Research on Kernel Functions of SVM for Line-of-sight Identification in Vehicle-to-Vehicle MIMO System

Chen Huang[†*], Andreas F. Molisch[†], Rui Wang[†], Pan Tang[†], Ruisi He[♮], Zhangdui Zhong[♮,§]
[†] Ming Hsieh Department of Electrical and Computer 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 Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
[§]Beijing Engineering Research Center of High-Speed Railway Broadband Mobile Communications,
Beijing Jiaotong University, Beijing 100044, China
E-mail: {ch_628, molisch}@usc.edu, {ruisi.he, zhdzhong}@bjtu.edu.cn

***Abstract*—Generally, propagation channels show significantly different characteristics for line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. Due to their good performance in classification problems, support vector machines (SVM) have been widely used for NLOS identification of propagation channels. In this paper, we investigate the impact of different kernel functions on the accuracy of SVM-based NLOS identification and validate the performance based on *measured* channel data. We find that a Gaussian kernel reduces the mis-identification rate by a factor 4 compared to a linear kernel, and also outperforms polynomial and sigmoid kernels.**

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

I. Introduction

Channel modeling is one of the fundamental research topics for radio propagation. Characteristics of the propagation channels are usually significantly different for line-of-sight (LOS) and non-line-of-sight (NLOS) conditions, respectively. Therefore, most channel models, such as COST 2100 [1] or 3GPP have separate channel descriptions for LOS and NLOS conditions. Whether a particular channel is in a LOS or NLOS situation becomes a crucial assumption to accurately model the propagation channels.

In the past, there were several solutions to identify LOS/NLOS conditions, including i) Key parameter thresholding: Since the radio propagation in LOS and NLOS conditions shows substantially different characteristics, e.g., Rician-K-factor and kurtosis, it is possible and practical to identify the LOS and NLOS condition by comparing some key parameters extracted from measurement data to a threshold. Nevertheless, it is hard to find a general such threshold for the key parameters especially when it needs to be valid in a large set of measurement data. , especially for time-varying channels where the channel properties changes with the variation of environments. ii) Machine learning-based LOS/NLOS identification: machine learning-based LOS/NLOS identification algorithms have been developed to automatically distinguish the LOS and NLOS condition based on measurement data. In this research, the support vector machines (SVM) method is the most widely used algorithm to identify LOS/NLOS due to its good performance in classification problems [2], [3], and can significantly outperform thresholding of single key parameters [5].

The basic idea of SVM-based classification is to analyze a set of training data with a known ground truth and seek a hyperplane to divide the data. To find a proper hyperplane, a kernel function is used to map the data into a higher dimension, where the training data are easier to separate. If a proper hyperplane cannot be found, the accuracy of the SVM-based classifier will decrease considerably. Therefore, the selection of the kernel function is crucial to the performance of SVM-based LOS/NLOS identification. It is for this reason that we investigate the performance of SVM-based LOS/NLOS identification with different kernel functions. To obtain realistic results, we use *measured* vehicle-to-vehicle (V2V) channel data.

II. SVM-based NLOS Identification

To identify the LOS and NLOS situations, we first need to collect the LOS and NLOS propagation characteristics in a database, and then use the SVM to learn the difference of the distribution of different training data. The design of the input feature vector is one of the key factors impacting the classification performance.

A. Channel Characteristics for Input Vector

A LOS scenario has a variety of different channel characteristics compared to a NLOS scenario. The ones used for the input vector in our study are as follows:

Maximum received power over delay samples ($\max(|h_j(t)|^2)$): 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 LOS/NLOS.

Kurtosis of the received power ($\mathcal{K}_j$) 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 probability density function. Usually, the amplitude of signals in the NLOS scenario is less concentrated than in the LOS scenario, thus the kurtosis is generally larger for a LOS condition.

Skewness of the received power ($\mathbb{S}_j$) 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.

Maximum excess delay ($\Delta\tau_j$) measures the time interval between the first MPC and the last MPC. It is typically larger in NLOS scenarios than in LOS scenarios.

RMS-delay spread ($\tau_{j,rms}$) measures the rms delay spread of all MPCs in the current snapshots. In a LOS channel a single strong component (LOS) is present, 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.

Rician-K-factor ($K_{\mathrm{r},j}$) 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.

Angular difference ($\Delta\lambda_{j,l}$) measures the difference between the AOA and AOD of the strongest MPC in each snapshot. In the LOS scenario, the strongest MPC should be the LOS MPC.

Angular spread of departure/arrival ($\lambda_{\mathrm{ASD}}/\lambda_{\mathrm{ASA}}$) 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 it in the NLOS scenarios.

A more detailed description of the input feature vector, as well as our general SVM setup, can be found in [5].

B. Kernel functions for SVM

As mentioned before, to find a proper hyperplane to divide different training data, the kernel functions are used to project the training data into a higher dimension, where it may be easier to find a proper hyperplane. Nevertheless, different kernel functions have a different impact on the performance of classification. Hence, we investigate the impact of different kernel functions in this paper. The selected kernel functions include:

- Linear kernel function:

$$K(x^{(i)}, x^{(j)}) = x^{(i)\mathrm{T}} x^{(j)} \tag{1}$$

- Polynomial kernel function

$$K(x^{(i)}, x^{(j)}) = (\gamma x^{(i)\mathrm{T}} x^{(j)} + c)^n \tag{2}$$

- Gaussian kernel function

$$K(x^{(i)}, x^{(j)}) = \exp\left(-\frac{(x^{(i)} - x^{(j)})^2}{2\sigma^2}\right) \tag{3}$$

- Sigmoid kernel function

$$K(x^{(i)}, x^{(j)}) = \tanh(\eta < (x^{(i)}, x^{(j)}) > +c) \tag{4}$$

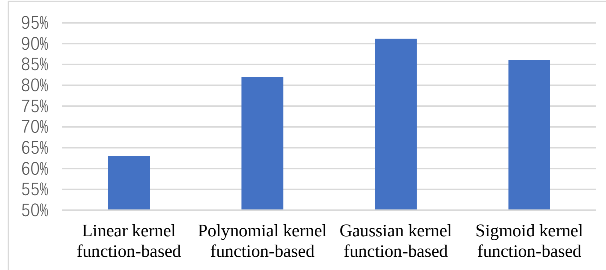

Fig. 1. The performance of using different kernel functions solution

Different kernel functions are able to project the data into a higher dimension in different ways, which leads to different hyperplane and different accuracy of classification.

III. Performance Evaluation

A. Measurement campaign

To validate the performance of SVM-based LOS/NLOS identification algorithm, a V2V channel measurement was conducted at 5.9 GHz with a self-built real-time MIMO channel sounder [4]. 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. More details about the measurement route and sample data can be found in [5].

B. Valuation based on measurement data

To validate the performance, 5000 sets of training data are randomly selected from the collected measurement data, whereas the remaining data (4350) are used as validation data. The performance of using different kernel functions solution is summarized in Fig. 1. From the results, it is found that the Gaussian kernel function-based SVM classification achieves the best performance, whereas the linear kernel function based solution performs the worst. The difference is very significant, with the Gaussian kernel reducing the mis-identification rate by a factor of 4 compared to the linear kernel.

References

[1] L. Liu et al.,“The COST 2100 MIMO channel model,” *IEEE Wireless Communications*, vol. 19, no. 6, pp. 9299, December 2012.

[2] S. Marano, W. M. Gifford, H. Wymeersch, and M. Z. Win, “NLOS identification and mitigation for localization based on UWB experimental data, *IEEE J. Sel. Areas Commun.,* vol. 28, no. 7, pp. 10261035, 2010.

[3] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom and J. Frolik, “Non-Line-of-Sight Identification and Mitigation Using Received Signal Strength,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1689-1702, March 2015.

[4] R. Wang, C. U. Bas, O. Renaudin, S. Sangodoyin, U. T. Virk and A. F. Molisch, “A real-time MIMO channel sounder for vehicle-to-vehicle propagation channel at 5.9 GHz,” In *Proc. 2017 IEEE International Conference on Communications (ICC)*, Paris, 2017, pp. 16.

[5] C. Huang, A. F. Molisch, R. Wang, P. Tang, R. He and Z. Zhong , “Angular information-based NLOS/LOS identification for vehicle to vehicle MIMO system,” submitted to *2019 IEEE International Conference on Communications (ICC)*, 2017.