Angular Information-based NLOS/LOS Identification for Vehicle to Vehicle MIMO System

Chen Huang†‡, Andreas F. Molisch†§, Rui Wang†, Pan Tang‡, Ruisi He‡§, Zhangdui Zhong‡
† Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA
‡ School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China
§ State Key Lab of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, 100876, China

Abstract—Distinguishing between line-of-sight (LOS) and non-LOS (NLOS) based on measured channel impulse responses is a fundamental problem for localization systems, as well as channel modeling. In this work we show that the use of angular information can considerably enhance the accuracy of such a scenario discrimination, and we propose the use of support vector machines (SVM) for performing this task. We demonstrate, using real-world channel measurements, that the new method provides significantly improved detection, reducing in our measured channels the mis-identification rate from 10% to less than 4%.

Index Terms—NLOS identification, channel measurement, Rician-K-factor, channel modeling, support vector machines.

I. INTRODUCTION

Location-awareness has become an essential need in a variety of commercial and military applications [1], [2]. Many localization systems, such as the Global Positioning System (GPS), cellular 911 localization, and indoor localization systems, are based on measuring the runtime to the runtime between anchor nodes with known location, and the agent nodes whose location has to be determined. However, the line-of-sight (LOS) connection between the nodes may be blocked by obstacles in the environment, which leads to a NLOS condition. NLOS propagation channels introduce a (positive) bias in distance estimation algorithm [3], [4]; it is thus critically important to distinguish such channels from LOS channels, based on quantities that can be easily measured, i.e., the impulse responses; such a distinction allows to either discard NLOS measurements, or perform more advanced processing to reduce their harmful impact. For this reason various NLOS identification methods have proposed in the literature, e.g., [5], [6].

Another important application for LOS/NLOS discrimination is channel modeling: since most channel models propose to use different channel parameters in LOS and NLOS, respectively, e.g., COST 2100 [7], distinguishing between those situations is important.

An example for the above issues are V2V communication system, where due to the mobility of cars and the diversity of environments, the wireless signals may be blocked or reflected by other cars or buildings in the street; therefore there is often only a NLOS connection between two cars that need to communicate with each other and/or need to find each others position.

In the past, the most widely used solutions to distinguish the LOS and NLOS scenarios included:

• Design the measurement campaign to inherently distinguish the scenarios. For channel modeling, the most direct way is to individually measure the channel in the LOS and NLOS scenarios, respectively. For example [8] measured the V2V channel in an intersection scenario, where the Tx- and Rx-cars first move in convoy and then separate to different directions at the intersection. Thus the measurement campaign has separate LOS and NLOS environments.

• Visual inspection by cross-checking video data and measurement data. For some environment, the LOS and NLOS conditions occur randomly due to blockage by cars or pedestrians. For example [9] used recording from cameras mounted near the Tx and RX to distinguish the LOS and NLOS situations.

• Using some characteristic parameters of the propagation channel to determine whether it is a LOS or NLOS scenario, e.g., [10] used the Rician-K-factor to estimate whether it is a LOS or NLOS scenario.

• Using machine learning tools to distinguish the LOS/NLOS scenarios. Due to the good classification performance of many machine learning tools, many studies conduct machine learning methods to distinguish the LOS and NLOS scenarios, e.g., support vector machine (SVM) [11], relevance vector machine (RVM) [12], and neural network [13].

It is noteworthy that the first two solutions above are not applicable for localization applications, because they require cumbersome human intervention. While using certain channel characteristics can be automatically deployed, a number of investigations have shown the most commonly used criteria (such as Rice factor) result in a rather high mis-identification rate. The now common use of multiple-input-multiple-output (MIMO) systems provides an opportunity to observe the propagation channel in the angular domain, which allows to investigate new discrimination criteria. At the same time, the good accuracy of machine learning tools for classification
makes the machine learning-based LOS/NLOS algorithms a promising approach to further enhance the accuracy.

Inspired by these considerations, we propose in this paper an angular information-based LOS/NLOS identification algorithm that combines conventional channel features, e.g., Rician-K-factor, kurtosis, and RMS-delay spread, with angular information as the training data for a SVM to generate an accurate classifier to distinguish the LOS and NLOS conditions, and demonstrate its efficacy in measured V2V channels.

The main contributions of this paper are summarized as follows:

- To the best of our knowledge, the proposed scheme is the first to use the angular information to identify the LOS/NLOS scenario.
- We propose a novel identification solution by using an SVM that combines angular and traditional channel parameters. In other words, instead of simply using the conventional Rician-K-factor, Kurtosis or Skewness, we also extract the angular features as training data and use SVM with all of them to generate the classifier model to distinguish the LOS and NLOS scenario.
- We use real V2V measurement data to assess the performance of our new algorithm; to the best of our knowledge, this is the first use of real measurement data in this context. In our experiments, the proposed solution reduces the identifying error rate from 10% for conventional solutions to 5%.

The rest of this paper is organized as follows. We first introduce the measurement campaign and parameter extraction in Section II. The details of the proposed algorithm are elaborated in Section III. Section IV presents the algorithm performance analysis based on the measurement data. Finally, Section V presents conclusions and an outlook to future work.

II. LOS/NLOS IDENTIFICATION

In this section, we present the angular information-based LOS/NLOS identification in two steps: the feature selection and data training of using SVM.

A. Feature Selection

Generally, the main factors affect the accuracy of an SVM-based classifier include: i) the design of the input feature vector and ii) the parameter settings of the SVM method. Through our experiments, the former factor most influences the LOS/NLOS identification, which indicates that feature selection is a key point to generate an accurate NLOS classifier.

From measured (or simulated) propagation channels, it is possible to extract a number of features expected to capture the salient differences between LOS and NLOS conditions for each snapshot. The selected features for the \( j \)-th snapshot include:

**Maximum received power over delay samples** \( \left( \max(||h_j(t)||^2) \right) \): The LOS multipath component (MPC) generally contains more power than the NLOS MPCs; therefore, the maximum received power of each snapshot \( \max(||h_j(t)||^2) \) is a referential information for NLOS.

**Kurtosis of the received power** \( (\kappa_j) \): measures the peakedness of the probability distribution, which is defined by the ratio between the fourth and the second order moments of the received signal’s amplitude. Usually, the amplitude of signals in the NLOS scenario decreases slower than the LOS scenario, thus the kurtosis is generally larger for a LOS condition, which is also expressed in Fig. 2(a). The kurtosis can be calculated as:

\[
\kappa_j = \frac{E[(|h_j(t)| - \mu_{|h_j(t)|})^4]}{E[(|h_j(t)| - \mu_{|h_j(t)|})^2]^2},
\]

where \( \mu_{|h_j(t)|} \) and \( \sigma_{|h_j(t)|} \) are the mean and standard deviation of \( |h_j(t)| \), which can be expressed as:

\[
\mu_{|h_j(t)|} = \frac{\sum_{l=1}^{L} ||h_j(t_l)| - |\bar{h_j}|||}{L},
\]

\[
\sigma_{|h_j(t)|} = \sqrt{\frac{\sum_{l=1}^{L} (||h_j(t_l)| - |\bar{h_j}||)^2}{L}}.
\]

**Skewness of the received power** \( (\beta_j) \): measures the asymmetry of the probability distribution, thus the skewness of a Rayleigh distribution is generally larger than that of a Rician distribution. More generally, the NLOS data usually have a higher skewness than the LOS data. The skewness can be calculated as:

\[
\beta_j = \frac{E[(|h_j(t)| - \mu_{|h_j(t)|})^3]}{\sigma_{|h_j(t)|}^3},
\]

where \( \mu_{|h_j(t)|} \) and \( \sigma_{|h_j(t)|} \) are given in (2) and (3), respectively. **Rising time** \( (\Delta \tau_j) \): measures the time interval between the first MPC and the last MPC. Considering the NLOS scenarios usually contain more reflections and scatterings, which means propagation distance is generally longer, the rising time in NLOS scenarios is usually larger than it in the LOS scenarios. The rising time can be calculated as:

\[
\Delta \tau_j = \arg \max_{\tau} [h_j(t) - \min(h_l)]
\]

where \( l \) is the index of the MPCs.

**RMS-delay spread** \( (\tau_{rms}) \): measures the rms delay spread of all MPCs in the current snapshots. In an NLOS channel single strong component (LOS) is absent, which tends to lead to a higher concentration of the power in delay; thus the RMS-delay spread is generally higher in the NLOS scenarios than in the LOS scenarios. The RMS-delay spread can be calculated as:

\[
\tau_{rms} = \sqrt{\frac{\sum_{l=1}^{L} (\tau_l - \tau_m)^2||h_j(t_l)||^2}{\sum_{l=1}^{L} ||h_j(t_l)||^2}}
\]

where \( \tau_m \) is the means excess delay, which can be expressed as:

\[
\tau_m = \frac{\sum_{l=1}^{L} \tau_l ||h_j(t_l)||^2}{\sum_{l=1}^{L} ||h_j(t_l)||^2}.
\]

**Rician-K factor** \( (K_{1,2}) \): is defined as the ratio between the power of a (possible) dominant MPC (typically the LOS) and the power in the remaining MPCs. Existing theoretical and
empirical studies have shown that there is a link between the Rician-K factor and the presence of LOS conditions. The Rician-K factor can be roughly approximated as:

$$K_r = \frac{(|h(t)|_{\text{max}})^2}{2\sigma_h^2}$$  \hspace{1cm} (8)$$

where $|h(t)|_{\text{max}}$ represents the amplitude of the main peak, which physical meaning is interpreted here as the LOS transmission peak (though strictly speaking it can also be created by a dominant reflected component). $\sigma_h$ is the variance of amplitude, representing the intensity of the multipath transmission signal. Specifically, in NLOS conditions where no direct path exists, the Rician K factor is expected to be close to zero.

Angular difference ($\Delta \lambda_{j,t}$) measures the difference between the AOA ($\theta$) and AOD ($\phi$) of the strongest MPC in each snapshot. In the LOS scenario, the strongest MPC should be the LOS MPC. For a LOS MPC, the signal transmitted from the Tx directly propagates to the Rx, thus the angular difference ($\theta - \phi$) is calculated as by using Euclidean distance in Cartesian coordinate.

Note that, to avoid the impact of the periodicity of the angle, the angular difference is calculated as by using Euclidean distance in Cartesian coordinate.

Angular spread of departure/arrival ($\lambda_{\text{ASD}} / \lambda_{\text{ASA}}$) measures the angle spread of departure/arrival of all MPCs in the current snapshot. Since the propagated signals more concentrate on the LOS MPC in the LOS scenarios, the angle spread of the LOS scenarios is generally smaller than in the NLOS scenarios. The angular spread of departure/arrival can be calculated as [14]:

$$\lambda_{j,\text{ASD/ASA}} = \sqrt{\frac{\sum_{l=1}^{L} \exp (i \theta_l/\phi_l - \mu_{\theta/\phi,j})^2 |h(t_l)|^2}{\sum_{l=1}^{L} |h(t_l)|^2}}$$  \hspace{1cm} (9)$$

where $\mu_{\theta/\phi}$ is the mean of the angular power spectrum, which can be calculated as:

$$\mu_{j,\theta/\phi} = \frac{\sum_{l=1}^{L} \exp (i \theta_l/\phi_l) |h(t_l)|^2}{\sum_{l=1}^{L} |h(t_l)|^2}.$$  \hspace{1cm} (10)$$

Note that while the first parameters have been used for LOS/NLOS detection for a considerable time, use of the angular parameters is novel.

Consequently, the input (feature) vector for the $j$-th snapshot of the static solution can be expressed as:

$$x_j = \{ \max(|h_j(t)|^2), \mathcal{X}_j, S_j, \Delta \mathcal{T}_j, \tau_{j,\text{rms}}, K_r,j, \Delta \lambda_{j,t}, \lambda_{j,\text{ASD}}, \lambda_{j,\text{ASA}} \},$$  \hspace{1cm} (12)$$

B. NLOS/LOS Identification based on Support Vector Machines

The SVM is a supervised learning method which has a great advantage on classification, especially for binary classification problem [15]. The SVM has been widely used for classification as its robustness and few requirements of pre-defined parameters. Specifically, least squares support vector machines (LS-SVM) [16] is used in this study to avoid the quadratic programming problem, which simplifies the optimization to learn the weights in the SVM.

A linear classifier can be expressed as a function of $X \Rightarrow \{-1,+1\}$ with the following form.

$$\mathcal{L}(x) = \text{sign}[w^T \varphi(x_j) + w_0]$$  \hspace{1cm} (13)$$

where sign is the signum function, $\varphi(\cdot)$ is a predetermined function, $w$ and $w_0$ are weight parameters learned from the training data $\{x_j, \mathcal{L}_j\}_{j=1}^N$, where $x_j \in \mathbf{X}$ and the input labels $\mathcal{L}_j \in \{-1,+1\}$. In this case, the LS-SVM separates two classes $\{-1,+1\}$ by determining the separating hyperplane that maximizes the margin between the two classes. Therefore, the LS-SVM classifier is obtained by solving the following optimization problem.

$$\arg \min_{w,w_0,c} \frac{1}{2}||w||^2 + \frac{1}{2} \sum_{j=1}^{N} c_j$$  \hspace{1cm} (14)$$

$$\text{s.t.} \quad \mathcal{L}_j[w^T \varphi(x_j) + w_0] = 1 - e_j, \forall j$$  \hspace{1cm} (15)$$

where $c$ is the weighting factor that controls the trade-off between training error and model complexity. Moreover, considering that the extracted characteristics of the LOS/NLOS scenario are not linear separable, as shown in Fig. 2, we use a Gaussian radial basis function (RBF) [17] rather than a linear mapping function to better classify the LOS/NLOS scenarios:

$$k(x, x_j) = \varphi(x)^T \cdot \varphi(x_j) = \exp \left[ -\frac{||x - x_j||^2}{2\sigma^2} \right]$$  \hspace{1cm} (16)$$

where $\sigma^2$ is the hyper-parameter learned from the training data using (14). It has been proven that the optimization problem (14) is a linear programming problem [16], which can be solved with its Lagrangian dual and Karush-Kuhn-Tucker conditions to obtain the prediction of the LS-SVM as

$$\mathcal{L}(x) = \text{sign} \left[ \sum_{j=1}^{N} \alpha_j \mathcal{L}_j k(x, x_j) + w_0 \right]$$  \hspace{1cm} (17)$$

where $\alpha_j$ is the Lagrange multiplier.

To distinguish the LOS and NLOS scenarios, we train an LS-SVM classifier with the input $x_j$ and corresponding labels $\mathcal{L}_j = -1$ and $+1$ for the NLOS and LOS data, respectively.

It is noteworthy that not only are we interested in the performance of the LS-SVM for certain features, but we are also interested in which subsets of the available features give the best performance. The details of the experiments are present in Section IV.
### III. Measurement Campaign and Parameter Extraction

This section provides details of the measurement campaign and parameter extraction procedure that provides the channel characteristics on which the evaluation of the proposed algorithm is based.

#### A. Measurement Campaign

The V2V measurement campaign was conducted with a self-built real-time MIMO channel sounder described in [18]. The sounder includes a pair of NI-USRP RIOs as the main RF transceivers, two GPS-disciplined rubidium references as the synchronization units and a pair of 8-element uniform circular arrays (UCAs) that are connected to the USRPs via electronic switches. The key parameters of the setup are given in Table I, and Figs. 1(a) and (b) show pictures of the measurement cars and antenna array.

The measurements analyzed here were conducted on the campus of the University of Southern California, LA, and public roads near the campus, as shown in Fig. 1(d). We used $360^\circ$ cameras located next to the Rx and Rx to record videos of the environment during the whole measurement. Thus, determination of the "ground truth", i.e., whether LOS or NLOS was valid at a particular time, can be accurately determined. This will be used both for the training, and the performance assessment. Fig. 1(c) gives an example snapshot of the video at the RX side. More details of the sounder and the measurement campaign can be found in [18], [19].

#### B. Parameter Extractions and Problem description

To utilize the angular information to distinguish the LOS and NLOS scenario, the relevant channel characteristics need to be extracted first. Therefore, RiMax, a high-resolution parameter estimation algorithm [19] is implemented to estimate the parameters of MPCs in each snapshot, including the power, delay, azimuth of arrival (AOA), azimuth of departure (AOD), which are denoted as $\{|h_i(t)|^2, \tau_i, \Theta_i, \Phi_i\}$, where $i$ is the index of the MPCs. RiMax is essentially an iterative maximum-likelihood estimator that provides both the MPC parameters, and an estimate of the diffuse background. For the current paper, we will only exploit knowledge of the discrete MPCs.

Besides these parameters, the most widely used characteristics to identify a NLOS condition include kurtosis and Rician-K-factors. However, for time-varying V2V channels, it is hard to find a certain threshold of the characteristics to distinguish the LOS and NLOS scenario. As shown in Fig. 2, both the kurtosis or the angle spread of arrival show different distributions in the LOS and NLOS scenarios, but there exists no threshold to separate the LOS and NLOS data perfectly. In contrast, the SVM based identification can well employ the different features within different dimensions and achieve a high accuracy of identification as the result. Therefore, our proposed algorithm employs different characteristics to build a feature vector and uses SVM to generate the classifier. The details of the proposed algorithm are given in Section III.

### IV. Algorithm Validation from Experimental Results

In this Section, we present the details of the experiments based on the V2V measurement data introduced in Section II, including the contribution of different features to the classification and the final results.

#### A. Feature Contribution

To better analyze the contributions of different characteristics, which are extracted from the V2V measurement data, to the LOS/NLOS identification, we analysis the identification performance for each characteristic as shown in Fig. 4, where (a)-(i) give the probability distribution functions (PDF) of different extracted characteristics. From the result, all of the extracted characteristics show different features in the LOS and NLOS scenarios, respectively. Specifically, the angular-based characteristics (angular distance, ASD, and ASA) show a good separation of the LOS and NLOS data, where the LOS
data are generally centralized to a certain value and the NLOS data have a relatively scattered trend.

Nevertheless, most of these characteristics also have some overlap areas between the LOS and NLOS scenarios. Thus, it is hard to directly determine whether it is a LOS/NLOS scenario based on single characteristic.

**B. Results and Discussion**

To evaluate the performance of our solution, we validate the algorithm by using the V2V measurement data and make a comparison of: i) the conventional single characteristic-based identification solution, ii) the conventional SVM-based identification solution, which uses the characteristics except for the angular-based features as the training data for the SVM, and iii) the proposed solution, which includes the angular characteristics.

In this part of the validation, we divide the measurement results into two sets for training and validation, respectively. We consider two types of dividing the data. First, we randomly select data sets for training, and validate the algorithm by using the remaining data, which is the usual approach for SVM validating and named Partition A. As a comparison, we select the data collected at two certain roads in and out the campus, respectively, for training, whereas the data collected at the other roads are used for validation, which is more practical and named Partition B. The validation layout parameters are given in Table II.

Fig. 3 gives the error rate of the LOS/NLOS identification for Partition A and B by using the proposed solution, the conventional SVM-based solution, and the single characteristic-based identification results. From the result, it is found that both the proposed and the conventional SVM-based solutions outperform all the single characteristic-based solutions, and that the proposed angular-information based LOS/NLOS identification achieves the lowest error rate. Meanwhile, all solutions for Partition B have higher error rate compared with Partition A, since in Partition B, we used the data collected at different streets for training and validation, respectively. In other words, the training and validation data for Partition B have less correlation, which make the identification harder but more practical. In addition, among the single characteristics, the maximum power has the best identification accuracy, where the angular distance also shows a good performance on identification.

**V. Conclusion**

In this paper, we propose an angular information-based LOS/NLOS identification algorithm, which not only uses the conventional NLOS identification features but also employs the channel characteristics in the angle domain, including ASA, ASD, and angular distance. The LOS/NLOS classifier is generated by using an SVM based on the V2V measurement data. From our validation, the proposed solution reduces the identifying error rate from 10% for conventional solutions to 5% in classical setup. In the future work, the contributions of the time-varying channel characteristics to LOS/NLOS identification will be further studied.
Fig. 4. Probability distribution function (PDF) of different extracted characteristics, including: (a) Maximum power, (b) Kurtosis, (c) Skewness, (d) Rising time, (e) RMS delay spread, (f) Rician-K-factor, (g) Angular distance, (h) Angular spread of departure, and (i) Angular spread of arrival.

REFERENCES


