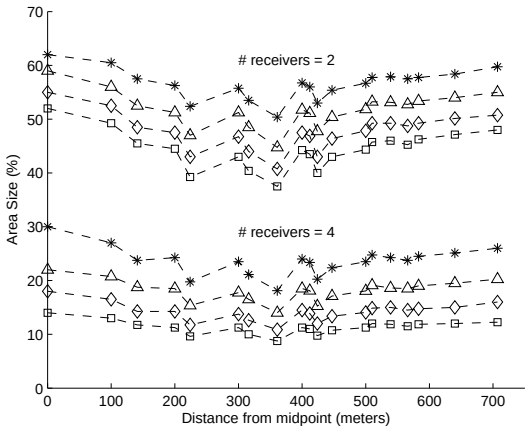

Figure 11: Area Size for Distances from Midpoint

Figure 12 illustrates the average area size for the success rates achieved with each confidence level in the four-receiver scenario. In general, the size of the candidate area is larger for given success rate as the confidence level increases. For example, a success rate of 80% yields a candidate area of 23% for $\mathcal{C} = 0.95$, an area size of 18% for $\mathcal{C} = 0.90$ and an area of 17% for $\mathcal{C} = 0.85$. Thus the average area size clearly decreases with the confidence level, due to the reduced shadowing interval.

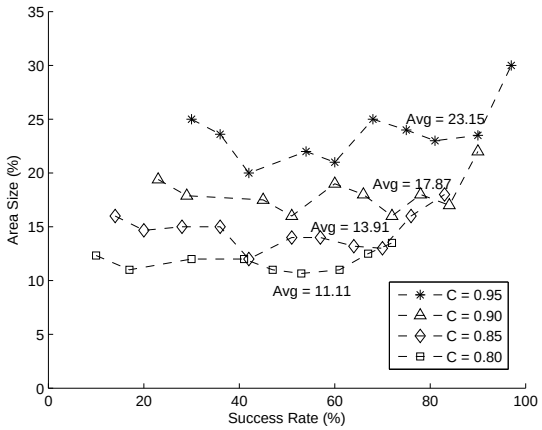

Figure 12: Area Size per Success Rate

4.3 Discussion

The localization algorithm can be applied to multiple types of wireless networks for the purpose of localizing a transmitting device that may not be trusted. Applications for wireless network security include the attribution of attack messages as originating from a particular candidate area which may contain devices whose identity is known to the trusted members of the network.

In access networks such as WiMAX/802.16, for example, BSs constitute trusted receivers which can collaborate with each other to localize a rogue. In such an infrastructure-based network, where each MS must register with the nearest BS, the candidate area determined by the localization algorithm can pinpoint the set of possible MSs among which the rogue is likely situated. If subsequent attack messages are generated, the rogue's location can be further narrowed. The use of directional antennas in WiMAX can serve to provide additional information for reducing the candidate area. For example, if an attack message is known to originate in a sector of a BS's range, the localization algorithm can be enhanced to only construct hyperbolic areas with other BSs whose range covers the same sector. If multiple BSs detect the attack message in a common sector, the size of the candidate area can be greatly reduced without negative impact on the success rate.

For vehicular networks, reductions in candidate

area size can be achieved using available navigation information, such as street and road layouts, building positions and so on. If a rogue device's location is restricted to the known navigable space, the size and shape of hyperbolic areas can be tailored to this layout. The receivers computing the hyperbolic areas can be comprised of trusted devices such as road-side units or state-controlled vehicles including police cars, ambulances and fire trucks. However, given the inherent mobility of vehicular network devices and dynamic nature of their topology, evidence may also be gathered from additional receivers such as private and commercial vehicles. While these do not enjoy the reliability implicit to the trusted devices, if we assume that the majority are honest, these additional receivers can greatly supplement the information gathered from the trusted devices in order to minimize the size of the candidate area while maximizing the rogue localization success rate.

5 Conclusion

In this paper, we presented a hyperbolic localization mechanism for attributing an attack message to an originating rogue device by estimating its position in a wireless network without the rogue's cooperation. The localization algorithm utilizes the relative RSS values of the attack message received at a set of trusted receivers to estimate the position of the transmitting device based on the aggregated RSS values, even though the transmitted EIRP power is unknown.

The scheme presented employs a large scale path loss statistical model to estimate the distances from the transmitter to a set of trusted receivers, with a selected confidence level. These distances are computed from the RSS values and yield a distance difference range between the transmitter and each pair of receivers. Hyperbolas are then constructed between each receiver pair at the minimum and maximum bounds of the distance difference range. The intersecting hyperbolic area between multiple pairs of receivers constitutes a candidate area for the location of the transmitting device with the given degree of confidence.

Performance evaluation through simulations revealed a success rate commensurate with the selected confidence level, although the size of the candidate area also increases with the success rate. The localization algorithm presented herein is sufficiently generic to be applicable to various types of wireless networks. Correspondingly, a confidence level of 95% yields an average candidate area size slightly below 25% of the simulated grid area. However, we foresee that each specific type of wireless network, such as WiMAX/802.16 access networks or vehicular networks, can exploit its particular characteristics to enhance the localization algorithm. It is expected that future research into the application of the localization algorithm to specific wireless technologies will result in the reduction of candidate area size for more precise rogue localization with higher success rates.

Acknowledgements

The authors gratefully acknowledge the financial support received for this research from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Automobile of the 21st Century (AUTO21) Network of Centers of Excellence (NCE).

References

[1] L. Ponemon, "National Survey on Managing the Insider Threat," Ponemon Institute, ArcSight White Paper, 2006. [Online]. Available: http://www.arcsight.com/

[2] IEEE 802 Committee of the IEEE Computer Society, "Draft Amendment to Standard for Information Technology - Telecommunications and Information Exchange Between Systems - LAN/MAN Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Wireless Access in Vehicular Environments (WAVE)," Draft IEEE Standard, IEEE P802.11p/D1.1, January 2005.

[3] C. Laurendeau and M. Barbeau, "Threats to Security in DSRC/WAVE," in *Proceedings of the 5th International Conference on Ad Hoc Networks and Wireless (ADHOC-NOW)*. Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 4104, 2006.

[4] K. Sampigethaya, L. Huang, M. Li, R. Poovendran, K. Matsuura, and K. Sezaki, "CARAVAN: Providing Location Privacy for VANET," in *Proceedings of the Conference on Embedded Security in Cars (ESCAR)*, 2005.

[5] M. Raya, A. Aziz, and J.-P. Hubaux, "Efficient Secure Aggregation in VANETs," in *VANET '06: Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks*. ACM Press, 2006, pp. 67–75.

[6] C. Laurendeau and M. Barbeau, "Secure Anonymous Broadcasting in Vehicular Networks," in *Proceedings of the 32nd IEEE Conference on Local Computer Networks*, 2007, pp. 661–668.

[7] J. R. Douceur, "The Sybil Attack," in *Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS)*. Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 2429, 2002.

[8] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free Localization Schemes for Large Scale Sensor Networks," in *MobiCom '03: Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*. ACM Press, 2003, pp. 81–95.

[9] P. Bahl and V. N. Padmanabhan, "RADAR: An In-building RF-based User Location and Tracking System," in *Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 2, March 2000, pp. 775–784.

[10] T. Roos, P. Myllymäki, H. Tirri, P. Misikangas, and J. Sievänen, "A Probabilistic Approach to WLAN User Location Estimation," in *International Journal of Wireless Information Networks*, vol. 9, no. 3, July 2002, pp. 155–164.

[11] A. M. Ladd, K. E. Bekris, A. Rudys, L. E. Kavraki, and D. S. Wallach, "Robotics-Based Location Sensing Using Wireless Ethernet," in *Wireless Networks*, vol. 11, no. 1–2, January 2005, pp. 189–204.

[12] C. Liu, K. Wu, and T. He, "Sensor Localization with Ring Overlapping Based on Comparison of Received Signal Strength Indicator," in *Proceedings of the IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, 2004, pp. 516–518.

[13] B.-C. Liu, K.-H. Lin, and J.-C. Wu, "Analysis of Hyperbolic and Circular Positioning Algorithms Using Stationary Signal-Strength-Difference Measurements in Wireless Communications," in *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, March 2006, pp. 499–509.

[14] Y. Chan and K. Ho, "A Simple and Efficient Estimator for Hyperbolic Location," in *IEEE Transactions on Signal Processing*, vol. 42, no. 8, August 1994, pp. 1905–1915.

[15] B.-C. Liu and K.-H. Lin, "Distance Difference Error Correction by Least Square for Stationary Signal-Strength-Difference-based Hyperbolic Location in Cellular Communications," in *IEEE Transactions on Vehicular Technology*, vol. 56, no. 5, 2007.

[16] S. Brands and D. Chaum, "Distance-Bounding Protocols (Extended Abstract)," in *EUROCRYPT '93: Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques*. Lecture Notes in Computer Science, Springer Berlin / Heidelberg, Volume 765, 1993, pp. 344–359.

[17] N. Sastry, U. Shankar, and D. Wagner, "Secure Verification of Location Claims," in *WiSe '03: Proceedings of the 2nd ACM Workshop on Wireless Security*. ACM Press, 2003, pp. 1–10.

[18] B. R. Waters and E. W. Felten, "Secure, Private Proofs of Location," Princeton University, Technical Report TR-667-03, 2003.

[19] B. Xiao, B. Yu, and C. Gao, "Detection and Localization of Sybil Nodes in VANETs," in *DIWANS '06: Proceedings of the 2006 Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks*. ACM Press, 2006, pp. 1–8.

[20] D. B. Faria and D. R. Cheriton, "Detecting Identity-based Attacks in Wireless Networks Using Signalprints," in *WiSe '06: Proceedings of the 5th ACM Workshop on Wireless Security*. ACM Press, 2006, pp. 43–52.

[21] M. Barbeau and J.-M. Robert, "Rogue-Base Station Detection in WiMax/802.16 Wireless Access Networks," in *Annals of Telecommunications*, vol. 61, no. 11–12, November–December 2006, pp. 1300–1313.

[22] LAN MAN Standards Committee of the IEEE Computer Society and the IEEE Microwave Theory and Techniques Society, "IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems - Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1," IEEE Std 802.16e-2005, 2005.

[23] National ITS Architecture, "DSRC 5GHz: Dedicated Short Range Communication at 5.9 GHz Standards Group." [Online]. Available: http://www.iteris.com/itsarch/html/standard/dsrc5ghz.htm

[24] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field Strength and its Variability in VHF and UHF Land-mobile Radio Service," in *Review of the Electrical Communication Laboratory*, vol. 16, no. 9–10, 1968, pp. 825–873.

[25] M. Miyashita, Y. Serizawa, and T. Terada, "Model Selection Method for Improving Path Loss Prediction of 400 MHz Band Land Mobile Radio," in *Proceedings of the 62nd IEEE Vehicular Technology Conference*, vol. 2, September 2005, pp. 1337–1341.

[26] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services," in *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, August 1980, pp. 317–325.

[27] M. Nakagami, "The m-Distribution – A General Formula of Intensity Distribution of Rapid Fading," in *Statistical Methods in Radio Wave Propagation*, W. Hoffman, Ed. Oxford, England: Pergamon, 1960, pp. 3–36.

[28] V. Taliwal, D. Jiang, H. Mangold, C. Chen, and R. Sengupta, "Empirical Determination of Channel Characteristics for DSRC Vehicle-to-Vehicle Communication," in *VANET '04: Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks*. ACM Press, 2004, p. 88.

[29] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice-Hall, 2002.

[30] L. Liechty, E. Reifsnider, and G. Durgin, "Developing the Best 2.4 GHz Propagation Model from Active Network Measurements," in *Proceedings of the 66th IEEE Vehicular Technology Conference*, September 2007, pp. 894–896.

[31] G. Durgin, T. S. Rappaport, and X. Hao, "Measurements and Models for Radio Path Loss and Penetration Loss In and Around Homes and Trees at 5.85 GHz," in *IEEE Transactions on Communications*, vol. 46, no. 11, November 1998, pp. 1484–1496.

[32] L. C. Liechty, "Path Loss Measurements and Model Analysis of a 2.4 GHz Wireless Network in an Outdoor Environment," Master's Thesis, Georgia Institute of Technology, 2007.

[33] H. T. Friis, "A Note on a Simple Transmission Formula," in *Proceedings of the I.R.E.*, vol. 34, no. 5, May 1946, pp. 254–256.

[34] A. P. Dempster, "Upper and Lower Probabilities Induced by a Multivalued Mapping," in *The Annals of Mathematical Statistics*, vol. 38, no. 2, April 1967, pp. 325–339.

[35] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.