Vehicle-to-Everything (V2X) Datasets for Machine Learning-Based Predictive Quality of Service

Published Version:
Skocaj, M., Di Cicco, N., Zugno, T., Boban, M., Blumenstein, J., Prokes, A., et al. (2023). Vehicle-to-Everything (V2X) Datasets for Machine Learning-Based Predictive Quality of Service. IEEE COMMUNICATIONS MAGAZINE, 61(9), 106-112 [10.1109/mcom.004.2200723].

Availability:
This version is available at: https://hdl.handle.net/11585/964871 since: 2024-03-01

Published:
DOI: http://doi.org/10.1109/mcom.004.2200723

Vehicle-to-Everything (V2X) Datasets for Machine Learning-based Predictive Quality of Service

Marco Skocaj*, Nicola Di Cicco†, Tommaso Zugno‡, Mate Boban‡, Jiri Blumenstein§, Ales Prokes§, Tomas Mikulasek§, Josef Vychodil§, Konstantin Mikhaylov¶, Massimo Tornatore† and Vittorio Degli-Esposti*

* DEI, University of Bologna, & WiLab, CNIT, Italy
† DEIB, Politecnico di Milano, Italy
‡ Huawei Technologies
§ Brno University of Technology, Czech Republic
¶ Centre for Wireless Communications (CWC), University of Oulu (UO), Finland
corresponding authors: *marco.skocaj@unibo.it, nicola.dicicco@polimi.it*

***Abstract*—We present two datasets for Machine Learning (ML)-based Predictive Quality of Service (PQoS) comprising Vehicle-to-Infrastructure (V2I) and Vehicle-to-Vehicle (V2V) radio channel measurements. As V2V and V2I are both indispensable elements for providing connectivity in Intelligent Transport Systems (ITS), we argue that a combination of the two datasets enables the study of Vehicle-to-Everything (V2X) connectivity in its entire complexity. We describe in detail our methodologies for performing V2V and V2I measurement campaigns, and we provide illustrative examples on the use of the collected data. Specifically, we showcase the application of approximate Bayesian Methods using the two presented datasets to portray illustrative use cases of uncertainty-aware Quality of Service and Channel State Information forecasting. Finally, we discuss novel exploratory research direction building upon our work. The V2I and V2V datasets are available on IEEE Dataport [1], [2].**

***Index Terms*—predictive QoS, V2X, V2V, V2I, machine learning, deep learning.**

I. Introduction

One of the core objectives of next-generation Beyond 5G (B5G) and 6G wireless systems is the minimization of human intervention in network management. On the one hand, network operators need new and more cost-effective solutions to match the increased management complexity. On the other hand, new services like autonomous driving set stringent requirements on reliability and guaranteed Quality of Service (QoS). As such, integrating autonomous capabilities and proactive decision making in network management constitutes a timely research challenge. Predictive Quality of Service (PQoS) was introduced as a real-time mechanism to provide autonomous systems with advance notifications about upcoming QoS changes [3]. In contrast to reactive strategies, PQoS allows for proactive decision-making and ensures agile adaptation and continuity of service following a *predict-adapt-inform* closed loop principle. In Vehicle-to-everything (V2X) applications (e.g., trajectory prediction, high-density platooning, teleoperated driving, etc. [3]) PQoS was introduced to accommodate various configurations (such as automation level, inter-vehicle gap, etc.) and allows for configuration adjustments in response to changes in QoS. Prompt adjustment may favor service continuity, comfort, and safety. Besides that, predictive knowledge of future network conditions might trigger a series of network procedures aimed at ameliorating channel conditions and QoS (e.g., early handover decision or Up Link (UL)/Down Link (DL) power control).

M. Skocaj and N. Di Cicco are co-first authors of this paper.

Recently, two main factors contributed to the rise of PQoS and autonomous adaptivity in network management: technological breakthroughs in the fields of artificial intelligence (AI) and machine learning (ML), and unprecedented availability of data and computing resources in many engineering domains [4]. Unfortunately, the lack of high-quality publicly-available datasets hinders progress in this domain. As such, with the goal of providing a reliable landmark to the research community, we propose two high-quality channel measurements testbed methodologies and datasets for Machine Learning (ML)-based PQoS in V2X applications. The V2X systems are primarily characterized by two distinct communication links, namely Vehicle-to-Vehicle (V2V) links for inter-vehicle communication and Vehicle-to-Infrastructure (V2I) links for backbone connectivity [5]. These links exhibit diverse characteristics, necessitating distinct technology, measurements, and optimization procedures. In the case of V2V links, performance is predominantly defined by channel conditions, which are susceptible to variations due to factors such as inter-vehicle distance, velocity, and the surrounding environment. mmWave communications have been envisioned as a potential communication technology for V2V communication, especially for applications that require high data transfer rates and low latency, such as cooperative driving scenarios where fast response times are critical [6], [7]. Novel beamforming and power control optimization techniques are timely research challenges associated to the use of mmWave frequencies in V2V communications, where Line-of-sight (LoS) is still nowadays an utmost requirement. In order to advance the study of proactive forecasting and optimization techniques, in Section II-B, we report a testbed methodology for *V2V Channel Measurement* and illustrative *Channel Impulse Response (CIR)s measurements* at 60 GHz collected in the campus of Brno University of Technology. Conversely, V2I communications are often performed via sub-6-GHz communications and rely on network Key-Performance Indicator (KPI) as the standard metrics. Thus, this study presents a KPI-oriented measurement campaign performed in the city of Munich. Overall, a combination of the two datasets enables studying the V2X connectivity from the points of view of both high-levels KPIs and raw channel measurements. One of the primary objectives of this study is to demonstrate how a similar algorithmic framework can be used to deliver PQoS for both V2I and V2V communications. We remark

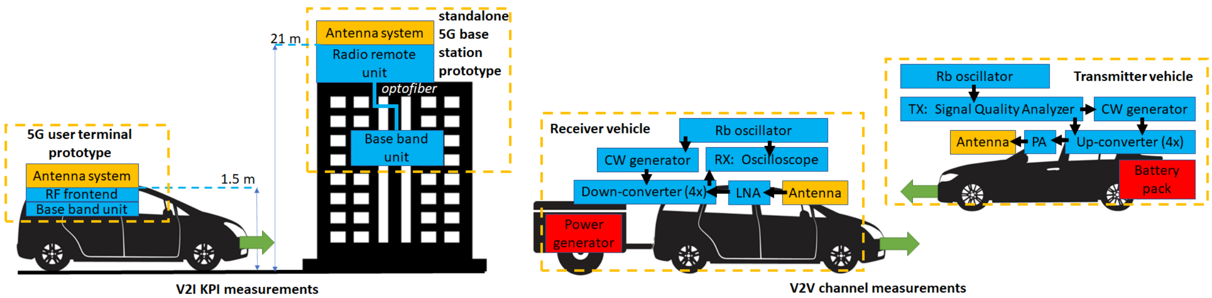


Fig. 1: Structural diagrams of the setups: for V2I KPI measurements on the left, for V2V channel measurements on the right.

that V2X trials should be performed on the field, since it is nearly impossible to obtain realistic results in laboratory environments. However, field results may significantly differ in different scenarios and urban environments. Therefore, we consider the contribution of this paper significant not only for disseminating specific data set(s), but also to propose a reference and repeatable testbed methodology for dataset collection. In particular, we enable the generation of comparable datasets, a better understanding of environmental effects on the V2X connectivity, and the use of datasets for novel ML-based predictive QoS methods. Moreover, we illustrate novel methods and experimental results for ML-based QoS forecasting, differentiating from previous literature which did not account for the evolution of QoS KPIs over time. In this regard, if new datasets are collected and structured similarly to the presented ones, methods and algorithms developed and tested in Section IV can be employed out-of-the-box. The two datasets have been originally presented in [8] and [9], respectively, and are shared for scientific collaboration within the European COST INTERACT 20120 cooperation action.

II. Testbed Design

This Section describes the testbed design and the measurement campaigns conducted for building the two presented datasets. The structural diagrams of the two setups depicted in Fig. 1 are discussed in more detail in the following.

A. V2I KPI Measurements

The measurement scenario illustrated in Fig. 1 was designed to resemble V2X use cases with stringent requirements in terms of UL throughput, such as teleoperated driving and/or local dynamic map update. We selected a commercial area located in the northwest part of Munich, Germany, characterized by the presence of relatively high buildings and two-lane streets with vegetation on the sides. In this urban area, the average speed limit is 50 km/h and the traffic is moderately heavy with stable flows and no congestion.

The measurements were conducted using a prototype 5G standalone system composed of one Huawei's base station and one user terminal. The base station was composed of three units, i.e, the baseband unit, the radio remote unit, and the antenna system. The radio remote unit and the antenna systems were placed on the rooftop of a building with a height of 21 m above the ground, at the location labeled with a black triangle in Fig. 2. The baseband unit was placed inside a server room of the same building, and was connected to the radio remote unit by means of optical fiber links.

Similarly, the user terminal was composed of a baseband unit, a radio frequency frontend, and an antenna system. All the units were installed on a compact car, with the antenna placed on the roof at a height of 1.5 m. The system was operating with a carrier frequency of 3.41 GHz and a bandwidth of 40 MHz. The antenna gain was 15.5 dBi at the base station side and 5 dBi at the user side. Both for the DL and UL channels we used a fixed modulation and coding scheme, with modulation order of 2 and 6, and code rate of 0.43 and 0.45, respectively. The transmission power of both the base station and user terminal was set to 28 dBm, and the UL channel was using two Multiple-Input Multiple-Output (MIMO) layers. The user terminal was connected to the base station and upstreaming data at a constant rate of 40 Mbps. Conversely, an application installed at the base station side was downstreaming data at a constant rate of 2.08 Mbps. Additionally, the car was equipped with a Global Positioning System (GPS) sensor measuring coordinates, speed, and time.

To build a rich dataset for ML training tasks as well as to ensure the consistency of the results, the vehicle traversed 10 times the 8-shaped track depicted in Fig. 2. In particular, the vehicle started from the position represented by the black star in the figure, turned left, and drove along the track clockwise. For training and testing of a ML model, repeated traversals of the same trajectory ensure that the data can be easily split both temporally (e.g., one round for training, and the remainder for testing) and spatially (e.g., south locations only for training, and the remainder for testing).

During the trial we monitored multiple metrics, including throughput, Signal-to-Interference-and-Noise-Ratio (SINR), and Modulation-Coding Scheme (MCS), both for the DL and the UL channels, together with the vehicle's GPS data. We measured these metrics for the entire duration of the experiment with a sampling period of 1 second. We note that, while the original study focused on predicting uplink throughput from uplink features via Supervised Learning, a similar study can be performed for predicting downlink throughput from downlink features. Furthermore, since location information (longitude, latitude, altitude), the speed of the vehicle, and uplink and downlink modulation and coding scheme are collected, it is also possible to use the dataset for uplink/downlink modulation scheme classification, trajectory prediction, etc. Overall, the composition of our dataset allows

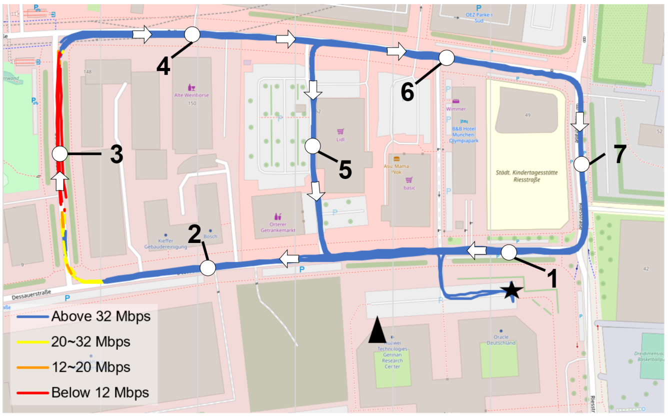


Fig. 2: Map showing UL throughput measurements. The test vehicle route traversed the double loop 10 times and is marked with GPS track in 4 colors representing throughput intervals. Base station antenna at top of the building (black triangle), 21m above ground. Vehicle antenna height: 1.5m.

for a diversified set of potential ML tasks.

B. V2V Channel Measurements

A measurement campaign for collecting Channel Impulse Responses (CIRs) in the 60 GHz frequency band has been carried out in Brno, Czech Republic to evaluate the time-domain dispersion characteristics of V2V channels [9]. Here, we present in further detail the testbed design and we extend the use of the presented dataset to the task of PQoS in V2V communications, as further elaborated in section IV-B. Given the crucial importance of LoS communication at 60 GHz, our experimental methodology was designed to ensure that this condition was met. Specifically, we conducted our investigation in a controlled traffic setting where LoS was consistently maintained throughout the duration of our CIR recording process. To further minimize potential sources of interference, no other vehicles or moving objects were present during the course of our measurements. It is worth noting that the two-lane road under examination is situated in a sparsely populated area, with no nearby structures that could impede signal transmission. The nearest edifice, a university building, is located approximately 50 m from the "meeting point" of the passing vehicles. The measurement setup is shown in Fig. 3. We use two vehicles marked as Cars #1 and #2, as transmitting and receiving endpoints, respectively. To power the transmitter (TX) and receiver (RX) hardware, the TX is powered by a large battery pack mounted in the trunk, while the RX is powered by a generator loaded on the trailer. The RX and TX are each equipped with a pair of omnidirectional Substrate-Integrated waveguide (SIW), as expounded upon in [10], [11]. The antenna is situated in the upper left-hand corner of the windshield of the test vehicle, as viewed from the driver's perspective, and is securely fastened using a suction cup.
For the sake of reproducibility, in the following, we elaborate on the design principles and commercial components of the transmitter and receiver ends of our setup. The key parameters of the time-domain sounder can be enumerated as follows: a maximum observable time of 163.8 ns, a maximum observable distance of 49.13 m, a pseudorandom binary sequence (PRBS) bit count of 2048, a sample count of 8192, a correlation gain of 39.1 dBm, a number of saved channel impulse responses (CIRs) equaling 468, and a total measurement duration of either 2.3 or 4.6 s.

1) Transmitter design: Our transmitter is based on an Anritsu MP1800A Signal Quality Analyzer (used as Pseudo-random binary sequence (PRBS) generator), whose baseband output signal is upconverted to the mmWave band using a SiversIma FC1005V/00 V-band up/down converter equipped with a programmable phase-locked loop local oscillator. To improve phase noise performance, the oscillators are kept separate and the reference signal for up-conversion is taken from an Agilent E8257B frequency-stable, low phase noise generator. The transmitter is powered by a DELL 5600W 4U 230 V uninterruptible power supply (UPS). As a transmit signal, we use seamless repetitions of complementary Golay sequences with a data rate of 12.5 Gbit/s with a frequency-limited bandwidth of 0-4 GHz, up-converted to the mmWave band with a centre frequency of 59.6 GHz. A bandpass filter (BPF) partially suppresses the leakage of the local oscillator signal into the RF path to prevent the PA saturation. Finally, the signal with a gain of 35±3 dB is fed to the power amplifier through a coaxial cable with an attenuation of 12 dB. The output powers of the PRBS and reference signal generators are 5 dBm and 10 dBm, respectively.

2) Receiver design: The receiver comprises a Tektronix MSO72004C (20 GHz, 50 GS/s) Mixed Signal Oscilloscope, and a SiversIma FC1003V/01 (without local oscillator) for downconversion. As a carrier generator, we employed an Agilent 83752A. We used LabView for downloading and processing the raw data from the oscilloscope and setting the operational parameters. The received signal is passed through an LNA with a gain of 33± 3 dB, a noise figure of 4.5 dB, and a coaxial cable with an attenuation of 12 dB. The front-end oscillations caused by the large gain of the amplifier and the subsequent SiversIma input circuit are suppressed with a waveguide isolator and a special three-screw waveguide tuner that acts as a matching circuit. A quadrature down-converter produces two baseband signals, I and Q, which are digitized and stored in the oscilloscope's memory for use as matched receivers. Since the converter contains a 4-multiplier, the frequency of the generator is set to 14.9 GHz. The Golay sequence was used as a excitation signals for their excellent cyclic correlation properties. To compensate for large propagation losses in the mmWave band and attenuation in coaxial cables, we employed a Quinstar's QLW-50754530-I2 Low Noise Amplifier (LNA) and a QPW-50662330-C1 PA. The transmitter and receiver were synchronized by a 10 MHz reference signal generated by a Rubidium 10 MHz oscillator.

III. Datasets

In this Section, we provide a concise description of the format of the presented V2I and V2V channel measurement datasets.

A. V2I KPI Measurements

The dataset is provided as a .csv file organized in a tabular format, containing 3034 rows and 12 columns. Each row

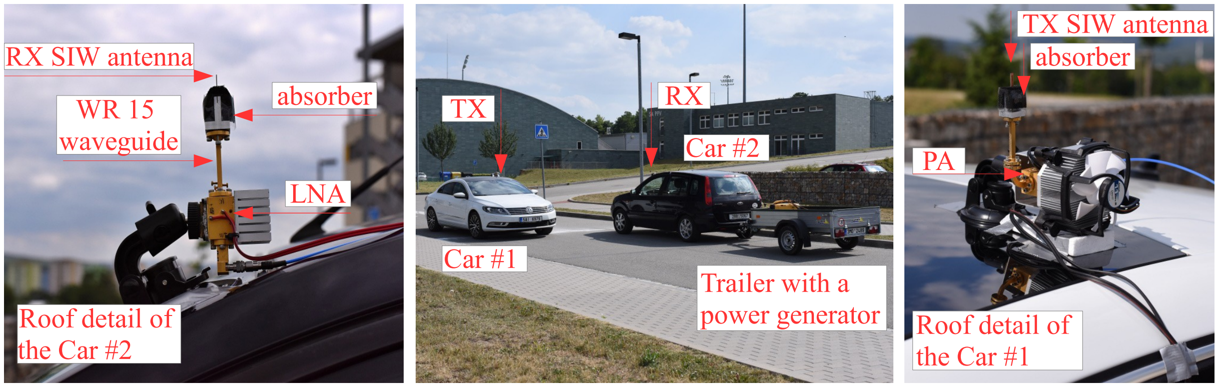


Fig. 3: Pictures of the two measurement vehicles, Car #1 with the TX and Car #2 with RX at their "meeting point". Detail photos of RX and TX, including Power Amplifier (PA) and Low Noise Amplifier (LNA) installation.

corresponds to a data sample, while each column reports a different metric, as follows:

- User Equipment (UEs) identifier (IDs)
- UEs GPS longitude
- UEs GPS latitude
- UEs GPS Speed (m/s)
- UEs GPS Time
- DL Signal to Noise Ratio (SNR) (dB)
- DL MCS
- DL Throughput (Mbps)
- UL SINR (dB)
- UL MCS
- UL Throughput (Mbps)

Since a single user terminal was used in the experiment, the UEs IDs are 0 for all observations. UEs Longitude and UEs Latitude values are expressed using a $xxyy.zzzz$ format, where xx are degrees, yy are minutes, and $zzzz \times 60$ are seconds. UEs GPS Time values are in a hh:mm:ss format. The MCS was fixed for both DL and UL channels.

As an illustrative data exploration example, Fig. 4 shows the relationship between the user throughput and UL SINR, and location. We note that the throughput distribution is heavily skewed towards the maximum value of 40 Mbps. In fact, approximately 95% of the collected samples are above 20 Mbps. Furthermore, there is a clear relationship to SINR, taking the shape of a sigmoid-like curve. This implies that, at sufficient SINR, the throughput jumps rapidly and fast reaches the maximum. Moreover, we observe that the points where the highest throughput was measured are concentrated in the South-East and North-East portions of the map.

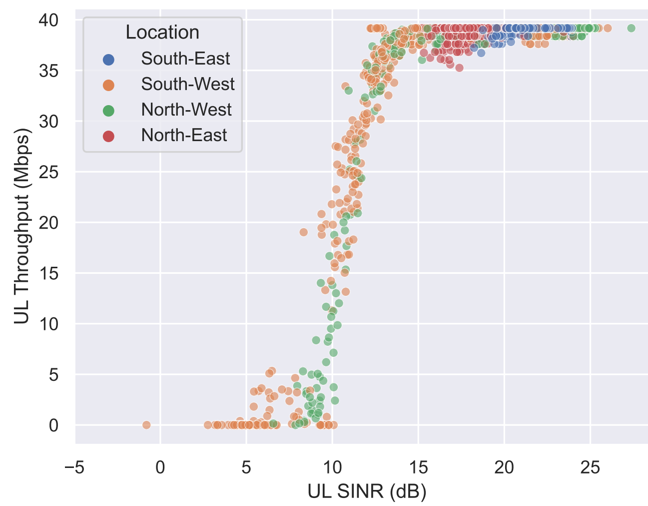


Fig. 4: UL SINR versus UL Throughput in the V2I measurements dataset.

B. V2V Channel Measurements

The measurement campaign consists of two independent runs of the vehicles as described in Section II-B. Specifically, the separate runs refer to a distinct set of CIRs captured at different relative TX-RX positions for a fixed duration of 4.6 seconds. During each measurement run, a set of 468 CIR realizations is acquired. Each experiment run lasts approximately 4.6 s, while each CIR measurement has a duration of 163 ns with 8192 samples each. In Fig. 5 we refer to the first time interval, comprising the whole length of the experiment, as "slow time" (x-axis), while the length of a single CIR measurement (y-axis) is referred to as "fast time". The dataset consists of two .mat files named `run*.mat`. We provide an illustrative MATLAB script (`V2V_CIR_show.m`) for visualizing the `run*.mat` files. We illustrate one CIR measurement in Fig. 5, which we can interpret as follows:

1) In $t \in [0, 2.4)$ s, the measurement vehicles are in motion towards each other from opposite directions. As a result, the received power increases together with the channel's delay dispersion.
2) The vehicles meet each other approximately at time $t = 2.4$ s having the minimal mutual distance of approx. 3 m. At this point, the received power and the delay dispersion are maximized.
3) In $t \in (2.4, 4.6]$ s, the measurement vehicles are driving apart from each other in opposite directions.

IV. Machine Learning Use-Cases and Research Directions

This Section discusses and illustrates potential ML use cases for the presented data. Our goal is to show that, even though the two sets of data yield very different kinds of measurements,

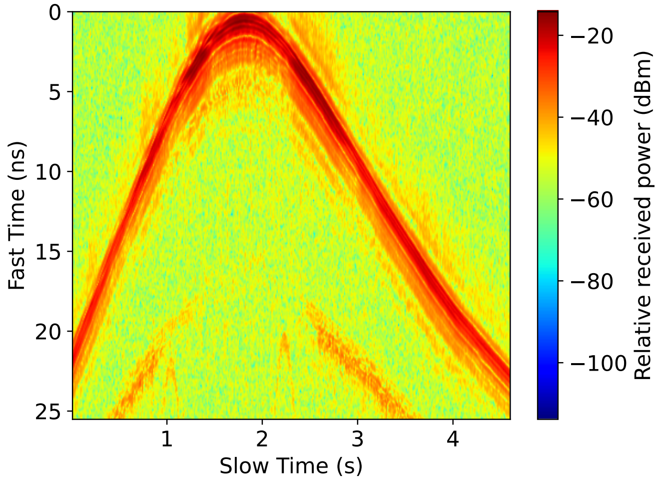


Fig. 5: Illustrative CIR of one pass of the vehicles from V2V dataset.

they can both be employed for proactive decision-making based on predictive QoS with a similar algorithmic approach.

A. V2I KPI Measurements

In [8] it was shown that the KPI measurements collected allow for satisfactory results in the task of UL throughput regression by employing classic ML methods such as linear regression, decision trees, and Deep Neural Network (DNN)s. We illustrate a novel use-case for the presented dataset, such that the data is not treated as independent KPI snapshots but as a KPI time series. In particular, we consider the task of multi-step KPI forecasting in place of instantaneous regression. This is a fundamental step towards PQoS, enabling reliable proactive decision-making in adaptive transmission systems (e.g., after forecasting an outage event tens of seconds in advance). To this end, we propose the use of approximate Bayesian ML models [12] capable of yielding probabilistic outputs, which can be interpreted as a measure of predictive uncertainty. This makes possible to leverage PQoS in risk-sensitive application scenarios.

To tackle the task of time-series forecasting with uncertainty estimation, we leverage Long-Short Term Memory (LSTM) networks with Variational Dropout [13]. In contrast with standard Dropout, Variational Dropout allows to perform approximate Bayesian inference in deep Bayesian LSTM networks. In particular, we implement our forecaster as an encoder-decoder architecture, where the encoder consists of multiple LSTM layers, and the decoder is a standard Multi-Layer Perceptron (MLP). We consider the task of forecasting future values of the UL throughput given a time series of past radio and location features. As in [8], we consider as input features the UL SINR and the position coordinates of the user equipment. Our forecasting task straightforwardly extends the work in [8]: the forecaster must learn not only a mapping between the input features and the user throughput, but its behavior over time too. For training and evaluating our model, we consider a 70/30 temporal train-test split. At test time, we perform approximate Bayesian inference via Monte-Carlo Dropout [13]. This provides not only a point-estimate prediction of the UL throughput, but also an estimation of the predictive uncertainty via the empirical standard deviation and/or confidence intervals.

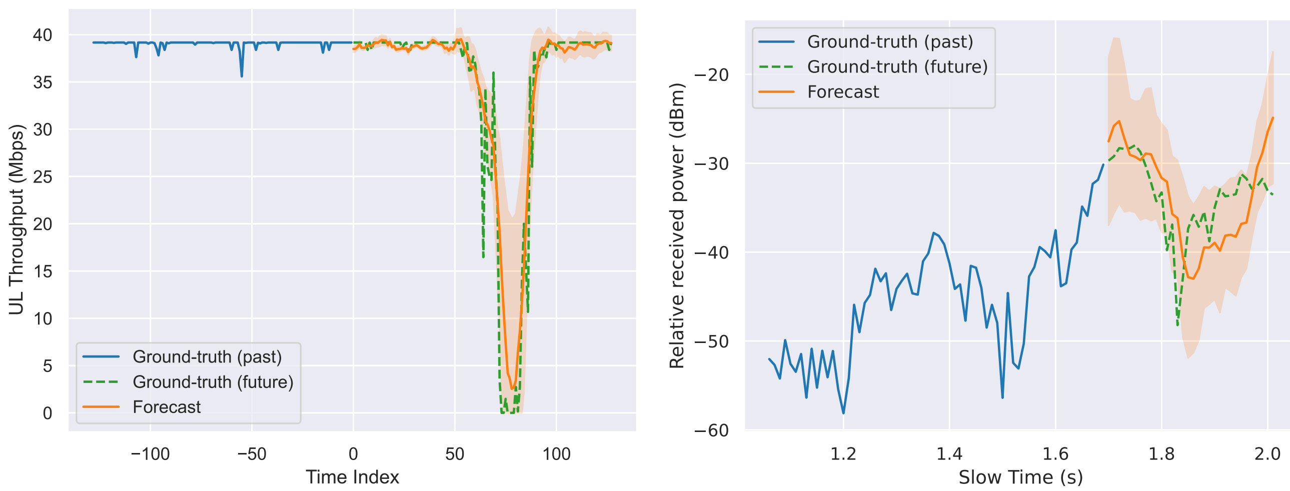
Fig. 6a, shows an illustrative forecast obtained on the prediction of the UL user throughput over 128 consecutive time steps, corresponding to a forecast prediction of 128 seconds. Our forecaster almost perfectly predicts an outage phenomenon approximately one minute in advance. As such, it is possible to act proactively to ensure that the user equipment will not experience the forecasted outage phenomenon in the future, for instance, by switching to a lower MCS or by initiating an early handover procedure. Moreover, outage predictions are useful in the context of high-density platooning to estimate when platoon members are out of coverage and ensure continuous operation even without connectivity [3].

Our approach can be easily generalized to a swarm of vehicles collaboratively collecting data over a vehicular network infrastructure. In this context, predictive uncertainty would provide indications on where (in terms of position coordinates) additional data is needed for improving the performance of the ML model [14]. In this scenario, each vehicle can share the same neural network model illustrated above, and rely on confidence bands to identify highly-uncertain regions where the available amount of training data is insufficient. This can be formally contextualized as an Active Learning framework [15], in which the action of labeling new data is associated with a cost (in our case, it would be the cost of a drive test, data collection, and data upload to a cloud server). The final goal would then be to train an accurate forecasting tool while minimizing the operational costs of data collection.

B. V2V Channel Measurements

The second measurements we discuss are intrinsically distinct from the previous dataset, as they collect CIRs in place of network KPIs. Nevertheless, using the same algorithmic framework described above, it is still possible to make use of these data in the context of PQoS. In [9], it is shown and discussed how in V2V communications the maximization of the SNR ratio is not always related to the best achievable performance in terms of Bit Error Rate (BER) and user throughput. This is because V2V communications are typically characterized by short distances, thus self-interference due to multipath components (MPCs) might lead to error floors and increase of the (uncoded) BER. In order to preserve a guaranteed level of QoS, it is therefore crucial in V2V scenarios to have proper predictive mechanisms as a support for adaptive algorithms. In particular, it is possible to represent raw CIR measurements as a collection of time-series of received power. As such, the same algorithm presented for PQoS can be leveraged for producing probabilistic CIR forecasts over time given past CIR measurements. As an illustrative application, CIR forecasts can be fed to a closed-loop power control algorithm, such as to guarantee that, e.g., the RMS delay spread never exceeds a given threshold.

We considered a forecasting task where the model is fed with the last 64 CIR samples and is tasked to predict the next 32, corresponding to a 0.32 s forecasting horizon. Moreover,



(a) V2I

(b) V2V

Fig. 6: (a) V2I UL throughput forecast from UL SINR and vehicle position. Each future prediction corresponds to 1 second in the time horizon. (b) V2V received power time-series forecast given past values of correlated received power time-series. The illustrated time-series corresponds to the top-left row of the CIR measurement shown in Fig. 5.

as CIR time-series are strongly correlated over fast time, we observed that including the six time-series ”closest in time” to the forecasted one yields a significant boost in forecasting performance. In Fig. 6b we show an illustrative forecast from our Encoder-Decoder LSTM for the earliest received power time-series, corresponding to the top row in Fig. 5. We can observe that our forecasting model not only captured correctly the temporal evolution of the time-series, but also provided accurate uncertainty estimates covering almost-perfectly the ground-truth, allowing for risk-aware decision making.

Finally, since V2V communications exhibit stronger variability compared to V2I, due to surrounding and traffic conditions, it makes sense to extend the framework from single-vehicle measurements to a swarm of vehicles. In particular, extending the presented dataset with new delay spread measurements conducted on highway scenarios would be of particular interest. Firstly, as vehicles in highways move in the same direction and the relative movement between them is less arbitrary, the typical evolution of the delay spread evolves at a slower pace. In turn, this makes the prediction task even more affordable. Secondly, given the large vehicular traffic in highways, a considerable amount of vehicles might contribute to measurements collection and sharing. Data or global ML models might be stored within roadside units.

V. Conclusion

In this work, we present two datasets comprising V2I and V2V measurements and show how these can be employed for PQoS. Besides discussing why it is important to employ both channel and KPI measurements for V2X systems, we discuss how similar predictive tasks can be performed using a similar algorithmic approach. To this end, we introduce novel experimental results and methods for predictive regression by leveraging ML techniques. In particular, we present numerical results showing that it is possible to predict an outage phenomenon approximately one minute in advance from KPI observations, and we discuss how the same methods apply to the task of CIR forecasting and adaptive power control in V2V communications. Moreover, by leveraging Bayesian Learning (BL) techniques, we articulate how it would be possible to efficiently extend the presented data from a single vehicle to a swarm of vehicles collaboratively collecting data over a vehicular network infrastructure. The datasets are available on IEEE Dataport [1] [2].

Acknowledgments

This article is based upon work from COST Action INTERACT, CA20120, supported by COST (European Cooperation in Science and Technology).

References

[1] M. Boban *et al.*, “Huawei MRC V2I measurement data,” 2022.
[2] J. Blumenstein *et al.*, “V2v mmwave channel measurements,” 2022.
[3] M. Boban *et al.*, “Predictive quality of service: The next frontier for fully autonomous systems,” *IEEE Network*, 2021.
[4] O. Simeone, “A very brief introduction to machine learning with applications to communication systems,” *IEEE Transactions on Cognitive Communications and Networking*, 2018.
[5] M.H.C. Garcia *et al.*, “A tutorial on 5g nr v2x communications,” *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1972–2026, 2021.
[6] C. Perfecto *et al.*, “Millimeter-wave v2v communications: Distributed association and beam alignment,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 2148–2162, 2017.
[7] B. Coll-Perales *et al.*, “Evaluation of ieee 802.11ad for mmwave v2v communications,” in *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2018, pp. 290–295.
[8] M. Boban *et al.*, “Measurement-based evaluation of uplink throughput prediction,” in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022.
[9] J. Blumenstein *et al.*, “Vehicle-to-vehicle millimeter-wave channel measurements at 56-64 GHz,” in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, 2019.
[10] T. Mikulasek *et al.*, “SIW slot antennas utilized for 60-GHz channel characterization,” *Microwave and Optical Technol. Lett.*, vol. 57, no. 6, pp. 1365–1370, 2015.
[11] T. Mikulasek *et al.*, “Antennas utilized for intra-vehicle 3-11 GHz and 55-65 GHz channel measurement,” in *2016 Progress in Electromagnetic Research Symposium (PIERS)*, Aug 2016.

[12] C. Blundell *et al.*, "Weight uncertainty in neural network," in *International conference on machine learning*. PMLR, 2015, pp. 1613–1622.
[13] Y. Gal *et al.*, "A theoretically grounded application of dropout in recurrent neural networks," in *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2016.
[14] L.V. Jospin *et al.*, "Hands-on bayesian neural networks—a tutorial for deep learning users," *IEEE Computational Intelligence Magazine*, vol. 17, no. 2, pp. 29–48, 2022.
[15] P. Ren *et al.*, "A survey of deep active learning," *ACM Comput. Surv.*, vol. 54, no. 9, oct 2021.

Marco Skocaj is a Ph.D. candidate within the Department of Electronic, Information, and Electrical Engineering at the University of Bologna and WiLab, CNIT. He is chair of HA1 (Datasets) working group in COST action 20120 INTERACT. His research interests include Radio Resource Management, Machine Learning, Autonomous Networks, and Optimization.

Nicola Di Cicco is a Ph.D. candidate within the Department of Electronic, Information, and Bioengineering, Politecnico di Milano, Milan, Italy. His research interests include Machine Learning for network optimization and automation.

Tommaso Zugno received his B.Sc. (2015) and M.Sc. (2018) in telecommunication engineering from the University of Padova. From May to October 2018 he was a postgraduate researcher with the Department of Information Engineering, University of Padova. His research focuses on protocols and architectures for 5G mmWave networks.

Mate Boban received his Ph.D. degree in electrical and computer engineering from Carnegie Mellon University. He is with Huawei Technologies, Munich Research Center, Germany. He is a co-chair of WG1 (Radio Channels) in COST action CA20120 INTERACT and an Associate Editor of IEEE Transactions on Mobile Computing. His current research interests include channel modeling, resource allocation, and machine learning applied to wireless communications.

Jiri Blumenstein received the Ph.D. degree from the Brno University of Technology, in 2013. In 2011, he was a Researcher with the Institute of Telecommunications, TU Wien. He is currently a Researcher with the Department of Radio Electronics, Brno University of Technology. His research interests include signal processing, physical layer of communication systems, channel characterization and modeling, and wireless system design.

Ales Prokes graduated from BUT in 1989, obtained the Ph.D. degree in 1999. In 2006 he habilitated and in 2011 he became Professor at BUT. He has been joint researcher in several grant projects concerned with problems of optical and RF communications.

Tomas Mikulasek received his master degree and doctoral degree from the Brno University of Technology in 2009 and 2013, respectively. At the present, he is with the Department of Radio Electronics, Brno University of Technology in the Czech Republic, as a researcher. His research interest is focused on design of centimeter and millimeter wave antennas.

Josef Vychodil received the master's degree from Brno University of Technology in 2013. Currently, he is pursuing his Ph.D. degree at the Department of Radio Electronics, Brno University of Technology. His Ph.D. thesis is focused on the ultra-wide band and millimeter wave band channel measurements techniques, especially dealing with the real devices imperfections.

Konstantin Mikhaylov is an Assistant Professor for Convergent IoT Communications for Vertical Systems with the CWC, UO. The focus of his research is on radio access and beyond-access technologies for massive and dependable IoT and the matters related to the design and performance of IoT devices and systems.

Massimo Tornatore is currently a Full Professor with the Department of Electronics, Information, and Bioengineering, Politecnico di Milano. His research interests include performance evaluation, optimization and design of communication networks, network virtualization, network reliability, and machine learning application for network management.

Vittorio Degli-Esposti is currently Associate Professor with the Department of Electrical Engineering of the University of Bologna and previously Director of Research at Polaris Wireless Inc., USA, in 2015-2016. He is Chair of the Propagation Working Group of the European Association on Antennas and Propagation (EurAAP) He is author or co-author of over 150 peer-reviewed technical papers and co-inventor of 7 international patents in the fields of applied electromagnetics, radio propagation and wireless systems.