



Towards Data-driven Simulation of End-to-end Network Performance Indicators

Benjamin Sliwa and Christian Wietfeld

Communication Networks Institute, TU Dortmund University, 44227 Dortmund, Germany

e-mail: {Benjamin.Sliwa, Christian.Wietfeld}@tu-dortmund.de

Abstract—Novel vehicular communication methods are mostly analyzed simulatively or analytically as real world performance tests are highly time-consuming and cost-intensive. Moreover, the high number of uncontrollable effects makes it practically impossible to reevaluate different approaches under the exact same conditions. However, as these methods massively simplify the effects of the radio environment and various cross-layer interdependencies, the results of end-to-end indicators (e.g., the resulting data rate) often differ significantly from real world measurements. In this paper, we present a data-driven approach that exploits a combination of multiple machine learning methods for modeling the end-to-end behavior of network performance indicators within vehicular networks. The proposed approach can be exploited for fast and close to reality evaluation and optimization of new methods in a controllable environment as it implicitly considers cross-layer dependencies between measurable features. Within an example case study for opportunistic vehicular data transfer, the proposed approach is validated against real world measurements and a classical system-level network simulation setup. Although the proposed method does only require a fraction of the computation time of the latter, it achieves a significantly better match with the real world evaluations.

I. INTRODUCTION

Designing and evaluating novel communication mechanisms for vehicular networks is a highly complex task as the performance of the radio link is severely impacted by mobility-related effects, the properties of the environment and interference [1]. Since comprehensive real world evaluations are very time-consuming, it is often not feasible to perform parameter studies in the real world.

Therefore, system-level network simulation is a commonly used approach, which enables to analyze various parameterizations and to evaluate novel schemes in different scenarios. In contrast to field tests, one of the main strengths of simulations is their system-immanent control of the environment, which allows to evaluate different methods under the exact same conditions. However, as the real world radio medium is shared with other cell users, for which the communication pattern is unknown, the generation of realistic simulation setups is itself a highly challenging – if not impossible – task [2]. Therefore, the results of these simulations differ significantly from real world evaluations.

The analysis of end-to-end network performance indicators (e.g., the resulting data rate) within mobile networks is even more complicated, as it is impacted by various cross-layer dependencies within the device itself, mobility-related phenomena and effects related to the coexistence of different

users and technologies in a shared radio medium (see Fig. 1) [3]. In this paper, we present a fundamentally different alternative to established analysis methods, which relies on a data-driven approach for machine learning-based modeling of end-to-end network performance indicators. In contrast to system-level network simulation, which is mainly intended for the *system conception phase*, the proposed approach targets the *system optimization phase* (see Fig. 2). Real world measurement data is exploited in an automated process to derive a *prediction model*, which is suitable for simulative evaluation as well as for online application within live systems. In addition, a second machine learning model is derived based on Gaussian Process Regression (GPR), which learns the statistical *distribution of the prediction errors*. Both models are brought together in a lightweight simulation setup – which is referred to as Data-driven Simulation (DDS) – for analyzing a novel Method Under Study (MUS).

With the proposed approach, novel communication methods can be analyzed under close to reality conditions in a fast and resource-efficient way. Overall, the proposed machine learning-based method brings together the accuracy of real world experiments, the environment control of simulation setups and the evaluation speed of analytical modeling. The contributions provided by this paper are as follows:

- An automated process for generating prediction and error

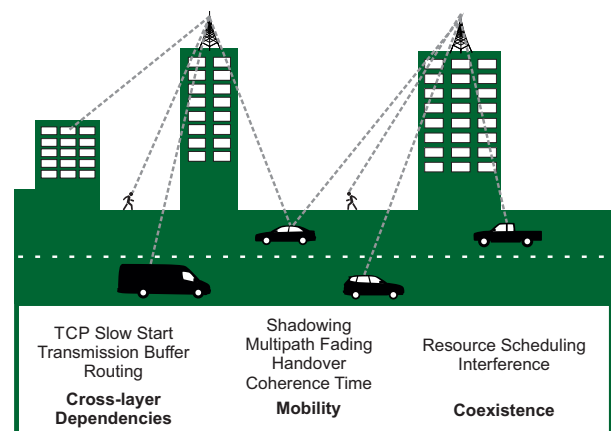


Fig. 1. Overview about different effects, which have an impact on the resulting end-to-end behavior of cellular vehicular networks. The strength of machine learning is the implicit consideration of hidden interdependencies, which are partly covered by measurable indicators.

models of end-to-end network performance indicators based on a **combination of multiple machine learning methods**.

- Proof-of-concept evaluation and **validation against system-level simulation and real world measurements** in a case study for anticipatory transmission of vehicular sensor data.
- The framework itself and the raw results of the considered case study are provided using an **Open Access** approach.

The remainder of the paper is structured as follows. After discussing relevant state-of-the-art approaches in Sec. II, we present the system model of the proposed solution approach in Sec. III. Afterwards, a case study for opportunistic data transmission is discussed, where the proposed DDS is validated based on real world measurements and compared to a purely simulative approach in Sec. IV. Finally, the limitations of DDS are discussed in Sec. V. Fig. 2

II. RELATED WORK

In recent years, machine learning has become an essential method for analyzing and optimizing communication systems and processes. Comprehensive surveys about methods and applications of machine learning within mobile wireless networks are provided by [4] and [5].

In [2], the authors present the results of a comprehensive literature review campaign with the aim of determining the most important topics, methods and trends of vehicular communication networks. One of the main findings is that many of the considered research works rely on unrealistic simulation models for both communication and mobility. Although the accuracy of simulation results can be increased by using more detailed – and more computation-intensive – simulation models [6], the required high-resolution data is often not available (e.g., material parameters for ray tracing simulation [7]). As a result, system-level network simulation is rather suitable for comparing novel methods in abstract environments than to derive conclusions for defined real world scenarios [2].

Anticipatory communication [8] is a novel wireless networking paradigm, which aims to optimize the performance of communication systems by integrating context information into decision processes. It is especially well-suited for systems

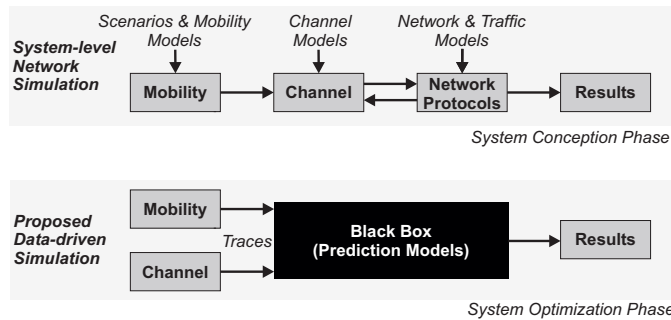


Fig. 2. Comparison of the proposed DDS approach to classical system-level network simulation.

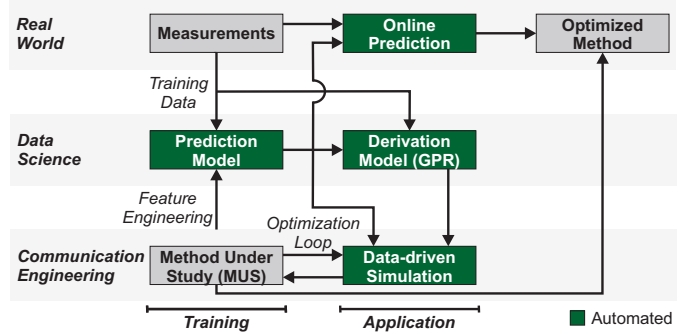


Fig. 3. System architecture model of the proposed multi-disciplinary machine learning-based analysis framework.

that experience highly dynamic radio conditions (e.g., due to the mobility behavior within vehicular networks). As anticipatory communication often aims to optimize a defined number of end-to-end performance indicators, it requires methods to describe and forecast the latter. In previous work, we have demonstrated the massive potentials of this approach for improving reliable mesh routing by mobility prediction [9] and optimizing the resource efficiency of vehicular sensor data transfer [10], [11], [12] by machine learning-based data rate prediction. One of the key strengths of these machine learning-enabled approaches is their *data driven* description of the environment. Instead of actually modeling a scenario with different interacting entities, only the interdependencies of multiple measurable variables are considered.

Within this work, we exploit the insights in data-driven modeling obtained by anticipatory communication to compensate the shortcomings of classical simulative evaluation. Recently, first results for learning end-to-end models of communication systems have been proposed by Ye et al. and [13] and Aoudia et al. [14]. While these works analyze the physical layer modeling, the proposed approach focuses on the application layer behavior.

III. MACHINE LEARNING-BASED SOLUTION APPROACH

The overall architecture of the proposed solution approach is shown in Fig. 3. The framework operates on three different logical domains: *real world*, *data science* and *communication engineering*. The overall goal is to optimize a novel MUS using DDS and bring it from theory to real world application. Although many of the components are derived by means of automation, the framework relies on two manual steps as prerequisites. In order to enable the data-driven modeling approach, a comprehensive data set needs to be obtained in the real world, which should cover all external influences on the considered Key Performance Indicator (KPI) as close as possible. Furthermore, expert knowledge from the communication engineering domain is required for the feature engineering phase, in order to monitor meaningful variables that describe the behavior of the considered end-to-end indicator. After the real world data acquisition, a DDS model is composed by the MUS, a derived prediction model (see Sec. III-A) and a derived machine learning-based model for the derivation

behavior of the prediction model itself (see Sec. III-B). Within an optimization loop, the MUS can now be analyzed and optimized based on DDS.

Finally, the optimized MUS can be transferred to a real world implementation. As the generated prediction model is represented as raw C++ code, the trained model can be directly exploited for performing online predictions in the real world deployment.

A. Generating and Analyzing Machine Learning-based Prediction Models

Network quality prediction is a regression task, which uses supervised learning based on a labeled data set \mathbf{Y} (measurements of the end-to-end network performance indicator) and a training set \mathbf{X} (measurements of related indicators). After the training phase, the trained model f can be utilized to make predictions \tilde{y} for unlabeled features sets \mathbf{x} such that $\tilde{y} = f(\mathbf{x})$. Within the proposed framework, this step is performed in an automated way, which is further described in Sec. IV-B.

As different evaluations have shown (e.g., [11], [15], [16]), for predicting network quality indicators, the application of Classification And Regression Tree (CART)-based models – e.g., Random Forest (RF) [17] and M5 Regression Tree (M5) [18] – often yields better performance than more complex regression models like Artificial Neural Network (ANN) and Support Vector Machine (SVM). In many cases, one of the considered network quality indicators has a dominant impact on the behavior of the considered KPI within a defined scope. As an example for Long Term Evolution (LTE), the analysis in [11] points out that at the cell edge – which can be identified by the Reference Signal Received Power (RSRP) – the interference level, which is partly identifiable by the Reference Signal Received Quality (RSRQ), has a strong impact on the data rate. In contrast to that, the Signal-to-interference-plus-noise Ratio (SINR) is of higher importance within the cell center. This segment-wise definition of ranges of validity matches the general structure of the CART models. In addition to their high prediction accuracy, one of the main advantages of these models is that their tree-like structure allows a resource-efficient implementation using a sequence of `if/else` statements.

It needs to be remarked that the accuracy of the resulting model is severely depending on the quality and the amount of the training data. Therefore, feature engineering is of tremendous importance in order to achieve a reusable model that works accurately within the considered evaluation scenario and provides a good generalization.

However, in contrast to system-level simulation, which requires complex evaluation setup phase, the derived model is suitable for making predictions in various scenarios after a sufficient amount of data has been obtained. For CART-based prediction models, the tree structure can be exploited to visually assess the impact of individual features. In addition, methods such as Mean Decrease Impurity (MDI) [19] allow to get a deeper insight into the importance of individual features within RFs.

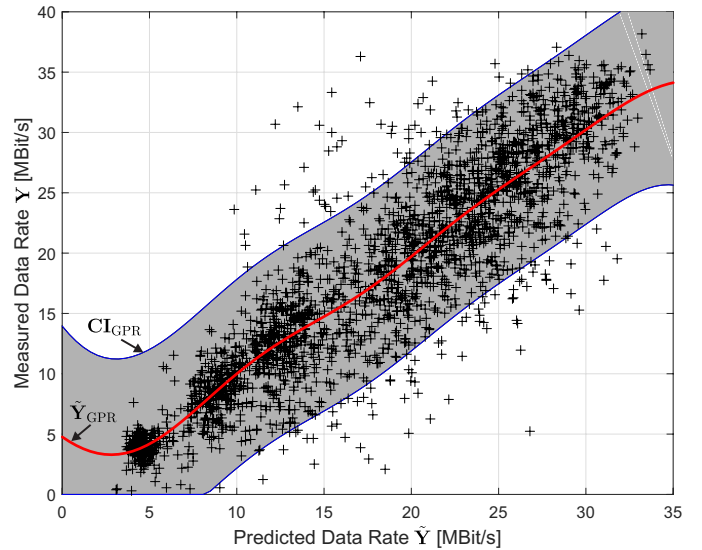


Fig. 4. For consideration of the error distribution of the trained prediction model, a GPR is trained on the prediction results of the prediction model itself. The shaded area illustrates the posterior confidence region.

For comparing the resulting prediction quality of different models, the *coefficient of determination* R^2 provides a statistical metric that describes how accurate measurement values fit to a regression model. It is calculated as

$$R^2 = 1 - \frac{\sum_{i=1}^N (\tilde{y}_i - y_i)^2}{\sum_{i=1}^N (\bar{y} - y_i)^2} \quad (1)$$

with \bar{y}_i being the mean measurement value, y_i being the current measurement and \tilde{y}_i being the current prediction.

B. Machine Learning-based Modeling of Derivations Between Prediction Model and Real World Measurements

The derived regression model can now be utilized to make predictions for unobserved values. However, for each input feature set \mathbf{x} , only a single predicted value \tilde{y} is obtained, which is constant if the same feature set is evaluated again.

Derivations within the distribution of the measured values are related to other real world influences, which are not considered by the model. However, in the targeted data-driven simulation, the results should represent the real world behavior and not just its predicted average. Therefore, a second machine learning model is applied for learning the characteristics of the derivations – which are assumed to have a gaussian distribution – between predictions and real world measurements. For this task, a GPR [20] is trained on the results of the prediction model such that $\tilde{\mathbf{Y}}_{\text{GPR}} = f_{\text{GPR}}(\tilde{\mathbf{Y}})$. A gaussian process $\mathcal{GP}(m, k)$ is a collection of random variables with a joint gaussian distribution and is described by its mean function m and its covariance function k . This method is considered *non-parametric*, as it learns a description for the data as a distribution over possible functions (in contrast to regular regression, which learns the distribution over function parameters). As illustrated by the example in Fig. 4, the GPR does not only compute predictions $\tilde{\mathbf{Y}}_{\text{GPR}}$ but also derives a confidence interval CI_{GPR} and a standard deviation function σ_{GPR} . Therefore, we can now compute a randomized sample

\tilde{y}_{GPR} , which follows a gaussian distribution \mathcal{N} of the prediction errors as

$$\tilde{y}_{\text{GPR}} = \mathcal{N}\left(\tilde{\mathbf{Y}}_{\text{GPR}}(\tilde{y}), \sigma_{\text{GPR}}^2(\tilde{y})\right) \quad (2)$$

For the data shown in Fig. 4, it can be seen that the confidence area for $\tilde{\mathbf{Y}} < 3$ MBit/s and $\tilde{\mathbf{Y}} > 33.5$ MBit/s is significantly larger than for the values in between. In these edge regions, the prediction model is not sufficiently covered by the supplied training data and therefore the resulting prediction accuracy is low. As a consequence to these statistical effects, unrealistic or even impossible values (e.g., negative data rates) can occur for \tilde{y}_{GPR} in these regions. To avoid this behavior, the results are *shaped to the value range of the measurements* in a final step with

$$\hat{y} = \begin{cases} \min(\mathbf{Y}) & \tilde{y}_{\text{GPR}} < \min(\mathbf{Y}) \\ \max(\mathbf{Y}) & \tilde{y}_{\text{GPR}} > \max(\mathbf{Y}) \\ \tilde{y}_{\text{GPR}} & \text{else} \end{cases} \quad (3)$$

For completeness, it is remarked that different machine learning models (e.g., the considered GPR) are able to derive confidence regions based on the input features themselves. However, the advantage of the presented separation of prediction model and derivation model is its modularity. As the GPR only relies on the measurements \mathbf{Y} and the prediction results $\tilde{\mathbf{Y}}$, the prediction model itself can be easily replaced by any other regression model.

IV. VALIDATION

In this section, we present a proof-of-concept evaluation of the proposed DDS approach and validate it against reference system analysis methods. As MUSs, we study the opportunistic sensor data transmission schemes Channel-aware Transmission (CAT) and the Machine Learning CAT (ML-CAT) [10], [11], which only have been analyzed empirically in previous work.

For the validation, the algorithms of the considered MUSs are implemented in three methodological setups: the proposed DDS, a classical system-level network simulator and in a real world system. The goal is to analyze how well the proposed DDS and the considered reference methods are able to represent the real world behavior of the transmission schemes.

A. Method Under Study: CAT-based Data Transfer

Straightforward periodic transmission methods within cellular vehicular networks do not consider the network quality within the transmission process itself. Therefore, cell- and energy resources are often wasted as data transmissions during low channel quality periods suffer from packet loss and lead to retransmissions.

To overcome these issues, CAT and ML-CAT [10], [11] have been proposed, which rely on a probabilistic process to compute a transmission probability $p(t)$ with respect to the current network conditions. Data is stored in a local buffer until a transmission decision is made for the whole buffer. As this approach introduces an additional delay – the transmission

scheme waits for better channel conditions – it is intended for delay-tolerant applications (e.g., intelligent traffic control, predictive maintenance, environmental monitoring) and not suitable for safety-critical real time applications. The resulting transmission probability is computed as

$$p(t) = \begin{cases} 0 & \Delta t < t_{\min} \\ 1 & \Delta t > t_{\max} \\ \left(\frac{\Phi(t) - \Phi_{\min}}{\Phi_{\max} - \Phi_{\min}}\right)^\alpha & \text{else} \end{cases} \quad (4)$$

with the exponent α for controlling the preference of high metric values and Δt being the passed time since the last transmission has been performed. t_{\min} is used to guarantee a minimum packet size and t_{\max} specifies an upper bound for the buffering delay, which has to be chosen with respect to the application requirements.

For the validation, an SINR-based metric $\Phi_{\text{SINR}} = \{0 \text{ dB}, 30 \text{ dB}\}$ and a throughput prediction-based metric $\Phi_{\text{RF}} = \{0 \text{ MBit/s}, 30 \text{ MBit/s}\}$ are applied. For comparison, straightforward periodic data transmissions with a fixed interval of $\Delta t = 10 \text{ s}$ are considered.

B. Methodological Setup

For the validation, the considered variants of CAT are analyzed using DDS and are compared to real world evaluations, which are carried out in different drive tests the public LTE network. An Android-based User Equipment (UE) (Samsung Galaxy S5 Neo, Model SM-G903F) executes the CAT process, which models a virtual sensor application. Data is transferred via Transmission Control Protocol (TCP) in the uplink from the UE through the cellular network to a cloud-based server, which measures the resulting data rate. Per configuration, five drive tests are performed on a 9 km long *suburban* track as well as on a 14 km long *highway* track. Tab. I summarizes further key parameters of the evaluation setup.

TABLE I
VALIDATION PARAMETERS

	Parameter	Value
General	Data source	50 kByte/s
	Evaluation interval	1 Hz
	t_{\min}	10 s
	t_{\max}	120 s
	α	6
	Metric for CAT	$\Phi_{\text{SINR}} = \{0 \text{ dB}, 30 \text{ dB}\}$
	Metric for ML-CAT	$\Phi_{\text{RF}} = \{0 \text{ MBit/s}, 30 \text{ MBit/s}\}$
SimuLTE	Carrier frequency	2100 MHz
	Bandwidth	20 Mhz
	UE transmission power	23 dBm
	eNB transmission power	43 dBm
	Channel model	WINNER II Urban Macrocell

The machine learning based data rate prediction metric computes its current values as $\Phi_{\text{RF}}(t) = f_{\text{RF}}(\mathbf{x}(t))$ with $\mathbf{x}(t)$ being a set of ten measured features: payload size, RSRP, RSRQ, SINR, Channel Quality Indicator (CQI), Arbitrary Strength Unit (ASU), Timing Advance (TA), carrier frequency, cell id, and velocity. The prediction model f_{RF} is a RF, which

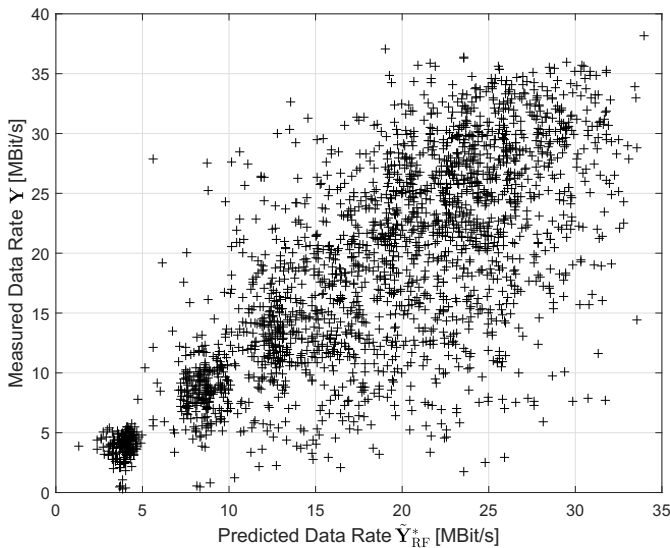


Fig. 5. Resulting data rate prediction performance for the simplified model f_{RF}^* , which is applied in the system-level network simulation.

consists of 100 random trees and is trained on the open data set of [21], which contains the results and features of 2342 uplink transmissions. On the supplied data set, the prediction model achieves a coefficient of determination of $R^2 = 0.8$ after 10-fold cross validation. For training the actual machine learning model, the Waikato Environment for Knowledge Analysis (WEKA) toolkit [22] is used. In order to automate the training phase and to generate implementations for performing online predictions, an interface application has been developed¹.

As a methodological reference, a simulation setup in Objective Modular Network Testbed in C++ (OMNeT++) 5.0 [23] with INET 3.4 and SimuLTE v0.9.1 [24] is analyzed, which replays the real world measurements. The evolved Node Bs (eNBs) are positioned with respect to their real world locations and the analyzed UE moves according to the recorded trajectory. Within the simulation setup, the prediction model f_{RF} cannot be directly implemented as the simulator only implements a fraction of the considered features. Therefore, a reduced prediction model $\hat{y}_{RF}^* = f_{RF}^*(\mathbf{x}^*)$ is derived, which utilizes the available features payload size and the SINR for the online data rate prediction. As a consequence to the model simplification, the resulting prediction accuracy is significantly reduced ($R^2 = 0.5655$), which can be observed in Fig. 5. All measurements are executed with an Intel Core i5-6200@2.3 GHz, 8 GB RAM and Ubuntu 18.04.1 LTS operating system.

C. Results

In the following paragraph, the results of different individual evaluations are presented and discussed.

1) *Modeling the Distribution of Real World Measurements:* Before the results for the CAT transmissions are analyzed, the impacts of the prediction and the GPR-based derivation model (see Sec. III) are studied. For this, all 2342 data transmissions of the training set are repeated by means of DDS and compared

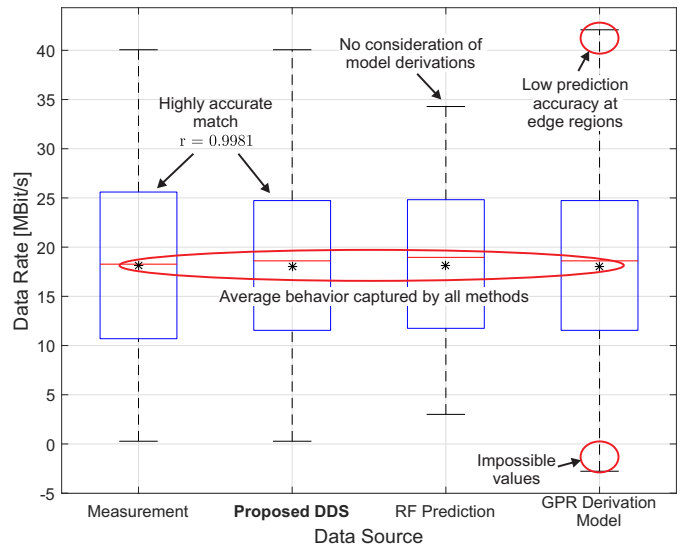


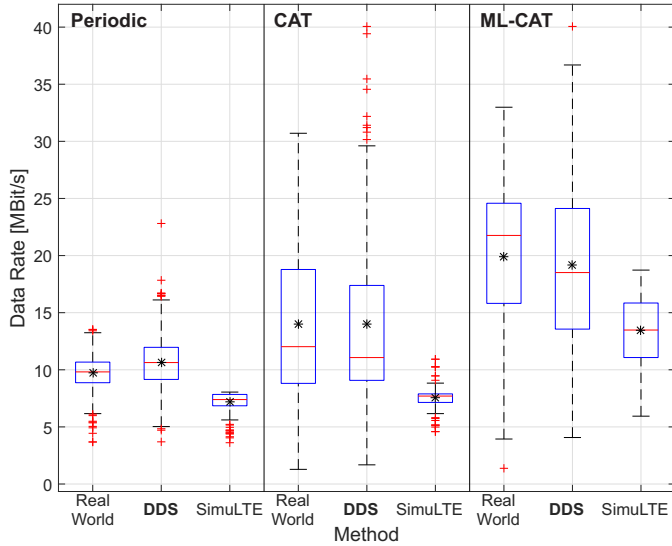
Fig. 6. Comparison of the data distribution modeling accuracy. The goal of each method is to achieve the highest congruency with the real world measurements.

to the achieved real world results. For each transmission, the recorded features are utilized by the RF-based prediction. Fig. 6 shows a comparison of the resulting accuracy for modeling the distribution of the measurement values. While all approaches are able to cover the average behavior of the observations, the resulting distributions show differences in the distribution of the data rate values. The RF prediction model does not consider model derivations, therefore the resulting value range is centered narrowly around the mean value of the measurements. In contrast to that, the raw GPR approach overemphasizes the statistical uncertainties in the edge regions. As a consequence, even impossible values occur (negative data rates)

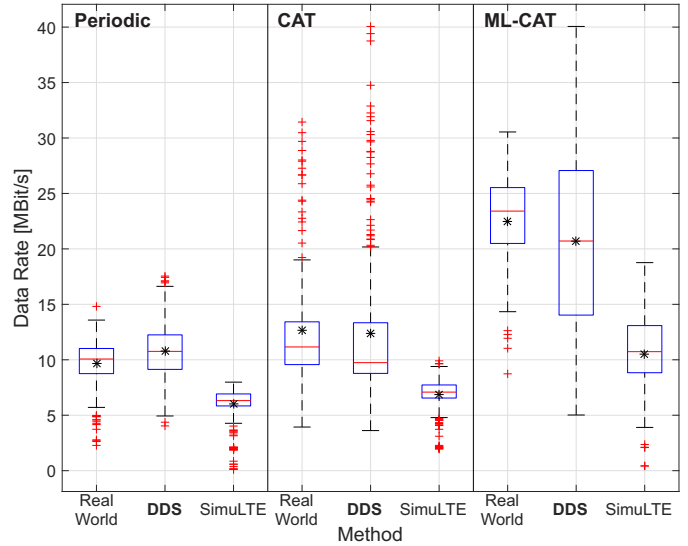
Finally, DDS brings together both components and is able to detect unrealistic derivations from the observed reality. The results show a high match with the real world measurements with a correlation coefficient of $r = 0.9981$.

2) *End-to-end Performance of the MUSs:* It is important to remark that DDS and system-level simulation utilize the data of the *training set* to replay the trajectory of the vehicle and in the case of DDS, also the channel conditions. The real world validation itself is performed against *additional measurements*, which are not part of the training set and are only used for validation. This data is captured on the same physical tracks but subject to different mobility and channel-related effects. The resulting end-to-end behavior for the considered methods and transmission schemes is shown in Fig. 7. For the real world behavior, several effects can be observed. In contrast to the periodic scheme, CAT and ML-CAT are able to achieve significant increases in the resulting data rate due to the opportunistic data transmission approach. The schemes behave differently with respect to the evaluation tracks. While the SINR-based CAT approach works better on the suburban track, ML-CAT achieves the highest benefits on the highway

¹The software is available at <https://github.com/BenSliwa/DDS>



(a) Suburban evaluation track



(b) Highway evaluation track

Fig. 7. Comparison of different performance evaluation methods for the considered MUSs.

track.

The proposed DDS is able to capture most of these phenomena very precisely. Especially on the suburban track, the results match well with the real world measurements. As discussed in previous work [10], the velocity-related network dynamics lead to a reduced prediction accuracy on the highway track, which results in an increased variance within the DDS. This effect has a doubled impact on ML-CAT, as the data rate prediction is executed twice (the first time within the transmission scheme itself and the second time within the DDS). However, the mean behavior of the data rate is still captured precisely.

In the SimuLTE evaluation, the effects of the implemented MUSs are less significant. For the suburban track, CAT only increases the mean data rate by 4 % (44 % in the real world) and ML-CAT achieves a boost by 86 % (105 % in the real world). During the real world measurements, the observed channel quality is significantly more dynamic than the channel model-based approach of the simulation. In addition, SimuLTE does not implement LTE Transmission Power Control (TPC), which explains the low variance of the results of each scenario.

It can be argued that a better congruency of simulation results and real world measurements can be achieved by increasing the modeling granularity within the simulation setup. However, in many cases the required parameters are either not easily accessible (e.g., operator internals) or impossible to obtain (e.g., the behavior of all other cell users). Therefore, the required modeling effort is massively increased for those intentions. In contrast to that, the strength of DDS is its ability to extract the parameterization implicitly from the data itself.

3) *Computation Time*: In addition to the achievable result validity, the computation time is another crucial factor for the simulative method. Often very detailed simulation environments lead to very high computation times, which limits the applicability for large-scale evaluations.

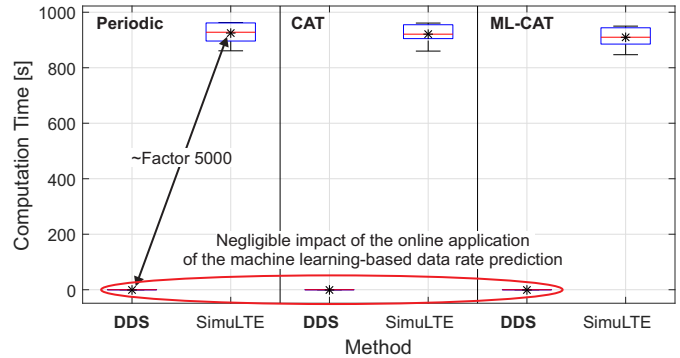


Fig. 8. Comparison of different performance evaluation methods for the considered MUSs.

Fig. 8 shows the resulting computation times per run for the proposed DDS and the system-level network simulator.

In the considered case study scenarios, the proposed DDS is more than 5000 times faster than the classical system-level simulation while being significantly more precise. Although the resulting RF model is very large – the generated RF contains 120533 leaves – it can be evaluated very resource efficiently, the computation time for online application of the trained model is negligible.

Within the case study, only the results of a single end-to-end indicator are considered. However, in contrast to the black box approach of the proposed DDS, classical network simulation spends a massive effort to derive the corresponding results. In addition to the focused UE and its protocol stack, the whole network infrastructure is simulated.

V. LIMITATIONS OF DDS

With this work, DDS is spotlighted as sophisticated alternative to classical system analysis approaches for specific evaluation tasks. Although the proposed approach has significant advantages in comparison to other scientific evaluation

methods, it has limitations that are related to its data-driven nature.

Due to the high dependency to the training set, the possible evaluations are fixed to the measured labels. In most cases, introducing additional performance indicators requires to perform additional measurements in the real world. Therefore, the exploitation of expert knowledge about the to be analyzed indicators is of tremendous importance to proactively optimize the data acquisition phase. For each considered performance indicator, dedicated prediction and error models need to be learned. In contrast to system-level simulation, the protocol stack is treated as a *black box*, therefore it is not possible to investigate intermediate indicators, which are not contained in the feature set.

VI. CONCLUSION

In this paper, we presented a data-driven approach for end-to-end network performance evaluation of novel communication methods. The proposed scheme – which is referred to as DDS – has been validated against real world measurements and system-level network simulation. DDS exploits the unique characteristics of different machine learning models, which are applied to derive a prediction model and a description for the derivations between predictions and real world measurements. The joint consideration of these models leads to a close to reality evaluation environment for the study of novel communication methods. In contrast to system-level network simulation, the scenario modeling overhead is very low as the parameterization is implicitly extracted from the data itself. Moreover, within an example case study, it was shown that DDS achieves a significantly better modeling accuracy of the real world behavior while only requiring a fraction of the computation time of regular simulation methods. In future work, we will evaluate the applicability of novel machine learning methods for both prediction and simulation model generation. Furthermore, we will utilize DDS to analyze novel anticipatory communication methods.

ACKNOWLEDGMENT

Part of the work on this paper has been supported by Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center SFB 876 “Providing Information by Resource-Constrained Analysis”, project B4.

REFERENCES

- [1] B. Sliwa, T. Liebig, T. Vranken, M. Schreckenberg, and C. Wietfeld, “System-of-systems modeling, analysis and optimization of hybrid vehicular traffic,” in *2019 Annual IEEE International Systems Conference (SysCon)*, Orlando, Florida, USA, Apr 2019.
- [2] E. R. Cavalcanti, J. A. R. de Souza, M. A. Spohn, R. C. d. M. Gomes, and A. F. B. F. d. Costa, “VANETS’ research over the past decade: Overview, credibility, and trends,” *SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 2, pp. 31–39, May 2018.
- [3] B. Sliwa and C. Wietfeld, “Empirical analysis of client-based network quality prediction in vehicular multi-MNO networks,” in *2019 IEEE 90th Vehicular Technology Conference (VTC-Fall)*, Honolulu, Hawaii, USA, Sep 2019.
- [4] 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, June 2018.
- [5] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen, and L. Hanzo, “Machine learning paradigms for next-generation wireless networks,” *IEEE Wireless Communications*, vol. 24, no. 2, pp. 98–105, April 2017.
- [6] Z. H. Mir, “Large-scale simulation of site-specific propagation model: Defining reference scenarios and performance evaluation,” in *Advances in Internet, Data and Web Technologies*, L. Barolli, F. Khafa, Z. A. Khan, and H. Odhabi, Eds. Cham: Springer International Publishing, 2019, pp. 492–503.
- [7] Z. Yun and M. F. Iskander, “Ray tracing for radio propagation modeling: Principles and applications,” *IEEE Access*, vol. 3, pp. 1089–1100, 2015.
- [8] 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*, 2017.
- [9] B. Sliwa, D. Behnke, C. Ide, and C. Wietfeld, “B.A.T.Mobile: Leveraging mobility control knowledge for efficient routing in mobile robotic networks,” in *IEEE GLOBECOM 2016 Workshop on Wireless Networking, Control and Positioning of Unmanned Autonomous Vehicles (Wi-UAV)*, Washington D.C., USA, Dec 2016.
- [10] B. Sliwa, T. Liebig, R. Falkenberg, J. Pillmann, and C. Wietfeld, “Machine learning based context-predictive car-to-cloud communication using multi-layer connectivity maps for upcoming 5G networks,” in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, USA, Aug 2018.
- [11] —, “Efficient machine-type communication using multi-metric context-awareness for cars used as mobile sensors in upcoming 5G networks,” in *2018 IEEE 87th Vehicular Technology Conference (VTC-Spring)*, Porto, Portugal, Jun 2018, Best Student Paper Award.
- [12] B. Sliwa, R. Falkenberg, T. Liebig, N. Piatkowski, and C. Wietfeld, “Boosting vehicle-to-cloud communication by machine learning-enabled context prediction,” *IEEE Transactions on Intelligent Transportation Systems*, 2019, Accepted for publication.
- [13] H. Ye, G. Y. Li, and B. Juang, “Power of deep learning for channel estimation and signal detection in OFDM systems,” *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114–117, Feb 2018.
- [14] F. A. Aoudia and J. Hoydis, “End-to-end learning of communications systems without a channel model,” in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, Oct 2018, pp. 298–303.
- [15] A. Samba, Y. Busnel, A. Blanc, P. Dooze, and G. Simon, “Instantaneous throughput prediction in cellular networks: Which information is needed?” in *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, May 2017, pp. 624–627.
- [16] F. Jomrich, A. Herzberger, T. Meuser, B. Richerzhagen, R. Steinmetz, and C. Wille, “Cellular bandwidth prediction for highly automated driving - Evaluation of machine learning approaches based on real-world data,” in *Proceedings of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems 2018*, no. 4. SCITEPRESS, Mar 2018, pp. 121–131.
- [17] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [18] J. R. Quinlan, “Learning with continuous classes.” World Scientific, 1992, pp. 343–348.
- [19] G. Louppe, L. Wehenkel, A. Suter, and P. Geurts, “Understanding variable importances in forests of randomized trees,” in *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1*, ser. NIPS’13. USA: Curran Associates Inc., 2013, pp. 431–439.
- [20] C. E. Rasmussen, *Gaussian Processes in Machine Learning*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 63–71.
- [21] B. Sliwa, “Raw experimental cellular network quality data,” Februar 2019. [Online]. Available: <http://doi.org/10.5281/zenodo.2553832>
- [22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The WEKA data mining software: An update,” *SIGKDD Explorations*, vol. 11, no. 1, pp. 10–18, 2009.
- [23] A. Varga and R. Hornig, “An overview of the OMNeT++ simulation environment,” in *Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops*, ser. Simutools ’08. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, pp. 60:1–60:10.
- [24] A. Virdis, G. Stea, and G. Nardini, *Simulating LTE/LTE-Advanced networks with SimuLTE*. Cham: Springer International Publishing, 2015, pp. 83–105.