



**HAL**  
open science

# Learning-Based Infrastructure To Vehicle Link Quality Estimation

Raoua Chakroun, Thierry Villemur, Kokouvi Benoit Nougnanke

► **To cite this version:**

Raoua Chakroun, Thierry Villemur, Kokouvi Benoit Nougnanke. Learning-Based Infrastructure To Vehicle Link Quality Estimation. 31st International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2023), University of Split, FESB and Croatian Communications and Information Society (CCIS), Sep 2023, Split, Croatia. hal-04194050

**HAL Id: hal-04194050**

**<https://laas.hal.science/hal-04194050>**

Submitted on 4 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Learning-Based Infrastructure To Vehicle Link Quality Estimation

1<sup>st</sup> Raoua Chakroun  
LAAS-CNRS

Université de Toulouse, CNRS, INSA  
Toulouse, France  
raoua.chakroun@laas.fr

2<sup>nd</sup> Thierry Villemur  
LAAS-CNRS

Université de Toulouse, CNRS, UT2J  
Toulouse, France  
thierry.villemur@laas.fr

3<sup>rd</sup> Benoit Nougnanke

Internet Initiative Japan Inc.  
IJJ Research Laboratory  
Tokyo, Japan  
benoit@ijj.ad.jp

**Abstract**—Timely and up-to-date information about road link quality is essential for the efficient and reliable dissemination of urgent messages in dynamic vehicular network environments. However, existing literature lacks models that provide infrastructure-to-vehicle link quality estimations. This is due to vehicular communication networks’ high variability and complexity (channel variations, complex interference patterns), making it difficult to integrate into a practical analytical model. To address this, we propose a supervised-machine-learning-based prediction model that estimates the Packet Reception Rate (PRR) on the road. Our model updates communication zones to adapt to changes in traffic conditions. It uses a dataset generated from a realistic mobility scenario using the NETSIM simulator and SUMO for training and evaluation. Our performance tests show promising results in terms of prediction accuracy. This work represents the first step toward developing an efficient and reliable scheme for disseminating urgent messages that considers traffic conditions and vehicular mobility changes.

**Index Terms**—Vehicular communications, Quality of Service assessment, Machine Learning

## ACKNOWLEDGMENT

Initial description of this work has been made in the PhD thesis entitled “An Efficient emergency information dissemination scheme for emerging infrastructure-based vehicular networks.” [1].

We also extend our appreciation to Slim Abdellatif for his help. His insightful input and assistance were instrumental in our progress.

## I. INTRODUCTION

In order to support Intelligent Transportation Systems (ITS) services with strict Quality of Service (QoS) requirements, despite rapidly changing traffic conditions, the network must be able to anticipate possible changes related to the evolution of traffic.

Moreover, in vehicular networks, the conditions of the radio signal propagation channel can vary considerably in time and space, which affects the Infrastructure-to-Vehicle (I2V) link quality on the road [2]. In order to ensure reliable and durable performance in such networks, efficient estimation of the link quality on the road is required by some communication services so that alternative or more reliable routes/areas can be selected for data retransmission [3].

The high variability of communication networks with many changing factors seems to us difficult to be integrated into a realistic analytical model. On the other hand, to our opinion, effective link estimation or prediction, started by network state observation, offers a more adaptable and usable approach.

The estimated view of the network’s future state [4] can be integrated into a proactive network control function to guarantee the continuity of services within the required QoS.

This work proposes an Infrastructure to Vehicle (I2V) link quality estimation technique on the road. It relies on Machine Learning techniques to estimate the link quality (Packet Reception Rate (PRR)) in each zone defined by a  $40 \times 40m^2$  square, which allows us to determine the reception quality of each square area. A supervised learning model has been proposed, with the particularity of using traffic information and “Hello” default exchange messages between vehicles and infrastructure, thus limiting transmissions’ overhead.

In order to train and evaluate the proposed model, a dataset has been generated, as the ones available from the literature do not fit the needs of our study. This dataset has been created using mainly the NETSIM framework and is based on mobility traces.

This paper is organized as follows. Section II presents a synthesis of existing works in the scientific literature. Section III gives a general model overview for the proposed link quality estimation model. Section IV details the proposed framework, while section V describes the dataset and the model. Section VI focuses on the experimental part. It first presents the metrics considered, then analyzes the evaluation results for the proposed model. Finally, section VII describes the more global application schema.

## II. RELATED WORK

During the last decade, wireless link quality estimation and characterization have been studied in wireless mobile networks and are considered crucial for reliable communication [2], [5]. Due to the dynamic nature of the vehicular environment, the statistical channel models do not predict wireless link quality with high accuracy [5].

We have not found any standard definition for link quality nor a standard unit of measure for the metric [6] [7]. In this

work, we consider that link quality can be generally thought of as a link description derivative of throughput or reliability.

Machine Learning (ML) techniques [2] have been recently used for predicting link quality in wireless environments [8].

In [9], [10], authors propose ML-based algorithms to predict Vehicle-to-Vehicle (V2V) path loss, proving that such models' application offers better performance than traditional analytical models utilization (e.g., log distance path loss model). Benrhaim and al. [11] method relies on periodically beacons exchanges between vehicles to estimate V2V link reception quality. They propose a Bayesian network-based scheme at different locations in the zone covered by the transmission range of the sender for the estimation. This is the only work that estimates the road links' quality. Estimation results show good accuracy. However, the sample of parameters considered in all the simulations remains small and limited.

Most of the works proposed in the literature to estimate the vehicular communication links' quality assume simplifications of vehicle mobility. We note these works concern V2V links, where only one is interested in estimating V2V link quality on the road, and approaches based on machine learning techniques generally present the best performances for both problems. Our work is the first one that focuses on I2V wireless quality links on the road. The proposed method for estimating road links' quality excludes any vehicle mobility or communication range assumption [1].

### III. PROPOSED MODEL OVERVIEW

The proposed work aims to estimate the quality of I2V links on the road, precisely, the Packet Reception Rate (PRR) of RSUs' transmission for each small square zone in a predefined region. Each region is represented by a matrix of small squares of  $40 \times 40m^2$  as described in [12] (see Fig. 2). These squares are used to identify gray zones in the region, where their PRR is less than a predefined threshold (90% in our case) [13]. The gray zones are identified within the controller. Fig. 1 shows the network architecture with the key elements of the proposed approach. A Software Defined Network (SDN) controller is assigned to each region as described in [14]; it manages all the Road Side Units (RSU) that provide V2I wireless connectivity within its region.

We assume all vehicles are equipped with a GPS module and can send information, such as their position ( $P(x, y)$ ) and the packet response of the "Hello" message received from the RSU at the time of association and beacon messages exchanges. This information is periodically collected by each RSU and hosted in the cloud or shared directly with the network controller to extract features used as input to the model in order to estimate the PRR. Finally, we assume that an RSU entity reports information about each newly associated vehicle to the SDN controller.

The proposed model  $M_{PRR}$  based on machine learning techniques runs within the SDN controller. It mainly merges the information reported by the RSUs. Centralizing the  $M_{PRR}$  model within the SDN controller, which manages a collection of nearby RSUs, allows the  $M_{PRR}$  model to learn the link

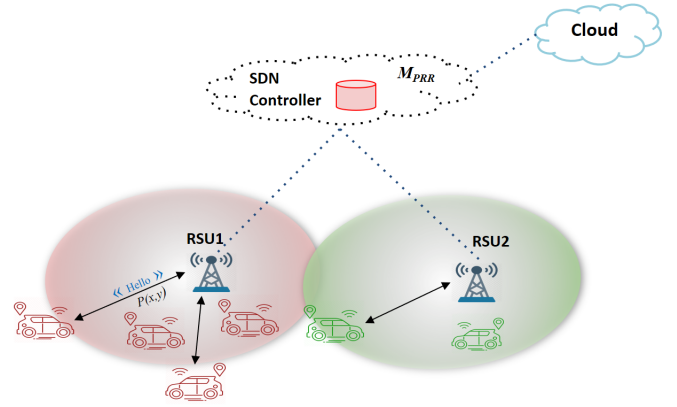


Fig. 1. Key elements of the proposed approach [1].



Fig. 2. Considered geographic map [1].

quality variations in the concerned coverage area. Finally, this information is combined with other data to extract features used as input to the model to estimate PRR in each small square zone. The design of the  $M_{PRR}$  model is detailed in the next sections.

### IV. SDN-ENABLED MACHINE LEARNING ROAD I2V LINK QUALITY PREDICTION

#### A. Supervised Machine Learning

Our work consider supervised machine learning, where learning is performed from a labeled dataset so the training data contain the desired solutions. Formally, in a dataset  $D$  defined by  $D(x_1, y_1), \dots, (x_n, y_n)$ , the training of the model  $M$  aims to find the best relationship between the inputs  $X$ , called predictors, and the outputs  $y$ , called labels,  $y = M(X)$ , such that, for new input data  $X_n$  whose outputs are unknown, the model can predict the corresponding output  $\hat{y}_n = M(X_n)$  with good accuracy. We distinguish two types of supervised learning: regression, when the value to be predicted is a continuous real number,  $y \in \mathbb{R}$ , and classification when  $y$  belongs to a finite set  $C = c_1, c_2, \dots, c_n$  called classes.

Furthermore, we consider the so-called ensemble learning techniques [15], allowing us to train several models (of the

same or different techniques) and combine their predictions. This technique represents one of the most popular and powerful supervised algorithms. It allows for the design of a generalized model and avoids overfitting. The model retained is the *Random Forest* technique [16], which trains a set of decision tree models.

The PRR estimation problem has been modeled as a regression problem. The estimated variable is the PRR in each map's little zone (square) under the SDN controller coverage. In the following sections, we present the considered training variables (or features) to design our models. These features are designed according to the objectives, i.e., i) features that require a minimum of information from the vehicles and ii) features that are independent of the network technology used. Next, we detail the techniques used to train and calibrate the model parameters.

### B. Proposed framework

The propagation channel characteristics of vehicular networks differ significantly from other wireless systems. The physical environment in vehicular channels is expected to experience random variations caused by several factors, including mobility patterns and rapid changes in traffic density, path loss, and environmental effects. The rapid temporal variability and non-stationary channels are rationales for developing a unique framework for predicting the link quality of vehicular networks. In this work, we aim to build a machine learning model capable of predicting the PRR on the road with minimum error.

The classical machine learning workflow [17] is composed of the following steps: Problem formulation, Data Collection, Data Analysis, Model Construction, and Model Validation, and the last step is Deployment and Inference. By combining this workflow and leveraging SDN [18], we come up with the SDN-enabled machine learning PRR prediction framework in Fig. 3. The framework learns some properties of an historical dataset and leverages the learned properties to provide good estimations on new observations.

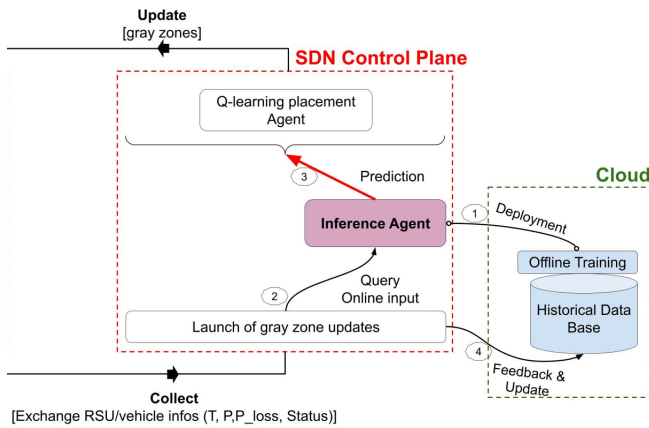


Fig. 3. Packet Reception Ratio inference framework.

From Fig. 3, the framework's workflow is as follows. Firstly, the prediction model  $M_{PRR}$  is constructed offline by training and tuning the parameters of the historical data. The historical dataset may be composed of a large number of samples. Each sample represents a combination of features' values and the associated target value. Collected data include vehicle positions, message identifications, and quality of communication exchanges (signal power, losses, packet status). The features' description, collection, and processing are detailed in the next section.

The designed model  $M_{PRR}$  based on machine learning techniques is then deployed (Fig. 3, step (1)) in the Inference Agent instance. The PRR inference is used to predict gray zones in the region. We can suppose that this approach can be used proactively, i.e., we assume the controller has a historical idea about traffic conditions change (e.g., traffic density at peak hours), so it updates these zones according to these schedules (e.g., every two hours or three times a day).

The online input for each zone after processing (Fig. 3, step (2)), composed of vehicle density and the packet loss exponent in the concerned zone (These input parameters are detailed in Tab. II in the next section), is got when the controller launches the updates of gray zones for the geographic map under its coverage. Taking this input, an inference of the PRR in each zone (small square) is made, so gray zones are identified (Fig. 3, step (3)).

Finally, when the process is complete, its observed output is also collected, and the historical dataset can be updated with the newly collected data (Fig. 3, step (4)). Therefore, having the database up-to-date is essential and will allow considering new dynamics from the traffic changes. The historical data gathering and online update of the historical data with the newly collected data form a base for our framework. The historical data could be enriched from the cloud [1].

## V. LEARNING-BASED MODELING

This section begins with the problem formulation. For the PRR inference, the target metric is a continuous variable; its prediction is a regression problem.

### A. Dataset and analysis

The case study comes from an urban environment. In order to evaluate the proposed model in a realistic framework, it is necessary to have a data set that meets the following requirements:

- Network coverage: it considers cells with varied communication ranges (small and large coverage) by manipulating the transmission power of RSUs. In addition, it is necessary to know the geographical positions of these entities to calculate the distance.
- Road traffic: we collect vehicle data (location) driving on most roads (main and secondary) covered by a given RSU.
- Size of the dataset: a collection of data for a long duration by varying the density of the vehicles, the transmission powers, the position of the RSUs, and the coefficients

of path loss, allowing the exploration of the temporal variations of the measured metrics.

TABLE I  
COLLECTED DATASET

Name	Feature Description
Packet Id	Sent packet identifier
RSU Id	Sender RSU identifier
Vehicle Id	Receiver vehicle identifier
$P(x, y)$	Vehicle position
$T(i)$	Transmission power of the sender RSU
$P_{loss}(i)$	Path loss exponent
Packet status	"received" or "not-received"

The dataset used in this work was generated using the microscopic road traffic simulator SUMO coupled with the event-based network simulator NETSIM as described in [19]. NETSIM simulates the DSRC protocol stack (signal strength, handover, connectivity), while SUMO handles vehicle mobility. We extract the area of  $2 \times 2 km^2$  of a European-like city (namely, Toulouse, France), with high traffic densities (Urban environment), large buildings affecting the quality of signals on the road, and irregular road structures, using Open Street Maps (OSM). The area is decomposed in regular  $40 \times 40 m^2$  square zones. In each zone, the PRR parameter will be estimated. For network simulation parameters, we run 196 simulations where each RSU broadcasts a control message every 100ms for 500s (27 hours in total), varying each time the transmission power, the positions of the RSUs, vehicle densities, and the path loss coefficients.

The dataset generated and collected after each simulation is described in Tab. I. Path loss is the reduction in the power density of an electromagnetic wave as it propagates through space. It may be due to many effects, such as free reflection, aperture-medium coupling loss, and absorption. The path loss exponent varies between 2–5 depending on the surrounding environment coherence bandwidth and Doppler shift. In reality, the path loss exponent can be calculated according to several parameters (e.g., the transmission power, the distance from the sender, the wavelength of the carrier wave (which depends on whether and obstacles), etc.), but as we work with data collected by a simulator, we use this parameter directly from the simulator. However, an SDN controller can use its global view and knowledge to calculate this parameter based on the real-time information (distance between RSUs and zones, urban or other environments, weather, etc.) shared by the RSUs and the cloud [3]. After each packet, we record the vehicle position (to identify the zone identifier), the vehicles that received the message, and the packet status (success or error). The results of simulations generate a dataset of 52007 observations. Figure 4 shows the number of samples generated for each PRR range.

We predict the PRR by a small road zone/square (only squares containing road areas are considered) instead of a road because the quality of links at the beginning of the road could differ from the end and the center. Hence, predicting the PRR by zone (a small portion of the road) is more precise and

accurate. We realize some data processing to have our dataset suitable for our model, which allows us to predict the PRR in each zone as described in Tab. II. First, we consider the vehicle's position to define the concerned zone. Then for each simulation, we calculate the density of vehicles per zone and hour, the average distance between the zone and the sender, and the PRR (the number of packets received by vehicles in the zone compared to the number sent by RSUs to those vehicles).

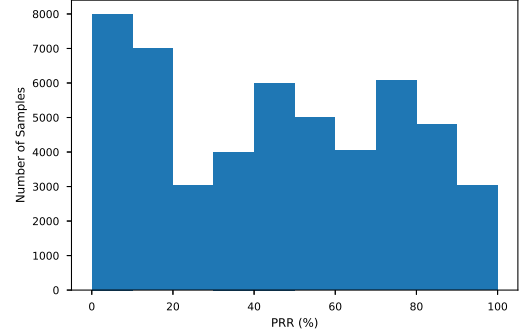


Fig. 4. Number of samples generated for each PRR range [1].

## B. Model training

The features of the  $M_{PRR}$  model (vehicle density, distance, transmission power, and path loss exponent) are designed according to two main criteria i) those requiring a minimum of information from the vehicles and ii) those independent of the network technology used.

TABLE II  
PARAMETERS AND NOTATIONS

Name	Feature Description
$Z_i$	Zone identifier
$V_d(i)$	Vehicle density per hour in a zone
$D(i)$	Average distance between the sender and the zone
$T(i)$	Transmission power of the sender RSU
$P_{loss}(i)$	Path loss exponent
$\widehat{PRR}(Z_i)$	Packet Reception Rate in the concerned zone

---

### Algorithm 1 Road PRR Estimation

---

**Require:**  $Z_i, i \in 0, \dots, N$  List of road zone ID

$P_{loss}(i)$  per zone: path loss per zone

$T(i)$ : Transmission power of the RSU that covers the zone

$V_d(i)$ : Vehicle density in the concerned zone

**Ensure:**  $\widehat{PRR}(Z_i)$ : the Packet reception ratio by zone

**for**  $i = 0$  to  $i = N$  **do**

$$D(i) = \sqrt{(x_{RSU} - x_i)^2 + (y_{RSU} - y_i)^2}$$

/\* distance between the zone  $i$  of coordinates  $(x_i, y_i)$  and

the RSU that covers it of coordinates  $(x_{RSU}, y_{RSU})$  \*/

$$\widehat{PRR}(Z_i) = M_{PRR}(D(i), P_{loss}(i), T(i), V_d(i))$$

**end for**

---

Algorithm 1 summarizes the information used as input and the features considered by the  $M_{PRR}$  model to estimate the

packet reception ratio for each road zone covered by the SDN controller. From a dataset composed of the various attributes listed previously, labeled by the desired results (PRR), a training of the models is performed offline to find the best representation between features and labels. As mentioned previously, we use the Random Forest (RF) technique to train a set of Decision Tree (DT). First, each tree is trained with a randomly chosen subset of the dataset. Furthermore, during this process, the concerned attribute is randomly chosen when splitting a node. Finally, predictions made by each tree are aggregated to produce a global prediction at the end: the average of the values estimated by all the trees is used in the regression case [1].

## VI. PERFORMANCE EVALUATION

The goal is to evaluate the model's capabilities to estimate the packet reception ratio in each square to identify the region's gray zones. We analyze the results based on the data visualization of the studied scenario. In addition, we identify the strengths and limitations of the proposed model [1].

We use 75% of the dataset for model training and the remaining 25% for testing, a commonly used ratio. Then, for each input  $i$  in the test set  $X$ , we compute the corresponding output  $\hat{y}_i = M(X_i)$  using the model  $M$  (model after the training phase). Then we compare this output with the real value  $y_i$ . In this way, we compute the prediction accuracy of each model using the performance metrics presented below.

### A. Performance metrics

To evaluate the accuracy of the proposed ML performance models, we consider two evaluation metrics:

- **Prediction score**  $R^2$  (Eq. 1): it represents which part of the variance in the dependent variable is predictable from the independent variables. In other words, it is the proportion of correctly predicted samples. The most precise regression model would be the one with a relatively high  $R$  squared, close to 100 when expressed in percentage.
- **Normalized Mean Absolute Error (NMAE)** (Eq. 2): it defines the average of the absolute differences between the estimated and observed values of  $PRR$ . We want the  $NMAE$  to be as small as possible.

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \text{ with } \bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (1)$$

$$NMAE(y, \hat{y}) = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\bar{y}}, \text{ with } \bar{y} = \frac{\sum_{i=1}^n y_i}{n} \quad (2)$$

### B. Results

For a subset with  $n\_samples$  of 10000, where 7500 samples are used to train the random forest model, we obtain a score  $R^2 = 90.70\%$  and  $NMAE = 5.20\%$  on the 2500 remaining sets. Using all the datasets (52007 samples) generated through NETSIM simulations give scores around 95%, with  $NMAE$  of around 5%. These results prove the ability of the ML-performance approach to provide accurate predictions.

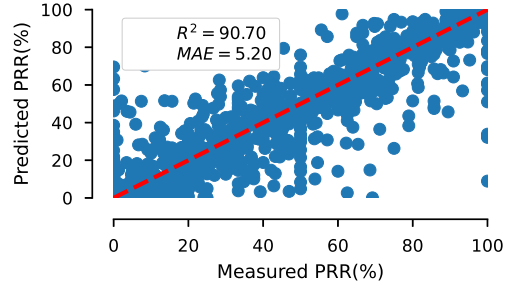


Fig. 5. Predicted versus real observations ( $K = 10000$  samples) [1].

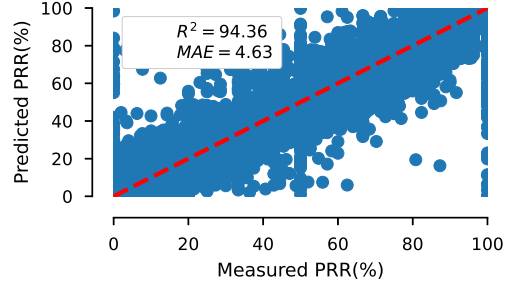


Fig. 6. Predicted versus real observations ( $K = 52007$  samples) [1].

Figure 5 shows the real observed test points and their corresponding predictions with the random forest model for  $K = 10000$  samples. But, of course, when the subset contains all the initial dataset ( $K = 52007$  samples), the prediction accuracy is better, with a score almost equal to 0.95, as shown in Figure 6.

The performance of the proposed model mostly meets the needs of a proactive network control based on the estimation of the quality of I2V links on the road, which allows defining and updating gray zones periodically in the region in order to update the rebroadcast zones.

Figure 7 represents the features' importance, where the Mean Decrease in Impurity (MDI) is used to calculate each feature's importance. It is calculated as the sum over the number of splits (across all trees) that includes the feature, proportionally to the number of samples it splits. This shows that the distance and the packet loss coefficient (where we consider the geographic area, urban/non-urban, the size and presence of the buildings in the area, the weather, etc., that influence communication channels) are the most important features. Indeed, the absence or disregard of these features will significantly decrease the model score accuracy on the test set (e.g., by removing the path loss from the features, the score accuracy can decrease up to 55%).

## VII. APPLICATION

This work constitutes the first step of a more global communication service chain to improve the broadcast of critical messages. The proposed Machine Learning algorithm is used to precisely assess the communication quality of  $40 \times 40m^2$

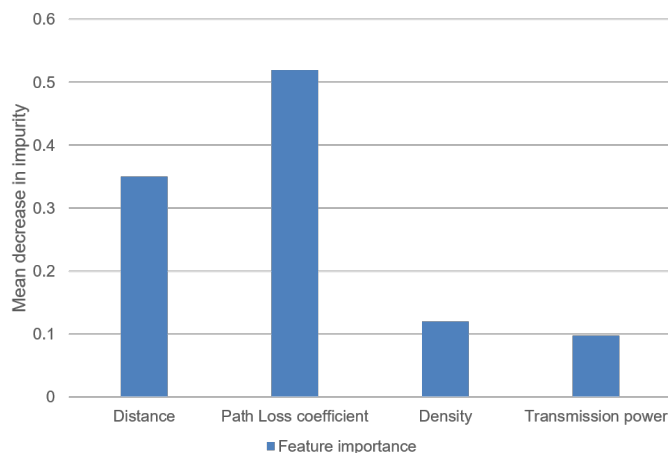


Fig. 7. Feature importance using Mean Decrease in Impurity (MDI) [1].

square zones in a geographical area. Then, it can be used as an input of a Reinforcement Learning (RL) placement algorithm proposed in [12]. According to their communication characteristics, the goal of the RL placement algorithm is to select the optimal number of rebroadcast square zones where relay vehicles could retransmit messages to vehicles located in gray zones (estimated by our method). This approach has been tested in the case of critical communications by designing and assessing an emergency service to ensure reliable alert message dissemination [20].

### VIII. CONCLUSION. FUTURE WORK

Estimating link quality on the road paves the way for intelligent and efficient network control. In the proposed approach, we mainly exploited the road packet reception ratio identification as our model's main learning variable. The performance tests showed excellent results in the vast majority of cases.

The model is trained offline using data collected under varying traffic conditions to approximate real mobility and traffic conditions that differ from one hour to another daily. However, the trends captured by the models during real-world training can vary even more when considering new installations and reconstructions in the area (new buildings, facades, parking, etc.), thus changing several parameters such as the communication medium and path loss. Therefore, the controller must be updated by these changes in order to update its road zones and be able to recalculate the path loss by each zone efficiently.

On the other hand, the service provider (ITS) can modify the network parameters in order to optimize its network (modification of the cell coverage, addition, or deletion of a cell). This can affect the model performances. Re-training can occur if the prediction error exceeds a certain threshold (set according to the service using the predictions) with the new data collected and obviously considering these new conditions.

The presented work focuses on a specific urban environment (Toulouse, France) and utilizes simulated data. The next step

will be to consider other urban areas and real-world scenarios. Real data should reinforce the model's applicability. Network parameters, as path loss exponent calculation, should be fine-tuned with real data feedbacks.

### REFERENCES

- [1] R. Chakroun, "an Efficient emergency information dissemination scheme for emerging infrastructure-based vehicular networks," PhD. thesis, University of Toulouse, France, Nov. 2022.
- [2] G. Cerar, H. Yetgin, M. Mohorčić, and C. Fortuna, "Machine learning for wireless link quality estimation: A survey," *IEEE Communications Surveys & Tutorials* Vol. 23 No. 2, pp. 696–728, 2021.
- [3] Y. A. Zakaria, E.K.I. Hamad, A.S. Abd Elhamid, and K.M. El-Khatib, "Developed channel propagation models and path loss measurements for wireless communication systems using regression analysis techniques," *Bulletin of the National Research Centre*, Vol. 45, No. 54, 2021.
- [4] N. Bui, M. Cesana, S.A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques," *IEEE Communications Surveys & Tutorials*, Vol. 19 No. 3, pp. 1790–1821, 2017.
- [5] L. Tang, K.C. Wang, Y. Huang, and F. Gu, "Channel characterization and link quality assessment of IEEE 802.15. 4-compliant radio for factory environments," *IEEE Transactions on industrial informatics*, Vol. 3, No. 2, pp. 99–110, 2007.
- [6] N. Baccour, A. Koubâa, L. Mottola, M. A. Zúñiga, H. Youssef, C. A. Boano, and M. Alves, "Radio link quality estimation in wireless sensor networks: A survey," *ACM Transactions on Sensor Networks (TOSN)*, Vol. 8, No. 4, pp. 1–33, 2012.
- [7] C. J. Lowrance, and A. P. Lauf, "Link quality estimation in ad hoc and mesh networks: A survey and future directions," *Wireless Personal Communications*, Vol. 96, No. 1, pp. 475–508, 2017.
- [8] H. Ye, L. Liang, G. Ye 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] B. Turan, A. Uyrus, O. Nuri Koc, E. Kar, and S. Coleri, "Machine Learning Aided Path Loss Estimator and Jammer Detector for Heterogeneous Vehicular Networks," *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2021.
- [10] P. M. Ramya, M. Boban, C. Zhou, and S. Stańczak, "Using learning methods for v2v path loss prediction," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1-6, 2019.
- [11] W. Benrhaiem, and A. Senhaji Hafid, "Bayesian networks based reliable broadcast in vehicular networks," *Vehicular Communications*, Vol. 21, pp.100–181, 2020.
- [12] R. Chakroun, S. Abdellatif and T. Villemur, "Q-Learning Relay Placement for Alert Message Dissemination in Vehicular Networks," *19th International Conference on Mobile Systems and Pervasive Computing (MobiSPC)*, Niagara Falls, Canada, pp. 222–230, August 9-11 2022.
- [13] R. Meireles, M. Boban, P. Steenkiste, O. Tonguz, and J. Barros, "Experimental study on the impact of vehicular obstructions in VANETS," *IEEE Vehicular Networking Conference*, pp. 338–345, 2010.
- [14] S. Toufga, S. Abdellatif, H.T. Assouane, P. Owezarski, and T. Villemur, "Towards Dynamic Controller Placement in Software Defined Vehicular Networks," *Sensors*, Vol. 20, No. 6, pp. 1701, March 2020.
- [15] O. Sagi, and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2018.
- [16] L. Breiman, "Random forests," *Machine learning*, Vol. 45, No. 1, pp. 5 – 32, 2001.
- [17] M. Wang, Y. Cui, X. Wang, S. Xiao, and J. Jiang, "Machine learning for networking: Workflow, advances and opportunities," *IEEE Network*, Vol. 32, No. 2, pp. 92–99, 2017.
- [18] N. Foster, N. McKeown, J. Rexford, G. Parulkar, L. Peterson, and O. Sunay, "Using deep programmability to put network owners in control," *ACM SIGCOMM Computer Communication Review*, Vol. 50, No 40, pp 82–88, 2020.
- [19] J. S. Weber, M. Neves, and T. FERRETO, "VANET simulators: an updated review," *Journal of the Brazilian Computer Society*, Vol. 27, No. 1, pp. 1–31, 2021.
- [20] R. Chakroun, S. Abdellatif and T. Villemur, "LAMD: Location-based Alert Message Dissemination scheme for emerging infrastructure-based vehicular networks," *Internet of Things*, Vol. 19, Article 100510, 2022.