Kalman Filter Based Spatial Prediction of Wireless Connectivity for Autonomous Robots and Connected Vehicles

Ramviyas Parasuraman  Petter Ögren  Byung-Cheol Min

Abstract—This paper proposes a new Kalman filter based online framework to estimate the spatial wireless connectivity in terms of received signal strength (RSS), which is composed of path loss and the shadow fading variance of a wireless channel in autonomous vehicles. The path loss is estimated using a localized least squares method and the shadowing effect is predicted with an empirical (exponential) variogram. A discrete Kalman Filter is used to fuse these two models into a state-space formulation. The approach is unique in a sense that it is online and does not require the exact source location to be known a priori. We evaluated the method using real-world measurements dataset from both indoors and outdoor environments. The results show significant performance improvements compared to state-of-the-art methods using Gaussian processes or Kriging interpolation algorithms. We are able to achieve a mean prediction accuracy of up to 96% for predicting RSS as far as 20 meters ahead in the robot’s trajectory.

I. INTRODUCTION

Autonomous mobile robots rely on wireless communications to coordinate between vehicles and to successfully accomplish their tasks. Communications are particularly important when the task itself is about gathering information and sharing this information with others, such as connected self-driving vehicle coordination and urban search and rescue (USAR) missions. There, the environments are often unknown a priori and dynamically changing.

Furthermore, wireless connectivity is complex to predict not only because the signal decreases over distance (path loss) and when large objects are obstructing the line of sight between the transmitter and the receiver (shadowing), but also because these effects are combined with more complex phenomena such as spatial and temporal dynamics in the environments (multipath fading). The combined effect has been shown to sometimes create isolated no-connectivity areas scattered throughout an environment.

Temporal link quality predictions are mainly beneficial for routing or protocol level algorithms such as channel trade-off and handover, whereas spatial prediction of wireless connection quality is useful in mobile ad-hoc networks and connected vehicles, where wireless nodes keep changing their positions. Thus, for autonomous robots, both spatial and temporal prediction is important, and while there has been a lot of research in temporal RSS prediction, relatively few investigations deal with spatial prediction due to the nature of multi-dimensional complexity.

II. RELATED WORK

For assessing wireless connection quality, we use the received signal strength (RSS) metric, which has been used successfully in previous work [2], [3], [4], [5]. In our work, we follow a data-driven approach to predict the RSS in terms of mean and variance values corresponding to path loss and shadow fading components, respectively. We use a localized estimation of large-scale path loss fading element using the recent measurements in the spirit of [2], where the region-to-region wireless channel variations are modeled.

More precisely, we employ a localized linear regression for the path loss component by applying differential path loss between RSS measurements at several previously-visited positions of the robot. This enables a fast prediction and gives modest requirements on the data and computation needs. The shadow fading, on the other hand, is modeled using an empirical spatial variogram method. For this purpose, we employ the exponential model used in [6]. We then use a Discrete Kalman Filter (DKF) to fuse the models above, filter the data, and provide spatial extrapolations of the RSS.

The main contributions of the paper are that we introduce a new online method based on a DKF to quickly predict the RSS in yet unexplored regions ahead of the robot and show that the proposed framework performs better than the state-of-the-art methods on the real-world experiment datasets [7] collected in both indoor and outdoor environments.
This work inspired our design of an adaptive localized log-linear regression scheme to model the path loss component. Additionally, we take inspirations from the probabilistic approaches in [13], [15], [3] and propose an empirical spatial covariance estimation method to model the small-scale dynamics in the RSS (mainly shadow fading) as a first order Markov-Gaussian process [3]. We depart from [13], [15] by using an empirical variogram. This allows us to create a non-parametric framework that will support dynamic environmental conditions. We rely on practical assumptions such as that the robot is able to localize itself (even a simple dead reckoning system should suffice) and that the RSS measurements are readily available.

Kalman filters (KF) have been successfully used in filtering and smoothing of RSS measurements in both temporal [4] and spatial predictions [3]. In this paper, we will not only use a KF based framework for filtering the RSS measurements but also use it as an integration framework for the RSS prediction models formulated as state-space representations.

Moreover, in contrast to the above mentioned related works, which either estimate the source location or assume the source location to be known, we do not assume that the exact location of the radio signal source is known (i.e. no knowledge on router position), adding to the merits of the proposed approach. Besides, our framework provides a fast knowledge on router position), adding to the merits of the exact location of the radio signal source is known (i.e. no environmental conditions. We rely on practical assumptions for creating a non-parametric framework that will support dynamic spatial covariance estimation method to model the small-scale dynamics in the RSS (mainly shadow fading) as a first order Markov-Gaussian process [3].

The main focus of this work is to provide a solution for online spatial RSS prediction that is reasonably accurate yet computationally efficient. Assuming that the robot can localize itself, we have the following measurement tuple: \{p, ν, z\}. Here, \( p \in \mathbb{R}^2 \) is the robot’s position in a plane, \( ν \in \mathbb{R}^2 \) is the instantaneous velocity, and \( z \in \mathbb{R} \) is the RSS measured at \( p \). Given this measurement tuple, we will estimate a state \( x_k \), which is decomposed of path loss \( z_{pl}^k \) and shadowing \( z_{sh}^k \) components,

\[
x_k = \begin{bmatrix} z_{pl}^k \\ z_{sh}^k \end{bmatrix}. \tag{2}
\]

The structure of the proposed method is illustrated in Fig. 1, where each block is explained in the consecutive subsections.

### III. Background on Radio Signal Propagation

An RF signal propagating through a medium is subject to several environmental factors that impact its characteristics. In the absence of nearby obstacles, the signal strength will be reduced by the free-space path loss (FSPL) caused by the spreading out of the signal energy in space. A commonly used model to represent variation in radio signal strength as a function of distance and time is the Log-Normal Shadowing Model (LNSM) (Eq. (1)).

\[
RSS_{(d,t)} = RSS_{d_0} - 10\eta \log_{10}(\frac{d}{d_0}) - \Psi(d) - \Omega(d, t). \tag{1}
\]

where \( d_0 \) is a reference distance (usually 1m), \( RSS_{d_0} \) is the RSS at the reference distance, \( η \) is an environment-dependent path loss parameter, and \( d = |p_r - p_s| \) is the euclidean distance between the radio source position \( p_s \) and the receiver (robot) position \( p_r \). While the path loss fading is deterministic, the shadowing (\( Ψ \)) and multipath fading (\( Ω \)) are often modeled as stochastic processes with (zero-mean) Gaussian or Nakagami distributions.

The LNSM model requires constant updates of the parameters as and when the environment changes and could include the configuration of walls, floors, and obstacles on the path of radio waves to precisely model the shadowing effects [12]. In this work, we mitigate the multipath fading component (\( Ω \)) in the RSS by locally (spatially) averaging measurements. Moreover, we apply localized learning of the radio channel instead of a global source specific model [3].

### IV. Proposed RSS Prediction Method

The main focus of this work is to provide a solution for online spatial RSS prediction that is reasonably accurate yet computationally efficient. Assuming that the robot can localize itself, we have the following measurement tuple: \{p, ν, z\}. Here, \( p \in \mathbb{R}^2 \) is the robot’s position in a plane, \( ν \in \mathbb{R}^2 \) is the instantaneous velocity, and \( z \in \mathbb{R} \) is the RSS measured at \( p \). Given this measurement tuple, we will estimate a state \( x_k \), which is decomposed of path loss \( z_{pl}^k \) and shadowing \( z_{sh}^k \) components,

\[
x_k = \begin{bmatrix} z_{pl}^k \\ z_{sh}^k \end{bmatrix}. \tag{2}
\]

The structure of the proposed method is illustrated in Fig. 1, where each block is explained in the consecutive subsections.

#### A. Multipath fading mitigation

Multipath fading possess spatial correlation only over short distances (usually within a wavelength \( λ \) of the radio signal used) [11]. Therefore, we seek to mitigate the multipath fading effects on the measured RSS by applying a spatial moving average filter (MAF) with a window size \( L_{MAF} = \frac{10λν}{f_s} \), which is adapted online based on the robot’s velocity (\( ν \)) and the RSS sampling frequency (\( f_s \)) such that the sampled distance during that window is around 10\( λ \) (see [8]). The minimum value of \( L_{MAF} \) is empirically set to 5 samples so that the filter also removes much of the noise in the measurement. This minimum value is used also when the robot is not moving (\( ν = 0 \)).

#### B. Path loss component

The attenuation in the RSS due to path loss (PL) can be modeled using a (adaptive) floating-intercept model

\[
z_{pl} = α \log_{10}(d) + Γ. \tag{3}
\]

Here, \( d = |p - p_0| \) is the distance between the robot’s current position \( p \) and a reference position \( p_0 \). Comparing Eq. (1) and (3), we see that \( α = -10η \) and \( Γ = RSS_{d_0} \).
A drawback with this model fitting is that it cannot be directly treated as a state-space representation, and hence does not integrate well into a state prediction framework such as a Kalman filter. Therefore, we propose a new differential path loss model in Eq. (3) which holds all the properties of Eq. (2), yet can be characterized with state transitions as shown in Eq. (4), at an arbitrary instant k.

\[ z - z_0 = \alpha \log_{10}(\frac{d}{d_0}) + \beta, \quad (4) \]

\[ z_{pl}^k = z_{pl}^{k-1} + \alpha(\log_{10}(d_k) - \log_{10}(d_{k-1})) + \beta. \]

Here, \( d_0 \) is the distance between the initial position of the robot and the source. \( z_0 \) is the RSS at initial position \( p_0 \). \( \alpha \) remains the same as in Eq. (2), and \( \beta \) represents the noise in the PL model. We do not assume to know the exact source location. Instead, we assume that when the robot starts its journey, it is within a few meters from the source location (i.e., \( d_0 \approx 1m \)), which is more practical because in connected cars, the GPS positions of vehicles are typically broadcasted first and localization errors in real-world sensors can be accounted into the model. Because we assume erroneous position measurements, a **total least squares** (TLS) solution would better fit the model to the sampled data compared with an **ordinary least squares** (OLS) solution.

We use a reference position \( p_r \) that decides the initiation of the algorithm (s.t. \( d_r \gg d_0 \) and \( z_k > z_r \) ) and the number of training samples (\( N \)) for the algorithm as: \( N = \text{size}(p_k, p_{k+1}, \ldots, p_k) \) with \( p_k \) being the robot position at the \( k^{th} \) instant. \( d_k \) is the distance between the robot at its current instant \( p_k \) and its initial position \( p_0 \). We use a training size adaptation similar to the one in [8] to reduce the prediction error, especially around abrupt changes in wireless channel parameters such as a change of environment from the line of sight (LOS) to non-line of sight (NLOS). Note, the proposed PL prediction model is similar to the ones used in [2], [16], except for three notable differences: (1) a differential PL is used instead of an absolute PL to address state-space transitions; (2) the PL model is fitted with a TLS solution instead of an OLS solution to improve fitting accuracy; (3) exact source position is not required by the model. These differences make our model adaptable to dynamic environments and widen the range of applications.

### C. Shadow fading component

We use a partial geostatistic approach for modeling the shadowing element (\( z^{sh} \)) in the RSS. Given that the shadowing effect is a wide sense stationary (WSS) process [6], we adopt a variogram model to learn the shadowing covariance function. The variogram measures the spatial dependence (dissimilarity) of the \( z^{sh} \). An empirical semi-variogram with a spatial lag \( \lambda \) shown in Eq. (5) is used to assimilate the measurements into the model.

\[ \tilde{\gamma}(\lambda) = \frac{1}{2N} \sum_{k=1}^{N} [\frac{z^{sh}_k - z^{sh}_{k+\lambda}}{z^{sh}_k}]^2. \quad (5) \]

In theory, many valid analytical variogram models (such as exponential, spherical, Gaussian, etc.) exist but especially, the **exponential variogram** model in Eq. (6) has been found to be suitable with the geographical radio signal strength map [3], [6] and is computationally less expensive in fitting. Therefore, we have

\[ \gamma(\lambda) = c_v + c_m[1 - e^{(-\frac{\lambda}{\tau})}], \quad (6) \]

where, \( c_v \) represents nugget (offset), \( c_m \) represents still (variance), and \( \tau \) is the range (characteristic length-scale) of the variogram. The distances are converted to lags (\( \lambda \)) when creating the variogram. We deviate from the traditional approach of using the variogram model in the Kriging process [3], [11], and instead, use a covariance function (similarity measure) as it is easier to integrate into a KF framework. Let us express \( k(p, p') = \text{cov}(z^{sh}(p), z^{sh}(p')) \) as the covariance of shadowing components between the measurements at positions \( p \) and \( p' \). The relation between the variogram model and the covariance function is as follows:

\[ \gamma(\lambda) \equiv K(0) - K(\lambda); \quad (7) \]

By substituting Eq. (6) in Eq. (7), we obtain an **exponential covariance function** (Laplacian kernel) in Eq. (8) that is less sensitive to the correlation parameter (\( \tau \)), which tends to vary greatly in dynamic environments.

\[ K(\lambda) = k(p, p') = c_m \exp\left(-\frac{||p-p'||}{\tau}\right). \quad (8) \]

Here, \( c_m \) is the shadowing variance. Following [17], a state-space representation for the shadowing process is obtained in Eq. (9) with the shadowing prediction variance in Eq. (10).

\[ z_{k+1}^{sh} = z_k^{sh} \exp\left(-\frac{||p-p'||}{\tau}\right) + \sigma_{sh}, \quad (9) \]

\[ \sigma_{sh} = c_m (1 - \exp\left(-\frac{2||p-p'||}{\tau}\right)). \quad (10) \]

### D. DKF Framework

The objective is a data-driven non-parametric approach that is adaptive in real time. DKFs are known to be fast, efficient and useful for reliable online estimation and prediction [4], [3]. We combine the \( z_{pl}^k \) and \( z_{sh}^k \) models described above in a state-space model that can be subject to a DKF design with the following system dynamics:

\[ x_k = A_k x_{k-1} + B_k u_k + G \omega_{k-1}; \quad (11) \]

\[ z_k = H x_k + \epsilon; \quad (12) \]

where, \( x_k \) is the state, \( u_k \) is the control input, \( z_k \) is the measurement, \( \omega \) is the process noise \( \sim \mathcal{N}(0, Q) \) with process noise covariance \( Q \), \( \epsilon \) is the measurement noise vector \( \sim \mathcal{N}(0, R) \) with measurement noise covariance \( R \). The matrices \( A, B, G, H \) and \( \delta \) represent state, control, process noise, and observation transformation, respectively.

\[ x_k = [z_{pl}^k, z_{sh}^k], \quad u_k = \begin{bmatrix} \log_{10}(d_k) \\ \log_{10}(d_{k-1}) \end{bmatrix}, \quad z_k = [z_k]; \quad (13) \]

\[ A = \begin{bmatrix} 1 & 0 \\ 0 & e^{(-\Delta t)} \end{bmatrix}, \quad B = \begin{bmatrix} \alpha & -\alpha \\ 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 1 \end{bmatrix}. \quad (14) \]
The spatial sampling interval is the distance between two successive samples: \( \Delta s = d_k - d_{k-1} \). The PL parameters \( \alpha \) and \( \beta \) are obtained by applying TLS (Deming) regression.

The shadowing components of the training samples are observed as the residuals of the PL components (based on the fitted PL model omitting \( \beta \)) from the true values. The residuals are then processed to create an empirical variogram using Eq. (6). The analytical variogram parameters \( c_m, c_v, \) and \( \tau \) are learned from the data by equating Eq. (5) and (9) with a weighted least squares (WLS) fitting. The nugget (offset in variogram) \( c_v \) is treated as the measurement noise variance. The process and measurement noise covariances \( R \) and \( Q \) are dynamically updated as follows:

\[
R_k = c_v; \quad Q_k = \begin{bmatrix} \text{cov}(\beta) & 0 \\ 0 & c_m(1 - \exp(-2|\Delta s|/\tau)) \end{bmatrix}, \quad (15)
\]

Note the shadowing process is implicitly detrended from the path loss fading \((z^h_k = z_k - z^2_k)\), as advised in [11], so that the applied variogram model complies with the geostatistical principles. This implies that the proposed framework follows a hybrid geostatistic-KF approach. The recursive nature of the DKF and data sub-sampling are exploited to reduce the computational load in processing measured data.

E. Prediction

Assume that the RSS is sampled at a constant frequency \((f_s)\). The RSS spatial interval is directly proportional to the robot’s velocity, \( \Delta s = \frac{\nu}{f_s} \). At any instant \( k \), the prediction will be done for the \((k + P)^{th}\) instant, where \( P \) is the number of samples ahead. Translating this to the spatial domain, for a prediction distance of \( M \) meters ahead of robot’s current position \( p_k \), \( P \) is calculated as follows: \( P = M/\Delta s \). The following equations provides prediction mean and variances.

\[
\begin{align*}
\hat{z}^{P}_{k+P} &= \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] x_k + \left[ \begin{array}{cc} \alpha_k & -\alpha_k \\ 0 & 0 \end{array} \right] \left[ \begin{array}{c} \log_{10}(d_k + M) \\ \log_{10}(d_k) \end{array} \right] \\
\hat{z}^{2P}_{k+P} &= \text{cov}(\beta) + c_m(1 - \exp(-2P\Delta s/\tau)) \quad (16)
\end{align*}
\]

V. EXPERIMENTAL EVALUATION

We implemented the proposed RSS prediction method in MATLAB and tested it on two publicly available real-world measurements datasets [7, 18], whose hardware configurations and parameters are listed in Table I. In both datasets, there were instances of LOS from/to NLOS transitions (e.g., outdoor: \( \approx 90^{th} \) sample, indoor: \( \approx 800^{th}, 1200^{th}, \) and \( 1500^{th} \) samples). Note the RSS is sampled at a frequency of 10 Hz.

To compare our approach with the state-of-the-art methods in spatial RSS prediction, we implemented (i) Kriging interpolation (KI) following [3] (except that the temporal filtering is done using the MAF instead of a KF), (ii) Gaussian Process Regression (GPR) following [19] (assuming an unknown source location), and (3) Linear regression (LR) following [12] (without the influence of obstacles/walls).

As measures of the prediction performance, we used an evaluation metric, Mean Absolute Error (MAE) \((AE^p = |\hat{z}^{P}_{k+P} - z_{k+P}|)\) which considers only the prediction mean so as to provide a fair comparison with all the state-of-the-art methods including LR. Here, \( z_{k+P} \) is the true measured value and \( \hat{z}^{P}_{k+P} \) is the predicted mean value. Using the MAE, we estimate the Mean Prediction Accuracy (MPA = \( 100 \frac{1 - \sum \left| \hat{z}^{P}_{k+P} - z_{k+P} \right|}{\sum |z_{k+P}|} \)) which normalizes the prediction error with respect to the true RSS values.

VI. RESULTS

Fig. 2 and Table I offers summarized results of all the investigated methods in both the environments. The values MAE and MPA are averaged over prediction distances of 1 to 10m in indoor and 1 to 20m in the outdoor datasets. The proposed KF based method reached more than 96% accuracy in the outdoor experiments, whereas it reached around 84% accuracy in the indoor experiment, which had a more complex robot trajectory. Note that the KF accuracy was always greater than 95% for short distances (up to 5m). In all cases, our method performed significantly better than the other methods. Also, it can be seen from Fig. 2 that the proposed method handled the predictions well even after LOS/NLOS transitions, however with some deviations at those transitions (e.g., \( \approx 800 \) samples in indoor environment).

It appears that the predictions in the indoor environment were more challenging partly because of the experiment setup in which the robot was moving more irregularly compared to the linear (unidirectional) robot movements in the outdoor environment. Moreover, several settings including odometry errors (localization system) could have impacted the prediction results. Also, the impact of multipath component, source location errors, and estimation over even longer distances (> 20m) will be investigated in our future work.

VII. CONCLUSIONS

In this paper, we presented an online method for predicting RSS over long distances (over tens of meters) in unknown environments. We exploited the integration of a localized differential path loss model and an exponential variogram approach into a DKF to predict mean and variance of RSS in unvisited positions of the mobile robots and connected vehicles. All parameters in the model are automatically adapted to the measured data, to realize a fully data-driven non-parametric approach with limited model priors on the wireless channel, which is a physical process.

The proposed method was evaluated using real-world experimental data in indoor and outdoor scenarios. The method was able to attain a mean absolute prediction error (MAE) of less than 4 dBm for predicting the RSS up to 20m in advance in outdoor data, whereas in the indoor dataset, the MAE was less than 7 dBm. In all cases, the results of our approach outperformed the state-of-the-art methods such as the Gaussian process regression, linear regression methods, and the Kriging interpolation algorithms.
Outdoor experiment: Error distribution for 20m prediction distance

<table>
<thead>
<tr>
<th>Cumulative Distribution Function (CDF)</th>
<th>MAE (dBm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Indoor experiment: Error distribution for 8m prediction distance

<table>
<thead>
<tr>
<th>Cumulative Distribution Function (CDF)</th>
<th>MAE (dBm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
<td>0.7</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9</td>
</tr>
</tbody>
</table>

are presented for the KF based predictions along with comparison results of the prediction error and its distribution.

Fig. 2: Prediction performance results of the outdoor (top row) and indoor (bottom row) experiment datasets. Sample plots are presented for the KF based predictions along with comparison results of the prediction error and its distribution.

### Table I: Configurations and important parameters used in the measurement dataset [7].

<table>
<thead>
<tr>
<th>Robot</th>
<th>Radio source (Tx)</th>
<th>Radio receiver (Rx)</th>
<th>Tx antenna</th>
<th>Rx antenna</th>
<th>Tx position (w.r.t. robot frame)</th>
<th>Area covered</th>
<th>Average velocity</th>
<th>Number of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outdoor</td>
<td>Tracked UGV</td>
<td>Bullet M2 (2.4GHz Wi-Fi AP)</td>
<td>Bullet M2 (Station)</td>
<td>Omni (16dBi)</td>
<td>(-5.25)</td>
<td>2068</td>
<td>0.3</td>
<td>5839 (raw)</td>
</tr>
<tr>
<td>Indoor [18]</td>
<td>youbot</td>
<td>Axiom RTU200 (2.4GHz Wi-Fi AP)</td>
<td>TP-LINK WN722N (Station)</td>
<td>Directional Omni2DHi (8dBi)</td>
<td>(3.0)</td>
<td>108</td>
<td>0.2</td>
<td>18689</td>
</tr>
</tbody>
</table>

### Table II: Comparison of mean prediction accuracy.

<table>
<thead>
<tr>
<th></th>
<th>KF</th>
<th>Kriging</th>
<th>GPR</th>
<th>LR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outdoor (N=200)</td>
<td>2.59</td>
<td>5.52</td>
<td>5.08</td>
<td>5.88</td>
</tr>
<tr>
<td>Indoor (N=200)</td>
<td>7.01</td>
<td>10.31</td>
<td>12.13</td>
<td>10.91</td>
</tr>
<tr>
<td></td>
<td>96.12</td>
<td>87.75</td>
<td>92.42</td>
<td>91.48</td>
</tr>
<tr>
<td></td>
<td>93.70</td>
<td>75.39</td>
<td>72.97</td>
<td>76.48</td>
</tr>
</tbody>
</table>

### References


