Machine Learning Aided Path Loss Estimator and Jammer Detector for Heterogeneous Vehicular Networks

Bugra Turan[1], Ali Uyrus[1], O. Nuri Koc[2], Emrah Kar[2] and Sinem Coleri[1]

[1]Department of Electrical and Electronics Engineering, Koc University, Sariyer, Istanbul, Turkey, 34450
E-mail: bturan14@ku.edu.tr, auyrus18@ku.edu.tr, scoleri@ku.edu.tr
[2]Koc University Ford Otosan Automotive Technologies Laboratory, Sariyer, Istanbul, 34450, Turkey.
E-mail: okoc@ku.edu.tr, ekar@ku.edu.tr

GLOBECOM 2021 - 2021 IEEE Global Communications Conference | 978-1-7281-8104-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/GLOBECOM46510.2021.9685428

***Abstract*—Heterogeneous vehicular communications aim to improve the reliability, security and delay performance of vehicle-to-vehicle (V2V) communications, by utilizing multiple communication technologies. Predicting the path loss through conventional fitting based models and radio frequency (RF) jamming detection through rule based models of different communication schemes fail to address comprehensive mobility and jamming scenarios. In this paper, we propose a machine learning based adaptive link quality estimation and jamming detection scheme for the optimum selection and aggregation of IEEE 802.11p and Vehicular Visible Light Communications (V-VLC) technologies targeting reliable V2V communications. We propose to use Random Forest regression and classifier based algorithms, where multiple individual learners with diversity are trained by using measurement data and the final result is obtained by averaging outputs of all learners. We test our framework on real-world road measurement data, demonstrating up to 2.34 dB and 0.56 dB Mean Absolute Error (MAE) improvement for V-VLC and IEEE 802.11p path loss prediction compared to fitting based models, respectively. The proposed jamming presence detection scheme yields 88.3% accuracy to detect noise interference injection for IEEE 802.11p links, yielding 3% better prediction performance than previously proposed deep convolutional neural network (DCNN) based scheme.**

***Index Terms*—Vehicular visible light communication (V-VLC), vehicular communications, heterogeneous communications, IEEE 802.11p.**

I. INTRODUCTION

Vehicular communication is one of the essential building blocks of the next generation intelligent transportation systems (ITS) promising increased traffic safety and efficiency. Currently, vehicular communication systems aim to increase road safety through the exchange of road and traffic information messages by using standardized technologies including IEEE 802.11p and cellular vehicle-to-everything (C-V2X). Furthermore, V2X communications is evolving to support future applications, such as cooperative maneuvering and automated driving with additional standardization efforts such as 5G-NR-V2X, IEEE 802.11bd [1] and technologies including mmWave and V-VLC [2], [3]. However, the link reliability of the novel mmWave and V-VLC technologies is a major concern due to the high atmospheric attenuation of the mmWave signals and mainly line-of-sight (LoS), weather dependent transmission characteristics of the V-VLC. Therefore, the upcoming technologies are foreseen to complement the current vehicle to everything communications (V2X) schemes to increase bandwidth, provide security, and offload the network.

To date, V-VLC is already demonstrated to be a strong complementary candidate for RF based V2V communications. V-VLC relies on the optical modulation of readily available vehicle light emitting diode (LED) lights through intensity modulation and direct detection (IM/DD) scheme. V-VLC provides license free, secure LoS communications due to its directional transmission characteristics, and resiliency to RF interference, malicious jamming and spoofing. In the literature, the hybrid usage of the RF and V-VLC technologies for reliable vehicular connectivity is investigated [4]–[6]. Authors in [4] demonstrate that V-VLC usage helps to decrease the end-to-end delay while achieving high message delivery ratio under intentional RF jamming. [5] provides a secure visible light communication (VLC) scheme to support IEEE 802.11p networks for reliable platooning applications, resilient to security attacks such as platoon data packet injection, replay, and platoon maneuver attacks. [6] provides a performance analysis for video transmissions in a platoon of vehicles through IEEE 802.11p and V-VLC, where V-VLC is demonstrated to increase the quality-of-experience by 30 %. Even though the simultaneous utilization of different technologies increases link redundancy, optimum technology selection and load balancing between two technologies can help to increase throughput while assuring reliability. Therefore, estimation of the link reliability and performance is one of the key enablers for hybrid communication schemes. However, current literature lacks to address the channel reliability quantification for the concurrent usage of IEEE 802.11p and V-VLC technologies by using channel reliability metrics such as path loss and jamming presence.

Recently, machine learning (ML) aided prediction schemes are proposed to estimate V2V link parameters with high accuracy [7]–[9]. Authors in [7] demonstrate the superior performance of Random Forest regression to log distance path loss model for path loss estimation of V2V channels, [8] addresses the viability of throughput prediction and channel parameter regression through supervised learning methods. [9] provides a ML based quality-of-service (QoS) prediction model, which predicts C-V2X network QoS levels two seconds ahead with 85 % reliability. However, the currently proposed ML based schemes consider only one technology, where they lack to investigate the applicability of ML based prediction schemes for heterogeneous vehicular communications. On the other hand, jamming detection for vehicular networks is investigated in [10], [11] based on simulations. Moreover, ML based RF jamming detection schemes are investigated in [12], [13]. [10] studies differentiation between jamming and interference with the utilization of Random Forest and k-means clustering algorithms. [11] proposes a simulation based study by using k-means algorithm based jamming presence detection scheme which incorporates received signal strength (RSS), packet-delivery-ratio (PDR), signal-to-noise ratio (SNR), and relative vehicle speed features. [12] compares different machine learning methods to detect jamming attacks to IEEE 802.11 networks, and [13] proposes a DCNN based scheme that identifies jamming with 85% accuracy for orthogonal frequency division multiplexing (OFDM) based RF communications. However, none of the studies to date investigate the design of a data driven path loss prediction model yielding concurrent path loss estimations for hybrid IEEE 802.11p and V-VLC links together with the jamming presence prediction, which helps to select the most viable link for reliable V2V communications under the consideration of vehicle movements and continuous additive white Gaussian noise (AWGN) type RF jamming existence.

In this paper, we propose ML based estimation and classification schemes to predict the reliability of IEEE 802.11p and V-VLC channels through path loss estimation and radio jamming presence detection. We propose Random Forest based regression for V-VLC and IEEE 802.11p V2V link path loss prediction, where inter-vehicular distance, speed, and bearing are considered as input features. We further provide a Random Forest classifier for IEEE 802.11p radio jamming presence detection with the SNR and PDR input features. The main contributions of this paper are threefold. First, we design a Random Forest based V-VLC and IEEE 802.11p path loss predictor for V2V links by using real-world driving measurement data which incorporates multiple mobility dependent inputs such as distance, bearing and relative speed. Second, we propose a Random Forest classifier based constant jamming presence detection scheme for IEEE 802.11p link. Third, we evaluate the performance of the proposed path loss prediction schemes in comparison to the existing fitting based models and provide jamming presence classification accuracy.

The rest of the paper is organized as follows. Section II describes the path loss and jamming presence prediction problem. Section III presents the system model and the implementation details. Section IV provides performance evaluation of the path loss and jamming presence prediction models. Finally, conclusions are given in Section V.

II. Problem Statement

We consider a hybrid V2V link based on V-VLC and IEEE 802.11p interfaces between the transmit (TX) and receive (RX) vehicles in LoS leader-follower scenario. The path loss between the TX and RX vehicle for considered V2V links is given by $PL_{Tx,Rx} = P_{Tx} - P_{Rx}$, where P_{Tx} and P_{Rx} are the transmit and received power, respectively. For V-VLC links, P_{Rx} mainly depends on the inter-vehicle distance and the geometry between two vehicles due to vehicle LED light illumination pattern and optical receiver field of view (FoV). On the other hand, IEEE 802.11p V2V link path loss is mainly determined by the inter-vehicle distance and environment, with nearby reflector and scatter objects. Therefore, we model the path loss as a regression problem with the inputs of inter-vehicular distance (d), bearing (ϕ), and relative speed (υ), yielding path loss estimates. Considering the location and velocity information exchange between two vehicles, the objective is to estimate the path loss with minimum prediction error, through the usage of offline trained Random Forest regression path loss model. Moreover, the importance of the input features on the prediction performance is also investigated to demonstrate the model parameter input importance difference between V-VLC and IEEE 802.11p V2V links.

Considering the jamming and interference vulnerability of the RF links, the jamming presence detection scheme is modelled as a classification problem. Random Forest classifier is trained to yield jamming presence with the PDR and SNR values obtained at the receiver vehicle.

III. System Model and Implementation

The proposed learning based path loss estimation and IEEE 802.11p radio jamming detection schemes consist of data acquisition, data processing, offline training and evaluation phases. The evaluation phase includes path loss estimation and jamming detection at the TX and RX vehicles, respectively. The detailed design of the proposed schemes is elaborated in the remainder of this section.

A. Path Loss Prediction for VVLC and IEEE 802.11p V2V Links

The path loss prediction schemes yield path loss for V-VLC and IEEE 802.11p V2V links with the mobility dependent model inputs. The path loss prediction scheme is built upon the real-world measurement data, with the additional feature of bearing extracted at the data processing stage with respect to global positioning system (GPS) coordinates providing 2.5 m horizontal position accuracy [14]. Then, two different Random Forest regression models are trained to yield path loss estimates of V2V V-VLC and IEEE 802.11p links.

a) Data Acquisition: To train path loss prediction models, time stamped GPS location and speed information of TX and RX vehicles are recorded with the V-VLC and IEEE 802.11p RSS information at the RX vehicle. The GPS data is recorded with the GPS disciplined oscillator of the Software-Defined Radio (SDR) at the TX and RX vehicles. To measure RSS of V-VLC signals, a sinusoidal tone with fixed transmit power of -6 dBm is transmitted through the LED day time running light (DRL) of TX vehicle, and the optical signals are captured with a silicon photodetector (PD) during day time. The V-VLC RSS is measured at the spectrum analyzer. The RSS of IEEE 802.11p packets are acquired through on-board unit (OBU) device at channel 172 with center frequency of 5.86 Ghz. The measurements are recorded at 1 Hz frequency. Simultaneous data acquisition and recording is handled through LabView software running on a laptop. The path loss measurement setup parameters are as follows: IEEE 802.11p TX vehicle antenna height = 153 cm, IEEE 802.11p RX vehicle antenna height = 178 cm, PD height = 83.4 cm, TX vehicle LED height = 49.7 cm, V-VLC TX electrical power = -6 dBm, RF TX power = 20 dBm. To train and validate the proposed path loss models, 5200 measurement points which include 46800 samples of V-VLC RSS and IEEE 802.11p signal RSS values are collected at different inter-vehicle distance, bearing, relative speed, TX vehicle Latitude, TX vehicle Longitude, RX vehicle Latitude, RX vehicle Longitude. The dataset can be found online at $https://github.com/bugratu/HeterogenousVehicularComms$.

b) Data Processing: The outlier measurement data with inaccurate GPS coordinate information due to lacking satellite lock is filtered out at the data processing stage. Moreover, RSS values close to the noise floor of the measurement systems are excluded from the training data-set, considering the lower accuracy of the measurements. Finally, bearing information between TX (following) and RX (leading) vehicles are extracted by [15]

$$y = sine[RX_{Long} - TX_{Long}] * cosine[RX_{Lat}] \tag{1}$$

$$\begin{aligned} a &= cosine[TX_{Lat}] * sine[RX_{Lat}] - sine[TX_{Lat}] \\ x &= a * cosine[RX_{Lat}] * cosine[RX_{Long} - TX_{Long}] \end{aligned} \tag{2}$$

$$Bearing_{TX} = arctan(y, x) \tag{3}$$

where sine, cosine, and arctan are trigonometric functions, TX_{Lat}, RX_{Lat} are the latitude, and TX_{Long}, RX_{Long} are the longitude coordinates of the transmitter and receiver measurement locations in radians, respectively. The bearing value is obtained in radians.

c) Random Forest Regression Training: Random Forest is an ensemble learning method, where the output for each input is determined by averaging T random tree outputs, with superior generalization performance [16]. T Random Forest trees are trained by using binary recursive partitioning with bootstrap samples of size N randomly selected input and output training samples. For each unsplit node, k random features are selected from all features. Then, according to the selected split criteria, mean square error (MSE) or mean absolute error (MAE), is calculated to find the best binary split among all binary splits on the k features. The node is splitted into two descendant nodes using the best split. All prediction outputs of T trees are averaged to obtain Random Forest prediction. Therefore, the increased diversity of trees increases the prediction performance of Random Forest model. The easy implementation with parallel computing capability, insensitivity to the amount of input feature data, fitting ability to highly non-linear data, and robustness to noisy measurements are the main motivations to employ Random Forest regression for path loss predictions.

The Random Forest regressor for path loss is trained with the training dataset, $X = \{\mathbf{x_1}, \mathbf{x_2}, .., \mathbf{x_N}\}$ and path loss output $Y = \{PL_1, PL_2, .., PL_N\}$, where $\mathbf{x_i} = (d_i, \phi_i, \upsilon_i)$, for $i = 1, .., N$, N is the number of training samples, d_i is the inter-vehicular distance, υ_i is the relative speed, and ϕ_i is the bearing of the i^{th} training sample, as detailed below:

1) Inter-vehicular Distance (d, meter): The distance between TX and RX vehicles is obtained from GPS coordinates, in the interval of 2.54 m and 71.35 m.
2) Relative Speed (υ, km/h): The relative speed between TX and RX vehicles is measured with the GPS receiver of the SDR, in the interval between 0 km/h-25 km/h relative speed.
3) Bearing (ϕ, Degree): The bearing is calculated from GPS coordinates of TX and RX vehicles according to Eq.3.

To increase the prediction accuracy of the Random Forest path loss models, the maximum number of levels in each decision tree (maximum depth), minimum number of data points allowed in a leaf node (the minimum number of samples per leaf), minimum number of data points in a node before the node split (the minimum number of samples per split), and number of trees in the forest hyper-parameters are optimized through random-search-cross-validation in the pre-defined grid of hyper-parameter ranges. The larger tree with more maximum depth value, the better precision, and the smaller the value of the minimum number of samples per leaf parameter, the higher the precision, which also lead over-fitting. On the other hand, the generalization performance of the model increases with the increasing number of trees, at the cost of more computing resources. The hyper-parameter search is conducted by using $Scikit - Learn$ [17] library with 100 iterations, 5-fold cross-validations. The optimum hyper-parameters are determined as: number of trees = 400, maximum depth = 100, the minimum number of samples per leaf = 1, the minimum number of samples per split = 5 for V-VLC, and number of trees = 2000, maximum depth = 10, the minimum number of samples per leaf = 1, the minimum number of samples per split = 5, for IEEE 802.11p path loss prediction models.

The V-VLC and IEEE 802.11p path loss models are trained with the randomly selected 80 % of the measurement samples, and the rest of the data-set is used to test the performance of the models.

The path loss prediction accuracy of the models is evaluated by using MAE metric formulated as

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} |PL_i - \widehat{PL_i}|, \tag{4}$$

where N_{test} is the total number of test samples, PL_i is the measured and $\widehat{PL_i}$ is the predicted path loss value for the i^{th} model input values.

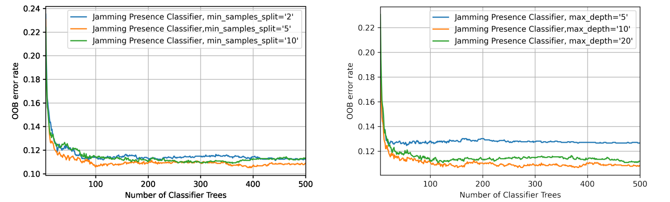
B. IEEE 802.11p V2V Link Jamming Detection

A Random Forest classifier based model is implemented to detect the jamming presence. To jam the IEEE 802.11p V2V link, continuous AWGN with 1 MHz bandwidth is transmitted with varying TX power levels. Higher the TX power of the jamming source leads lower PDR performance for the V2V link.

a) Data Acquisition: Radio jamming presence detection model is trained and validated with the SNR and PDR samples collected through two IEEE 802.11p standard compliant OBU modems. The jamming signals are generated through a Vector Signal Generator (VSG) (Rohde&Schwarz SMBV100A), and transmitted with Mobile Mark ECOM6-5900 antennas. One of the OBU devices is configured to transmit IEEE 802.11p packets on Channel 172 at various power levels of -10 dBm, 0 dBm and 23 dBm. Packet rate is set to 100 Hz and each packet consists of 193-byte test messages. PDR and SNR values are measured and recorded at the receiving OBU by an update rate of 1 Hz. Therefore, each SNR-PDR measurement sample is calculated from 100 packets. AWGN signal with 1 MHz bandwidth is generated by VSG at the center frequency of 5.86 GHz and injected to the channel at varying power levels between [-25 dBm, -45 dBm]. The measurements are performed in a laboratory environment, where TX, RX, and jammer antennas are located on an equilateral-triangle pattern with side lengths of 2 meters. The non-jammed data acquisition is conducted in a leader-follower driving scenario for varying inter-vehicular distances. To train and validate the Random Forest classifier, 9845 measurement samples of SNR, PDR, noise power, TX power and jamming status values are utilized.

b) Data Processing: The SNR and PDR values recorded at the RX OBU are processed for outlier detection and scenario labeling to classify jammed and non-jammed samples. Moreover, SNR values of different scenarios are normalized between $[0, 1]$ since the goal is to classify jamming presence with respect to the variation of SNR and obtained PDR, rather than relying on exact SNR values.

c) Random Forest Classifier: The Random Forest Classifier is an ensemble classifier, which includes set of decision tree classifiers fitted to the training data, using random sampling. For Random Forest classifiers, each individual tree yields a class prediction where the majority of the votes is the model prediction output [16].



(a) (b)

Figure 1: OOB error with varying minimum number of (a) decision tree splits and (b) tree depths.

The proposed Random Forest Classifier is trained with the training data-set, $X_{jam} = \{\mathbf{x_{jam,1}}, \mathbf{x_{jam,2}}, .., \mathbf{x_{jam,N}}\}$ yielding the jamming decision output $Y_{jammed} = \{J_1, J_2, .., J_N\}$, $J_i \in \{0, 1\}$, where 0 denotes non-jammed, 1 denotes jammed measurements, and $\mathbf{x_{jam,i}} = (SNR_i, PDR_i)$, $i = 1, .., N$, N is the number of training samples, SNR_i is the normalized SNR measured by the receiver, and PDR_i is the corresponding PDR obtained at the receiver for the i^{th} sample.

To train the classifier, maximum depth, minimum number of data points allowed in a leaf node and number of decision tree hyper-parameters are selected through the simulations, minimizing Out-of-Bag (OOB) error, which indicates the performance of the model on the data set as depicted in Fig.1.

To evaluate the classification performance of the model, OOB error metric is utilized. To calculate OOB error, one third of the input data-set is separated as OOB samples, and they are not used in the training set to build any tree. Then each tree is tested with the OOB samples, where the misclassification results are averaged over all classification trees.

IV. Performance Results

In this section, we demonstrate the prediction and classification performance of the proposed schemes through real-world measurement data. Moreover, the importance of the input features for path loss prediction models are evaluated.

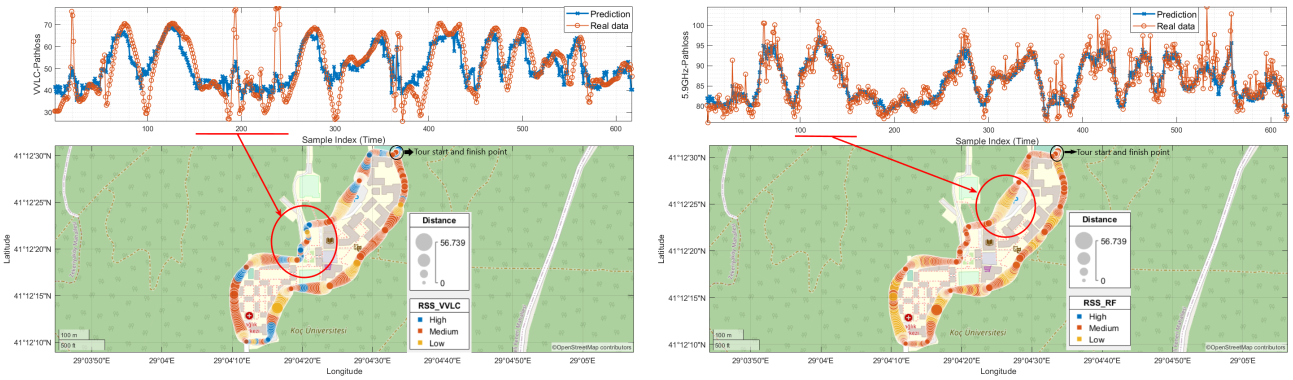
A. Path Loss Model Performance Evaluation

To evaluate the path loss prediction performance of the proposed models, simultaneously acquired RSS measurement data of V-VLC and IEEE 802.11p V2V links for 2 km path is considered. The overall measurements include 6 tours of the considered path at varying speeds, bearings and inter-vehicle distances. We compare the path loss prediction performance of the proposed models with the fitting based log-distance path loss model [18] for IEEE 802.11p and empirical path loss model proposed in [3] for V-VLC.

In the fitting based log-distance path loss model [18], measured path loss values of IEEE 802.11p signals are fitted to log-distance path loss model given by

$$P_L(d)\,[\text{dB}] = P_L(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right) + \chi(\mu, \sigma^2), \tag{5}$$

where $P_L(d)$ is the path loss at inter-vehicular distance d, n is the path loss exponent, d_0 is the reference distance,



(a) (b)

Figure 2: (a) V-VLC and (b) IEEE 802.11p path loss prediction performance for Tour 1

$\chi(\mu, \sigma^2)$ is a normal distributed random variable with mean μ and variance σ^2. Log-normal model fit parameters for the measurement data are as follows: n= 1.4256, d_0= 2.5449 cm, $P_L(d_0)$= 74.7562 dB, μ= -0.1682, σ= 3.0631.

In the empirical path loss model [3], V-VLC path loss values are fitted to the empirical V-VLC path loss model given by

$$P_{distance}[dBm] = \alpha + 10 * \beta * log_{10}\left(\frac{1}{d+\gamma}\right) \qquad (6)$$

where α, β, and γ are the fitting parameters obtained through least-squares fitting, and derived as $\alpha = 148.9$, $\beta = 10.94$, and $\gamma = 56.68$.

Table I demonstrates the MAE performance of the proposed path loss prediction models and the considered fitting based models. For V-VLC links, the proposed model outperforms the fitting based model up to 2.34 dB, whereas for IEEE 802.11p path loss predictions, the proposed model yields up to 0.56 dB better prediction performance.

Table I: Path Loss Models MAE Performance

Evaluation Scenario	**Path Loss Model MAE (dB)**			
	VVLC - RF	VVLC - Fit	IEEE 802.11p - RF	IEEE 802.11p - Fit
Tour 1	5.42 dB	7.50 dB	1.78 dB	2.23 dB
Tour 2	5.05 dB	7.39 dB	1.57 dB	2.09 dB
Tour 3	4.54 dB	6.28 dB	1.69 dB	2.19 dB
Tour 4	4.52 dB	6.00 dB	1.92 dB	2.36 dB
Tour 5	4.48 dB	6.15 dB	1.75 dB	2.21 dB
Tour 6	4.31 dB	5.85 dB	1.87 dB	2.43 dB

Fig. 2 shows the path loss prediction tracking performance of the proposed models for single tour of the measurements. The width of the circles on the map denotes the inter-vehicle distance. The RSS values in [-45, 0] dBm, [-65, -45] dBm , [-90, -65] dBm ranges are marked with blue, red and yellow colors, respectively. In Fig.2 (a), the path loss prediction performance dependency of the V-VLC on the road geometry is depicted, where the performance significantly varies on the curved and straight segments. The proposed model exhibits high path loss tracking accuracy for relatively straight paths, with respect to varying distance and speed. However, the path loss for instantaneous non-line-of-sight (NLoS) transmissions due to sharp maneuvers are under-estimated. On the other hand, Fig. 2 (b) depicts the path loss tracking performance of the proposed IEEE 802.11p path loss prediction model, where the model well tracks the PL variations at varying distances in straight segment. Moreover, under-estimated PL values are mainly observed due to road inclinations and bumps, yielding beam tilt in the elevation [19].

Table II: Normalized Feature Importance Values for Path Loss Predictions

Feature	**Importance**	
	IEEE 802.11p	V-VLC
Inter-vehicle distance	0.75	0.66
Bearing	0.12	0.18
Relative Speed	0.12	0.15

Table II demonstrates the feature importance for the proposed prediction models. The importance of features is extracted through the mean decrease impurity method, where the average reduced impurity by reduced feature is calculated during the training process [16]. The IEEE 802.11p path loss prediction highly depends on the inter-vehicle distance, whereas bearing and speed additionally play an important role for the prediction of V-VLC path loss. The bearing dependency is closely related to the radiation pattern of the LED lights and the FoV of the optical detector. The relative speed increase denotes the increased speed of the leader (RX) vehicle, since the follower vehicle proceeds with constant velocity. Considering the narrow aperture area of the PD at the RX vehicle, vibrations and angular variations due to increased speed yield higher path loss.

The proposed path loss models exhibit better prediction performance compared to fitting based models, since they directly learn from the data instead of relying on assumptions. Moreover, their ability to learn from multiple prediction input variables increases the accuracy, due to the multiple mobility induced variable dependency of the V2V links.

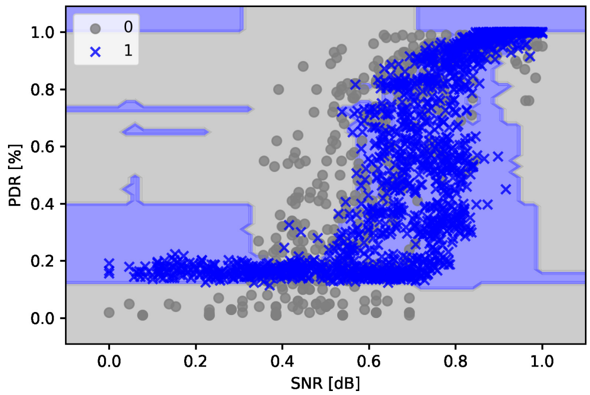


Figure 3: Random Forest classifier based decision regions for non-jammed (0) and jammed (1) SNR-PDR samples

B. Jamming Presence Detection Model Performance

The detection accuracy of the model is evaluated through OOB score and accuracy metrics. To select the optimum hyper-parameters, OOB error is observed for candidate hyper-parameter values as depicted in Fig.1. According to OOB error evaluations, the model is trained with the following hyper-parameters: Number of classification trees = 100, maximum number of levels in each decision tree = 10, and minimum number of data points allowed in a leaf node = 5.

The decision regions of the implemented model is demonstrated in Fig. 3, where the detection accuracy of the jammed samples is higher than the non-jammed samples. This can be explained with the higher correlation between jammed samples, since the same jamming signal is injected to the system with varying power levels.

The proposed model demonstrates 88.3 % OOB score, with the inputs of PDR and SNR, which outperforms the jammer detection accuracy of the Random Forest algorithm [10] and DCNN [13] based studies by 12 % and 3 %, respectively.

V. Conclusion

In this paper, we propose Random Forest regressor based V-VLC and IEEE 802.11p path loss estimation model, and a Random Forest classifier based jamming detection scheme for IEEE 802.11p. The proposed models are demonstrated to provide reliable estimations for path loss prediction and jamming detection. Moreover, the proposed path loss models for V2V V-VLC and IEEE 802.11p links outperform the traditional fitting based models in terms of prediction accuracy, due to their ability learn from the exact data. On the other hand, the jamming detection performance shows that the proposed models can be used to identify the continuous AWGN type jamming. The offline trained models are promising to be utilized in the vehicles, due to the low computing requirements and easily attainable input features. In the future, we plan to extend the proposed models with more input features, such as vehicle speed, jamming signal bandwidth, and jammer distance.

VI. Acknowledgement

This work was supported by CHIST-ERA grant CHIST-ERA-18-SDCDN-001, the Scientific and Technological Council of Turkey 119E350 and Ford Otosan.

References

[1] G. Naik, B. Choudhury, and J.-M. Park, "Ieee 802.11 bd & 5g nr v2x: Evolution of radio access technologies for v2x communications," *IEEE Access*, vol. 7, pp. 70169–70184, 2019.

[2] K. Z. Ghafoor, L. Kong, S. Zeadally, A. S. Sadiq, G. Epiphaniou, M. Hammoudeh, A. K. Bashir, and S. Mumtaz, "Millimeter-wave communication for internet of vehicles: Status, challenges, and perspectives," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8525–8546, 2020.

[3] A. Memedi, H.-M. Tsai, and F. Dressler, "Impact of realistic light radiation pattern on vehicular visible light communication," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2017.

[4] S. Ishihara, R. V. Rabsatt, and M. Gerla, "Improving reliability of platooning control messages using radio and visible light hybrid communication," in *2015 IEEE Vehicular Networking Conference (VNC)*, pp. 96–103, IEEE, 2015.

[5] S. Ucar, S. C. Ergen, and O. Ozkasap, "Ieee 802.11p and visible light hybrid communication based secure autonomous platoon," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8667–8681, 2018.

[6] W. L. Junior, J. Costa, D. Rosário, E. Cerqueira, and L. A. Villas, "A comparative analysis of dsrc and vlc for video dissemination in platoon of vehicles," in *2018 IEEE 10th Latin-American Conference on Communications (LATINCOM)*, pp. 1–6, IEEE, 2018.

[7] P. M. Ramya, M. Boban, C. Zhou, and S. Stańczak, "Using learning methods for v2v path loss prediction," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, IEEE, 2019.

[8] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks: Recent advances and application examples," *IEEE Vehicular Technology Magazine*, vol. 13, no. 2, pp. 94–101, 2018.

[9] D. C. Moreira, I. M. Guerreiro, W. Sun, C. C. Cavalcante, and D. A. Sousa, "Qos predictability in v2x communication with machine learning," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pp. 1–5, IEEE, 2020.

[10] D. Kosmanos, D. Karagiannis, A. Argyriou, S. Lalis, and L. Maglaras, "Rf jamming classification using relative speed estimation in vehicular wireless networks," *arXiv preprint arXiv:1812.11886*, 2018.

[11] D. Karagiannis and A. Argyriou, "Jamming attack detection in a pair of rf communicating vehicles using unsupervised machine learning," *Vehicular Communications*, vol. 13, pp. 56–63, 2018.

[12] O. Puñal, I. Aktaş, C.-J. Schnelke, G. Abidin, K. Wehrle, and J. Gross, "Machine learning-based jamming detection for ieee 802.11: Design and experimental evaluation," in *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, pp. 1–10, IEEE, 2014.

[13] S. Gecgel, C. Goztepe, and G. K. Kurt, "Jammer detection based on artificial neural networks: A measurement study," in *Proceedings of the ACM Workshop on Wireless Security and Machine Learning*, pp. 43–48, 2019.

[14] "NEO-M8." https://www.u-blox.com/sites/default/files/NEO-M8-FW3˙DataSheet˙UBX-15031086.pdf, accessed 2021-03-27.

[15] S. Dabiri and K. Heaslip, "Inferring transportation modes from gps trajectories using a convolutional neural network," *Transportation research part C: emerging technologies*, vol. 86, pp. 360–371, 2018.

[16] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[18] T. S. Rappaport, *Wireless communications: principles and practice*. Upper Saddle River, New Jersey: Prentice Hall, Inc., 2002.

[19] C. F. Mecklenbrauker, A. F. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernadó, T. Zemen, O. Klemp, and N. Czink, "Vehicular channel characterization and its implications for wireless system design and performance," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1189–1212, 2011.