An Information Theoretic Location Verification System for Wireless Networks

Shihao Yan¹, Robert Malaney¹
¹School of Electrical Engineering & Telecommunications
The University of New South Wales
Sydney, NSW 2052, Australia

Ido Nevat², Gareth W. Peters³,⁴
²Wireless & Networking Tech. Lab, CSIRO
³School of Mathematics and Statistics, University of NSW
⁴CSIRO Mathematical and Information Sciences
Sydney, Australia

Abstract—As location-based applications become ubiquitous in emerging wireless networks, a reliable Location Verification System (LVS) will be of growing importance. In this paper we propose, for the first time, a rigorous information-theoretic framework for an LVS. The theoretical framework we develop illustrates how the threshold used in the detection of a spoofed location can be optimized in terms of the mutual information between the input and output data of the LVS. In order to verify the legitimacy of our analytical framework we have carried out detailed numerical simulations. Our simulations mimic the practical scenario where a system deployed using our framework must make a binary Yes/No “malicious decision” to each snapshot of the signal strength values obtained by base stations. The comparison between simulation and analysis shows excellent agreement. Our optimized LVS framework provides a defense against location spoofing attacks in emerging wireless networks such as those envisioned for Intelligent Transport Systems, where verification of location information is of paramount importance.

I. INTRODUCTION

As Location-Based Services become widely deployed, the importance of verifying the location information being fed into the location service is becoming a critical security issue. The main difference between a Location Verification System (LVS) and a localization system is that we are confronted by some a priori information, such as a claimed position in the LVS [1]–[7]. In the context of a main target application of an LVS, namely Intelligent Transport Systems (ITS), the issue of location verification has attracted a considerable amount of recent attention [8]–[13]. Normally, in order to infer whether a network user or node is malicious (spoofs its claimed location) or legitimate (actually at its claimed location), we have to set a threshold for an LVS. This threshold is set so as to obtain a low false positive rate for legitimate users and a high detection rate for malicious users. As such, the specific value of the threshold will directly affect the performance of an LVS.

One traditional approach to set the threshold of an LVS is to search for a tradeoff between false positive and detection rates according to receiver operating characteristic (ROC) curve [14]. Another technique is to obtain the false positive and detection rates through empirical training data and minimize specific functions of the two rates to set the threshold [2] [4] [6]. For example, in [4], the sum of false positive and false negative rates were minimized. However, although successful in many scenarios, the approaches mentioned above do not specify in any formal sense what the ‘optimal’ threshold value of an LVS should be. In addition, in our key target application of our LVS, namely ITS, it is not practical to collect the required training data due to the variable circumstances.

The main point of this paper is to develop for the first time an information theoretic framework that will allow us to formally set the optimal threshold of an LVS. In order to do this, we first define a threshold based on the squared Mahalanobis distance, which utilizes the Fisher Information Matrix (FIM) associated with the location information metrics utilized by an LVS. To optimize the threshold, the Intrusion Detection Capability (IDC) proposed by Gu et al. [14] for an Intrusion Detection System (IDS) will be utilized. The IDC is the normalized mutual information (reduction of uncertainty) between the input and output data of an IDC. As such, the IDC measures the capability of an IDS to classify the input events correctly. A larger IDC means that the LVS has an improved capability of classifying users as malicious or legitimate accurately. From an information theoretic point of view the optimal threshold is the value that maximizes the IDC.

The rest of this paper is organized as follows. Section 2 presents the system model, which details the observation model and the threat model we utilize. In section 3, the threshold is defined in terms of the FIM associated with the location metrics. Section 3 also provides the techniques used to determine the false positive and detection rates, which are utilized to derive the IDC. Section 4 provides the details of how the IDC is used in the optimization of the threshold. Simulation results which validate our new analytical LVS framework are presented in Section 5. Section 6 concludes and discusses some future directions.

II. SYSTEM MODEL

A. A Priori Information: Claimed Position

Let us assume a user (legitimate or malicious) could obtain its true position, \( \theta_t = [x_t, y_t] \), from its localization equipment (e.g., GPS), and that the localization error is zero. Thus, a legitimate user’s claimed (reported) position, \( \theta_c = [x_c, y_c] \), is exactly the same as its true position \( \theta_t \). However, a malicious user will falsify (spoof) its claimed position in an attempt to fool the LVS. We denote the legitimate and malicious
hypothesis as $H_0$ and $H_1$, respectively, and the a priori information can be summarized as

$$\begin{align*}
H_0 : & \quad \theta = \theta_t, \quad \text{(Legitimate)} \\
H_1 : & \quad \theta \neq \theta_t, \quad \text{(Malicious)}. 
\end{align*}$$

\(1\)

**B. Observation Model based on \(H_0\)**

Although the framework we develop can be built on any location information metric, for purposes of illustration in this work we will solely investigate the case where the location information metric is the Received Signal Strength (RSS) obtained by a Base Station (BS) from a user. The RSS of the i-th BS from a legitimate user, $P^l_i$ (in dB), is assumed to be given by

$$P^l_i = P_0 - 10\gamma_0 \log_{10} \left( \frac{d^l_i}{d_0} \right) + w_\sigma, \quad \text{(2)}$$

where $P_0$ is a reference received power, $d_0$ is the reference distance, $\gamma_0$ is the path loss exponent, $w_\sigma$ is a zero-mean normal random variable with variance $\sigma^2_\sigma$, the Euclidean distance of the i-th BS to the user’s true position $[x_t, y_t]$ is

$$d^l_i = \sqrt{(x_t - x_B^i)^2 + (y_t - y_B^i)^2}, \quad i = 1, 2, \ldots, N,$$

where $[x_B^i, y_B^i]$ is the location of the i-th BS, and $N$ is the number of BSs. For $H_0$ in eq. (1), $d^l_i$ in eq. (2) can be replaced by $d^l_t$, where $d^l_t$ is the Euclidean distance of the i-th BS to the user’s claimed position $[x_c, y_c]$ and can be expressed as

$$d^l_t = \sqrt{(x_c - x_B^i)^2 + (y_c - y_B^i)^2}, \quad i = 1, 2, \ldots, N.$$

**C. Threat Model (Observation Model based on \(H_1\))**

Let us assume a malicious user knows the positions of all BSs, all the parameters in eq. (2), and is able to boost his transmit power according to his claimed position. The RSS of the i-th BS from a malicious user, $P^m_i$, can be written as

$$P^m_i = P_0 + \left| P_x \right| - 10\gamma \log_{10} \left( \frac{d^m_i}{d_0} \right) + w_\sigma, \quad \text{(3)}$$

where $P_x$ is the power adjustment. We assume the malicious user is equipped with only one omni-direction antenna, and thus $P_x$ is a constant.

The malicious user can adopt the following strategy to set a value for $P_x$. Constrained by the positions of all BSs, the observations $P^m_i$ are not exactly the same as the ideal observations $P_t$ calculated according to the malicious user’s claimed position as follows

$$P_t = P_0 - 10\gamma \log_{10} \left( \frac{d^m_t}{d_0} \right) + w_\sigma.$$ 

However, in order to avoid detection, the malicious user would like to adjust his transmit power so that all $P^m_i$ are as close as possible to the ideal observations $P_t$. Thus, the malicious user will set a value of $P_x$ to minimize the difference between $P^m_i$ and $P_t$. This difference can be defined by the Mean Square Error (MSE) as follows

$$\mathcal{D} = \frac{1}{N} \sum_{i=1}^{N} \left( E[P^m_i] - E[P_t] \right)^2 \quad \text{(4)}$$

$$\mathcal{D} = \frac{1}{N} \sum_{i=1}^{N} \left( P_x - 10\gamma \log_{10} \left( \frac{d^m_t}{d_0} \right) + 10\gamma \log_{10} \left( \frac{d^m_t}{d_0} \right) \right)^2,$$

where $E$ is the expectation operation. Then, the value of $P_x$ can be expressed as $P_x = \arg \min \mathcal{D}$. Taking the first derivative of $\mathcal{D}$ with respect to $P_x$ and setting it to zero, we can obtain

$$K^t = \frac{1}{N} \sum_{i=1}^{N} 10\gamma \log_{10} \left( \frac{d^m_t}{d_0} \right), \quad K^c = \frac{1}{N} \sum_{i=1}^{N} 10\gamma \log_{10} \left( \frac{d^c_t}{d_0} \right).$$

Substituting $P_x$ into eq. (3), the value of $P^m_i$ can be rewritten as

$$P^m_i = P_0 + P^l_i - K^c + w_\sigma, \quad \text{(4)}$$

$$P^m_i = K^l - 10\gamma \log_{10} \left( \frac{d^m_t}{d_0} \right).$$

Eq. (4) is based on the general threat model. However, this threat model is not practical since a malicious user’s true position is unknown. To make progress, we can assume a distribution for $\theta_t$. More specifically in this work, we assume a malicious user is effectively at infinity from all BSs (a more general case will be discussed later). Given this assumption, all BS distances from the user converge to one value. That is, the distance of a malicious user’s true position to every BS is effectively a constant number $d_{far}$, i.e., $d^m_i \approx d_{far}, \forall i \in [1, 2, \ldots, N]$. Therefore, the term $P^l_i$ can be approximated as zero. Accordingly, eq. (3) can be expressed as

$$P^m_i = P_0 - K^c + w_\sigma. \quad \text{(5)}$$

**III. THRESHOLD AND TWO RATES**

In this section, we first present our threshold determination based on the squared Mahalanobis distance (which utilizes the inverse FIM). Then, we provide techniques used to determine the false positive and detection rates of our LVS.

**A. Threshold**

The threshold is defined in terms of the squared Mahalanobis distance of an estimated position $\hat{\theta} = [\hat{x}, \hat{y}], \footnote{Note that an equivalent description of our LVS, which does not introduce the Mahalanobis distance, can be described in terms of the Cramer-Rao Lower Bound $\sigma_{\hat{\theta}}$. In this alternative description, an error ellipse is derived directly from the FIM, with the scale of the ellipse being set by $\sigma_{\hat{\theta}}$, and the orientation being set by the eigenvectors of the inverse FIM. For different values of the threshold $T$ the ellipse size scales as $T \sigma_{\hat{\theta}}$, and the detection algorithm decides the user is malicious if the estimated position returned by the location MLE lies outside of the ellipse.}$ The squared Mahalanobis distance can be expressed as [15]

$$\hat{D}_M = (\hat{\theta} - \theta)M^{-1}(\hat{\theta} - \theta)^T,$$
where $\hat{\theta} = [\hat{x}, \hat{y}]$ ($\hat{x}$ and $\hat{y}$ are the mean of $x$ and $y$, respectively), and $M$ is the covariance matrix of $\hat{\theta}$. According to the definition of $D_M$, it is a dimensionless scalar and involves not only the Euclidean distance but also the geometric information. In an LVS, we are interested in the 'distance' between a user’s estimated position $\hat{\theta}$ and its claimed position $\theta_c$. Thus, we will use $\theta_c$ instead of $\hat{\theta}$ to calculate $D_M$. In addition, without any a priori results from a localization algorithm, we can not obtain any estimate of the covariance matrix $M$. Therefore, we will use the inverse FIM, $M_e$, to approximate $M$. With this, the squared Mahalanobis distance in our LVS can be written as

$$D_M = (\hat{\theta} - \theta_c) M_e^{-1} (\hat{\theta} - \theta_c)^T.$$  
where $M_e = F^{-1}$ and $F$ is the FIM to be calculated as given below. In practice, the LVS works on the observation model based on $H_0$ with the likelihood function of received powers obtained using eq. (2). Let us assume the observations received by different BSs are independent, then the log-likelihood function can be expressed as

$$l(P^i|\theta_t) = -\frac{1}{2} \sum_{i=1}^{N} \left[ p_i^x - p_0 + 10 \log(d_i d_0) \right]^2 + \log C,$$

where $P^i = [P^i_1, P^i_2, \ldots, P^i_N]$ and the constant number $C$ is

$$C = \frac{1}{(2 \pi \sigma_{dB}^2)^{N/2}}.$$ 

Then, we can calculate the terms of the FIM through

$$F_{xy} = -E \left[ \frac{\partial^2 l(P^i|\theta_t)}{\partial x \partial y} \right],$$

where $E$ represents the expectation operation with respect to all observations. After some algebra, the FIM can be written as [16],

$$F = \begin{bmatrix}
    \frac{1}{2} \sum_{i=1}^{N} \frac{\sin^2 \varphi_i}{\sigma_i^2} & \frac{1}{2} \sum_{i=1}^{N} \frac{\sin 2 \varphi_i}{d_i^2} \\
    \frac{1}{2} \sum_{i=1}^{N} \frac{\sin 2 \varphi_i}{\sigma_i^2} & \frac{1}{2} \sum_{i=1}^{N} \frac{\cos^2 \varphi_i}{d_i^2}
\end{bmatrix},$$

where

$$b = \left( \frac{10 \gamma}{\sigma_{dB}^2 \ln 10} \right)^2,$$

$$\varphi_i = \arctan \frac{y_i - y_B}{x_i - x_B}.$$

After setting a threshold parameter $T$ for the squared Mahalanobis distance, the decision rule of an LVS (i.e. a malicious user or not) can be expressed as follows

$$\begin{cases}
D_M \leq T \Rightarrow H_0 \ (\text{Legitimate}) \\
D_M > T \Rightarrow H_1 \ (\text{Malicious}).
\end{cases} \quad (6)$$

Note that, we are able to transform any covariance matrix into a diagonal matrix by rotating the position vector [17]. Thus, the general form of $M_e$ can be expressed as

$$M_e = \begin{bmatrix}
    \sigma_x^2 & 0 \\
    0 & \sigma_y^2
\end{bmatrix}.$$ 

Then, the threshold $T$ can be encapsulated within the equation for an ellipse as follows

$$\frac{(x - x_e)^2}{T \sigma_x^2} + \frac{(y - y_e)^2}{T \sigma_y^2} = 1.$$ 

Therefore, the threshold $T$ can also be understood in the context of an ellipse, denoted as $\mathcal{T}$, which is determined by scaling the error ellipse provided by the FIM with the threshold parameter $T$.

Based on the above analysis, the overall process of our LVS includes four steps

- Collect observations of the RSS obtained by each BS from a user;
- Apply a localization algorithm to obtain an estimated position $\hat{\theta}$;
- Calculate the squared Mahalanobis distance $D_M$ of $\hat{\theta}$ to the user’s claimed position $\theta_c$;
- Infer if the user is legitimate or malicious according to the decision rule in eq. (6).

In practice, the above are all the steps of our LVS. However, to evaluate an LVS, false positive and detection rates, which are functions of the threshold parameter $T$ and other LVS parameters, are always investigated in theory. In the following subsections, we provide techniques used to determine false positive and detection rates in order to optimize the threshold parameter $T$.

### B. False Positive Rate

The false positive rate $\alpha$ is the probability by which legitimate users are judged as malicious. For a legitimate user, $\theta_c = \theta_t$. Then, in the 2-D physical space, the false positive rate can be expressed as $\alpha = e^{-t}$ [17].

In fact, the true positive rate $(1 - \alpha)$ is a well known metric that underlies the performance of unbiased localization algorithms. For example, in the 2-D physical space, it states that the probability by which an estimated position lies within the ellipse with $T = 1$ is no more than $39.35\%$.

### C. Detection Rate

The detection rate $\beta$ is the probability that malicious users are recognized as malicious. In order to calculate $\beta$, we have to obtain the posterior probability density function (pdf) for a location given the RSS observation vector, which can be expressed as

$$f(\theta|P^m) = \frac{f(P^m|\theta) f(\theta)}{f(P^m)},$$

where $\theta = [x, y]$ is a general location, and $P^m = [P^m_1, P^m_2, \ldots, P^m_N]$. Of course, if the user is malicious the observed signal vector $P^m$ will be one that has undergone a
is a metric that measures the capability of an IDS to classify the IDe. From an information theoretic point of view, the IDC optimization procedure is to find the value of the false positive rate (i.e., actually at his claimed location), and \( X = 1 \) represents the user is malicious (i.e., spoofs his claimed location). For the output, \( Y = 0 \) represents the LVS indicates the user is legitimate and \( Y = 1 \) represents the LVS indicates the user is malicious. Accordingly, the false positive rate \( \alpha \) is the probability \( \mathcal{P}(Y = 1|X = 0) \), and detection rate \( \beta \) is \( \mathcal{P}(Y = 1|X = 1) \). Therefore, the optimal value of \( T \) is that which maximizes the value of the \( C_{IDC} \) of the LVS.

The realizations of input and output data are denoted as \( z_x \) and \( z_y \), respectively. Given the base rate \( B \), the entropy of the input data \( H(X) \) can be written as [18]

\[
H(X) = \sum_{z_x} p(z_x) \log p(z_x) = -B \log B - (1 - B) \log (1 - B).
\]

The conditional entropy \( H(X|Y) \) can be expressed as [18]

\[
H(X|Y) = \sum_{z_x, z_y} p(z_x, z_y) \log p(z_x | z_y).
\]

Numerical methods are applied in order to search for the optimal value of \( T \) since there is no closed form for \( \beta \) in the above integral equation for \( \beta \) since there is no closed form solution. We note that based on the above analysis, \( \beta \) is a function of \( T \). As an aside it is worth mentioning that the false positive rate \( \alpha \) can also be written in a similar form as follows

\[
\alpha = 1 - \frac{1}{A_0} \int_{[x,y] \in T} f(\tilde{P}^m|\theta) f(\theta) dxdy,
\]

where \( \tilde{P}^m \) is the average value of \( P^m \) and \( A_0 = f(\tilde{P}^m) = \int_{[x,y]} f(\tilde{P}^m|\theta) f(\theta) dxdy, \)

where

\[
f(\theta|\tilde{P}^m) = \frac{f(\tilde{P}^m|\theta) f(\theta)}{f(\tilde{P}^m)}.
\]

IV. OPTIMIZATION OF THE THRESHOLD

In this section we will optimize the value of the threshold by maximizing the IDC, which is a function of the false positive rate \( \alpha \), detection rate \( \beta \) and the base rate \( B \) (the \textit{a priori} probability of intrusion in the input event data). That is, our optimization procedure is to find the value of \( T \) that maximizes the IDC. From an information theoretic point of view, the IDC is a metric that measures the capability of an IDS to classify the input events correctly and is defined as [14]

\[
C_{IDC} = \frac{I(X;Y)}{H(X)} = \frac{H(X) - H(X|Y)}{H(X)}, \tag{7}
\]

where \( H(X) \) is the entropy of the input data \( X \), \( I(X;Y) \) is the mutual information of input data \( X \) and output data \( Y \), and \( H(X|Y) \) is the conditional entropy. The mutual information \( I(X,Y) \) measures the reduction of uncertainty of the input \( X \) given the output \( Y \). Thus, \( C_{IDC} \) is the normalized mutual information (reduction of uncertainty) between the input and output data. Its value range is [0, 1]. A larger \( C_{IDC} \) value means that the IDS has an improved capability of classifying input events accurately.

Our LVS can be modeled as an IDS whose input data are binary numbers representing whether a user is legitimate or malicious, and the output data are the binary decisions. Then, for the input, \( X = 0 \) represents the user is legitimate and \( X = 1 \) represents the user is malicious (i.e., actually at his claimed location), and \( X = 1 \) represents the user is malicious (i.e., spoofs his claimed location). For the output, \( Y = 0 \) represents the LVS indicates the user is legitimate and \( Y = 1 \) represents the LVS indicates the user is malicious. Accordingly, the false positive rate \( \alpha \) is the probability \( \mathcal{P}(Y = 1|X = 0) \), and detection rate \( \beta \) is \( \mathcal{P}(Y = 1|X = 1) \). Therefore, the optimal value of \( T \) is that which maximizes the value of the \( C_{IDC} \) of the LVS.

The numerical methods are applied in order to search for the optimal value of \( T \) since there is no closed form for \( \beta \). In the following we refer to this optimal value as \( T_0 \).

V. SIMULATION RESULTS

Adopting a Maximum Likelihood Estimator (MLE) in our location estimation algorithm, we now verify our previous analysis via detailed simulations. The theoretical and simulated \( \alpha \), \( \beta \) and \( C_{IDC} \), all of which are dependent on \( T \), are utilized in order to find the value \( T_0 \) that maximizes \( C_{IDC} \).

A. Simulation Set-up

The simulation settings are as follows:

- \( N \) BSs are deployed in a 200m \( \times \) 200m square field and the legitimate and honest users can communicate with all BSs;
- The claimed positions of honest and malicious users are the same, denoted \( \theta_c \).
B. \( \alpha, \beta, C_{IDC} \) with Different Values of \( T \)

As shown in Fig.1, the solid lines are the theoretical \( \alpha, \beta \) and \( C_{IDC} \) while the symbols are the simulated \( \alpha, \beta \) and \( C_{IDC} \). The simulated values of \( \alpha \) and \( \beta \) are calculated directly from the realizations of estimated positions, and then the simulated \( C_{IDC} \) is obtained from eq. (7). The simulation parameters \( (\gamma, \sigma_{dB}, S, \theta_e) \) are shown in the figure captions (note that in all the figures explicitly shown in this paper the four BSs are fixed at the corners of a 200m x 200m grid). From Fig.1, we see that the theoretical optimal value \( T_0 \) can be seen to be 4.75. We also see that the comparison between simulation and analysis shows excellent agreement.

Beyond the simulations explicitly shown in Fig.1, we have investigated a range of other fixed BSs positions (up to 10 BSs where BSs positions are randomly simulated), and these simulations also show excellent agreement with analysis. Collectively, these simulation results verify the analysis we have provided earlier.

The simulation results with a malicious user having a certain distance to all BSs are shown in Fig.2. The true position of the malicious user in the simulations is set at 10km away from the claimed position. Although the simulation and theoretical values of \( \alpha, \beta \) and \( C_{IDC} \) do not match with each other exactly (the theoretical analysis approximates the user as being at infinity), the simulation and theoretical optimal values \( T_0 \) are effectively the same. We find this result holds down to distance where the malicious user is a few km away from the claimed position. This shows that our framework is tenable when the assumption that malicious user is infinitely far away is relaxed down to the few km range.

In order to verify the \( C_{IDC} \) with the optimal value \( T_0 \) is correct, we also simulated \( C_{IDC} \) for a range of \( \sigma_{dB} \). Fig. 3 shows such results for the case where the malicious user is effectively at infinity. Here the optimal value \( T_0 \) is derived from the proposed theoretical analysis, but in the simulations the threshold is set to the other values of \( T \) shown \((2T_0 \) and \( 0.5T_0)\). From the results shown we can see that these other values do provide simulated false positive and detection rates which result in lower values of \( C_{IDC} \) (and therefore suboptimal performance), which once again verifies the robustness of our analytical framework. Fig. 4 shows the same results except that the malicious user is again set at \( 10km \) away from the claimed position. Again we see a validation of our analysis.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a novel and rigorous information theoretic framework for an LVS. The theoretical framework we have developed shows how the value of the threshold used in the detection of a spoofed location can be optimized in terms of the mutual information between the input and output data.

In order to verify the legitimacy of our framework we have carried out detailed numerical simulations of our framework under the assumption of an idealized threat model in which the malicious user is far enough from the claimed location such that his boosted signal strength results in all BSs receiving the same RSS (modulo noise). Our numerical simulations mimic the practical scenario where a system deployed using our framework must make a binary Yes/No “malicious decision” to each snapshot of RSS values obtained by the BSs. The comparison between simulation and analysis shows excellent agreement. Other simulations where we modify the approximation of constant RSS at BSs also showed very good agreement with analysis.

The work described in this paper formalizes the performance of an optimal LVS system under the simplest (and perhaps most likely scenario), where a single malicious user
attempts to spoof his location to a wider wireless network. The practical scenario we had in mind whilst carrying out our simulations was in an ITS where another vehicle is attempting to provide falsified location information the wider vehicular network.

Future work related our new framework will include the formal inclusion of more sophisticated threat models, where the malicious user is closer to the claimed location and has the use of colluding adversaries. It is well known that no LVS can be made foolproof under the colluding adversary scenario, however, we will investigate in a formal information theoretic sense the detailed nature of the vulnerability of an LVS under such different threat models.

ACKNOWLEDGMENTS

This work has been supported by the University of New South Wales, and the Australian Research Council, grant DP120102607.

REFERENCES


2Note that location verification in the context of quantum communications systems has previously been considered e.g. [19], [20], [21], and it has been argued that such systems are able to securely verify a location under all known threat models [22] - although see [23] who argue otherwise. It is undisputed that classical communications alone cannot achieve secure location verification under all known threat models.