Abstract—In vehicular networks, vehicle claimed positions should be independently verified to help protect the wider network against location-spoofing attacks. In this work, we propose a new solution to the problem of location verification using the Cramer-Rao Bound (CRB) on location accuracy. Compared to known-optimal solutions, our technique has the advantage that it does not depend on a priori information on the probability of any vehicle being malicious. To analyze the performance of our new solution, we compare its operation, under Received Signal Strength (RSS) inputs, with a known optimal solution for this problem that assumes the probability of a vehicle being malicious is known. The results show that our new solution provides close to optimal performance under a wide range of anticipated channel conditions. Our solution is simple to deploy and can easily be integrated into a host of vehicular applications that use location information as an input.

I. INTRODUCTION

Location verification of mobile nodes has attracted a great deal of attention in recent years. Applications such as routing protocols, cellular systems, tracking of mobile devices, location-enabled security, and Vehicular Ad Hoc Networks (VANETs) are only some of the applications which would benefit from accurate location verification [1], [2].

In VANETs, location spoofing can lead to catastrophic outcomes if it leads to vehicle collisions [3]–[6]. A malicious vehicle in such networks could also spoof its location to seriously disrupt other vehicles or to greedily increase its functionality within the network. However, safety issues are perhaps the biggest reason behind the need for the location verification within VANET infrastructure. Recently, several information-theoretic Location Verification Systems (LVSs) for VANETs have been proposed (e.g., [7]–[11]). These studies have investigated the detection performance of an LVS in a range of possible VANET scenarios and configurations.

However, known LVS solutions, in spite of being optimal in an information-theoretic sense, depend on the a priori probability of a user being malicious. In practice, the knowledge of such a priori information is usually low.

In this work, we propose a new solution to the problem of location verification, using the CRB on location accuracy, that does not depend on the a priori probability of a vehicle being malicious. Given this new algorithm we are then interested in the question: How sub-optimal our location verification algorithm is without using a priori location information of a vehicle being malicious? We address this question by comparing the performance of our proposed solution, under RSS inputs, with a known optimal solution that assumes the probability of a vehicle being malicious is known.

To make progress, we first derive CRBs under the log-normal model for the RSS measurements between the Base Stations (BSs) and vehicles. We then investigate location verification under the quantitative impact of path-loss and shadowing variance. Although the precision level of RSS is smaller than of other metrics such as the time-of-arrival, RSS is simpler and less costly to deploy in terms of the hardware requirements and power consumption [12]–[14].

As we shall see, the location verification performance depends on the number of BSs and the geometry (relative position) of the BSs [15]–[18]. The use of CRBs also helps us understand the importance of the relative positions of the BSs on the position accuracy and position verification [19]–[21]. To this end, a CRB analysis provides insight into the issues surrounding location verification.

The remainder of this paper is organized as follows. Section II details our location verification algorithm. In Section III, the CRB analysis on the estimated location accuracy of a vehicle is derived. Section IV presents our location verification binary decision algorithm. In Section V, we present our numerical results on the performance of our solution relative to known-optimal LVS solutions. The known solutions are based on a Likelihood Ratio Test (LRT) which assumes the a priori probability of a vehicle being malicious is known. Finally, Section VII draws concluding remarks.

II. LOCATION VERIFICATION ALGORITHM

In this section, we present our proposed location verification algorithm. Our algorithm is designed such that it seamlessly integrates into any geographic routing protocol for VANETs. We assume, in a network, we have $N$ BSs with known location coordinates, $\theta_i = [x_i, y_i]$, and a vehicle (legitimate or malicious) with unknown location coordinates $\theta_0 = [x_0, y_0]$. We adopt a model for the RSS observations given by log-normal shadowing

$$P_r[dBm] = P_T[dBm] - (P_0[dB] + 10\gamma\log_{10}\left(\frac{d}{d_0}\right)) + X_{\sigma_m},$$  

where $P_r[dBm]$ is a received power at distance $d$, $P_T[dBm]$ is a transmitted power, $P_0[dB]$ is a reference received power corresponding to a reference distance $d_0$, $\gamma$ is the path loss exponent, $X_{\sigma_m}$ is a zero-mean normal random variable with

$$X_{\sigma_m} = \text{Normal}(0, \sigma_m),$$
where 
\[ (x_i - x_0)^2 + (y_i - y_0)^2 = d_i^2, \]
with a total of \( N \) equations. To solve (2), a reference BS can be chosen. All other equations are subtracted from the equation which incorporates the reference BS. Let 
\[ k_i = x_i^2 + y_i^2, \]
and let \( l \) denote the reference BS, after some algebra, (2) can be written as
\[ \sum_{i=1}^{N} x_i (x_i - x_l) + y_i (y_i - y_l) \theta_i = \sum_{i=1}^{N} d_i^2 - k_l + k_i. \]
We note that (4) can further be solved using linear algebra
\[ \theta_0 = \theta^T \theta c. \]
Putting (4) into (5) and applying the least squares method gives the possible solution
\[ \theta_0 = (\theta^T \theta)^{-1} \theta^T c, \]
where \([\cdot]^T\) denotes the matrix transpose. With these equations at hand, it is possible to authenticate the location of a vehicle in relation to the reference BS’s known coordinates. In the next section we find the lower bound on the covariance of our estimated location, \( \theta_0 \) by applying a CRB analysis.

III. CRB ANALYSIS

For a vehicle at an unknown position, the conditional probability distribution \( f_{\mathbf{S}_i} \) of RSS measurements \( S_i \), at BS \( i \), is given by (ignoring leading constants)
\[ -\ln f_{S_i|\theta_0, \theta} = \frac{(S_i + 10 \gamma \log_{10}(d_i/d_0))^2}{2 \sigma_{\mathbf{d}}^2}. \]

In a network with \( N \) BSs, the logarithm of the joint conditional probability density function can be written as
\[ \Gamma(S_i|\theta_0, \theta_i) = \sum_{i=1}^{N} \ln f_{S_i|\theta_0, \theta_i} (S_i|\theta_0, \theta_i). \]

To find the maximum likelihood estimate for \( \theta_0 \), we have to maximize (8) which is possible by taking the gradient derivative of (8) with respect to \( x_0 \) and \( y_0 \) given as
\[ \mathbf{J}_{x_0, y_0} = -\mathbf{E} \left[ \frac{\partial^2 \Gamma(S_i|\theta_0, \theta_i)}{\partial x_0 \partial y_0} \right]. \]
From this relation, we can find the terms of the Fisher information matrix (FIM) given as
\[ \mathbf{C}(\theta_0) = \left[ \begin{array}{ccc} I(\theta_0)_{x_0 x_0} & I(\theta_0)_{x_0 y_0} & I(\theta_0)_{x_0} \\ I(\theta_0)_{x_0 y_0} & I(\theta_0)_{y_0 y_0} & I(\theta_0)_{y_0} \\ I(\theta_0)_{x_0} & I(\theta_0)_{y_0} & I(\theta_0) \end{array} \right] = \mathbf{J}_{x_0, y_0}^{-1}. \]

The CRB is then the trace of the FIM matrix \( \mathbf{C}(\theta_0) \) given as
\[ \sigma_{\mathbf{d}}^2 = I(\theta_0)_{x_0 x_0} + I(\theta_0)_{y_0 y_0}. \]

IV. LOCATION VERIFICATION BASED BINARY DECISION ALGORITHM

In this section, we present our location verification binary decision algorithm as depicted in Fig. 1.

Inputs:
- Claimed location, \((x_c, y_c)\)
- Estimated location, \((x_0, y_0)\)
- Area bound by the error ellipse

Output: Binary decisions are \( D_0 \) and \( D_1 \)
- \( D_0 = 1, D_1 = 0 \), the reporting vehicle is legitimate
- \( D_0 = 0, D_1 = 1 \), the reporting vehicle is malicious

Body:
- If an estimated location is inside the area bounded by the standard error ellipse, i.e., \( \frac{(x_0-x_c)^2}{\lambda_1^2} + \frac{(y_0-y_c)^2}{\lambda_2^2} < 1 \), then the reporting vehicle is marked as legitimate. Here \( \lambda_1 \) and \( \lambda_2 \) are the eigenvalues of the covariance matrix \( \mathbf{C}(\theta_0) \), \( q = [-2 ln(1-p)] \) is the scale of the error ellipse, and \( p = 1 - e^{-\frac{1}{2} q^2} \) is the probability of an estimated location positioned inside the error ellipse.
- If an estimated location is outside the area bounded by the standard ellipse, i.e., \( \frac{(x_0-x_c)^2}{\lambda_1^2} + \frac{(y_0-y_c)^2}{\lambda_2^2} > 1 \), then the reporting vehicle is marked as malicious.
- If an estimated location is on the perimeter of the error ellipse, i.e., \( \frac{(x_0-x_c)^2}{\lambda_1^2} + \frac{(y_0-y_c)^2}{\lambda_2^2} = 1 \), in this case, we assume that the reporting vehicle is marked as malicious.
The area of the error ellipse can be calculated as 
\[ A = \pi \sigma_x \sigma_y, \]
where 
\[ \sigma_x = \sqrt{\lambda_1}, \]
and 
\[ \sigma_y = \sqrt{\lambda_2}, \]
is the length of semi-major axes and the length of semi-minor axes of the error ellipse. Similarly, the probability that an estimated location is outside (or on) the error ellipse boundary is given as 
\[ p_{o} = e^{-\frac{1}{2}d^2}. \]
The rotation matrix is given as
\[ R = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}, \]
where 
\[ 0 < \phi \leq 2\pi. \]
The scaling matrix is given as
\[ M = \begin{bmatrix} m_x & 0 \\ 0 & m_y \end{bmatrix}. \]
Using the FIM, we draw the error ellipse as
\[ C(\theta_0)R = RL, \]
where 
\[ R \]
is the rotation matrix defined above, which gives the orientation of error ellipse, whose columns are eigenvectors, and 
\[ M \]
is the scaling matrix or the diagonal matrix whose non zero elements are eigenvalues. Therefore, the FIM matrix 
\[ C(\theta_0) \]
can be represented as a function of the eigenvalues and eigenvectors, that is
\[ C(\theta_0) = RLR^{-1}. \]

The binary decision algorithm given in [23] is based on the LRT and considers the a priori likelihood of the RSS based LVS functions of \( \omega \). Moreover, the false positive rate can be calculated as 
\[ \alpha(\omega) = P(\gamma(n) \geq \omega|H_0), \]
and the detection rate can be calculated as 
\[ \beta(\omega) = P(\gamma(n) \geq \omega|H_1). \]
By minimizing the Bayesian average cost the specific value of \( \omega \) can be determined [24]. Here, we adopt the Bayesian average cost as the performance metric, which is known as the total error. The total error is a combination of the fraction of the false positive rate and missed detection rate. This can be written as
\[ \epsilon(\omega) = (1 - P_r)\alpha(\omega) + P_r(1 - \beta(\omega)), \]
where \( 1 - P_r \) and \( P_r \) are the a priori probabilities that the vehicle is legitimate and malicious, respectively. Based on the Bayesian framework, the optimal value of \( \omega \) that minimizes \( \epsilon(\omega) \) is given by \( \omega = P_r/(1 - P_r) \). We adopt in this work \( P_r = 0.5 \).

V. NUMERICAL RESULTS

In this section, we present numerical results to verify the accuracy of our provided analysis. We have conducted detailed Monte Carlo simulations of randomly selected BSs and calculated, on average, the difference in the total error (between the two binary decision algorithms) as our performance metric. In our simulations specifically shown here, the BSs and the claimed positions are deployed in a rectangular area 250m \( \times \) 250m. The claimed positions of both legitimate and malicious vehicles are generated randomly, however, the claimed position of a legitimate vehicle is also its true location. Each BS randomly collects 100 measurements from legitimate and malicious vehicles. The transmit power is set to \( P_T = 30 \) [dBm], the reference power is set to \( P_0 = -10 \) [dB] at \( d_0 = 1m \), and fading caused by multipath propagation and shadowing is averaged out.

The formal error ellipses constructed from the CRB analysis gives us the theoretical bound on the variance of some
The binary decision algorithm positions are outside the error ellipses which show that the true position of the vehicle is at a distance of 100m from its claimed position. We find in this case our binary decision algorithm in position. It can be seen that most of the estimated positions of the vehicle is at a distance of 55m from its claimed position. This small variation is due to the intrinsic GPS errors. The results in Fig. 5 show that when a malicious vehicle is far from its claimed position it is hard for both solutions (for the LRT based LVS and for our proposed location verification solution) to distinguish if it is a legitimate or malicious vehicle. The average of 100 trials is taken), $\sigma_{dB}^2 = 2$, and the path-loss is $\gamma = 3$.

In Fig. 5, we show the average total error calculated with $N$ BSs ($4 \leq N \leq 10$) randomly positioned. On the x-axis is the difference in distance between the claimed position and the estimated position. The shadowing variance is $\sigma_{dB}^2 = 2$, and the path-loss is $\gamma = 3$.

Fig. 2. Positions estimated by the system with four BSs for a claimed position (blue diamond). The true position of the legitimate vehicle is indicated by the red triangle. The shadowing variance is $\sigma_{dB}^2 = 2$ and the path-loss is $\gamma = 3$.

Fig. 3. Positions estimated by the system with four BSs for a claimed position (blue diamond). The true position of the malicious vehicle is 55m away from its claimed position (indicated by the blue triangle). The shadowing variance is $\sigma_{dB}^2 = 2$ and the path-loss is $\gamma = 3$.

Fig. 4. Positions estimated by the system with four BSs for a claimed position (blue diamond). The true position of the malicious vehicle is 100m away from its claimed position (indicated by the red triangle). The shadowing variance is $\sigma_{dB}^2 = 2$ and the path-loss is $\gamma = 3$.

Fig. 5. Average total error for different error ellipses and the LRT-based LVS with $N$ BSs ($4 \leq N \leq 10$) with random positions. On the x-axis is the difference in distance between the claimed position and the estimated position. The shadowing variance is $\sigma_{dB}^2 = 2$, and the path-loss is $\gamma = 3$.

Using (22), we have carried out detailed Monte Carlo calculations of the average total error for different error ellipses i.e., $1\sigma$, $2\sigma$, and $3\sigma$ confidence intervals and the LRT based LVS. In these simulations, we have used a varying number of BSs that are positioned randomly, with different values of shadowing variance, $\sigma_{dB}^2$ and path-loss exponents, $\gamma$.

In Fig. 5, we show the average total error calculated with $N$ BSs ($4 \leq N \leq 10$) randomly positioned (for each $N$ the average of 100 trials is taken), $\sigma_{dB}^2 = 2$, and $\gamma = 3$. The results in Fig. 5, show that when a malicious vehicle is at a 10m distance from its claimed position it is hard for both solutions (for the LRT based LVS and for our proposed location verification solution) to distinguish if it is a legitimate or malicious vehicle. The average total error is large when a malicious vehicle is near its claimed position and the average total error is small when a malicious vehicle is far from its claimed position. When the difference in claimed and true position is large both decision algorithms can easily detect malicious vehicles.

Different from Fig. 5, in Fig. 6, we extend our simulations estimated positions at some claimed position. These ellipses depend on the shadowing variance, path-loss exponent, and the standard deviation i.e., $\sigma_x$, $\sigma_y$, and $\sigma_{dB}$. We show in Fig. 2, 3, and 4, several estimated (and corresponding claimed) positions of a vehicle with four BSs (with fixed positions) deployed, with $\sigma_{dB}^2 = 2$ and $\gamma = 3$. In Fig. 2, a legitimate vehicle reports its position to all BSs in range. The reported (i.e. claimed) position is compared with the estimated positions. Here, the claimed position of the vehicle is 10m away from its true position. This small variation is due to the intrinsic GPS errors. We find for this case our binary decision algorithm always verifies the vehicle as a legitimate vehicle.

In Fig. 3, the reporting vehicle is malicious. The true position of the vehicle is at a distance of 55m from its claimed position. It can be seen that most of the estimated positions in $1\sigma_x$, $1\sigma_y$, $2\sigma_x$, $2\sigma_y$, and in $3\sigma_x$, $3\sigma_y$ are outside the error ellipse. We find in this case our binary decision algorithm mostly detects the vehicle as malicious.

In Fig. 4, the reporting vehicle is again malicious. The true position of the vehicle is at a distance of 100m from its claimed position. It can be seen in all three cases, the estimated positions are outside the error ellipses which show that the reporting vehicle is malicious. The binary decision algorithm in this case always detects the malicious vehicle successfully.
and show the average total error with additional averaging over a range of $\sigma_{2B}^2$ from $2 \leq \sigma_{2B}^2 \leq 5$ and a range of $\gamma$ from $2 \leq \gamma \leq 10$. Here, we find our solution and the LRT-based solution differ in the total error by 0.12, 0.09, 0.07, at a 10m distance and 0.10, 0.07, 0.03, at a 250m distance for different error ellipses i.e., $1\sigma$, $2\sigma$, and $3\sigma$, respectively. These values quantify, at different ellipse thresholds, the small sub-optimality of our solution relative to the LRT based LVS.

It is evident from the above results that our proposed solution with no a priori information about a vehicle being legitimate or malicious is close to optimal relative to the LRT based LVS. Our proposed solution is designed such that it seamlessly fits into a location verification-based hybrid routing protocol recently proposed by us [25]. It will also be straightforward to deploy in any geographic routing protocol.

VI. CONCLUSIONS

This work investigated the sub-optimality of a proposed new solution to the problem of location verification. The new solution, based on the CRB for the RSS-derived location accuracy in VANETs, is simple to deploy in that it requires less dependency on unknown information. We proved the efficiency of our solution by comparing it with an optimal LRT based LVS decision solution. The results obtained through detailed Monte Carlo simulations showed that the average total error of our location verification solution compared to an optimal LRT based LVS decision solution is of order 10% in most scenarios. We therefore conclude that our new solution performs well in that it is close to optimal in spite of the fact that the a priori probability of a vehicle being malicious was not utilized.

REFERENCES


