



# System-of-Systems Modeling, Analysis and Optimization of Hybrid Vehicular Traffic

Benjamin Sliwa<sup>1</sup>, Thomas Liebig<sup>2</sup>, Tim Vranken<sup>3</sup>, Michael Schreckenberg<sup>3</sup> and Christian Wietfeld<sup>1</sup>

<sup>1</sup>Communication Networks Institute, <sup>2</sup>Department of Computer Science VIII

TU Dortmund University, 44227 Dortmund, Germany

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

<sup>3</sup>Physics of Transport and Traffic, University Duisburg-Essen, Germany

e-mail: {Vranken, Schreckenberg}@ptt.uni-due.de

**Abstract**—While the development of fully autonomous vehicles is one of the major research fields in the Intelligent Transportation Systems (ITSs) domain, the upcoming long-term transition period – the hybrid vehicular traffic – is often neglected. However, within the next decades, automotive systems with heterogeneous autonomy levels will share the same road network, resulting in new problems for traffic management systems and communication network infrastructure providers. In this paper, we identify key challenges of the upcoming hybrid traffic scenario and present a system-of-systems model, which brings together approaches and methods from traffic modeling, data science, and communication engineering in order to allow data-driven traffic flow optimization. The proposed model consists of data acquisition, data transfer, data analysis, and data exploitation and exploits real world sensor data as well as simulative optimization methods. Based on the results of multiple case studies, which focus on individual challenges (e.g., resource-efficient data transfer and dynamic routing of vehicles), we point out approaches for using the existing infrastructure with a higher grade of efficiency.

## I. INTRODUCTION AND RELATED WORK

Autonomous vehicles are often envisioned as the savior that solves all issues [1] of current traffic systems. They are expected to be able to drive with minimum safety distances even at high speeds, which enables massive increases of the usable road capacity. They do not dawdle, which results in a better traffic flow and they coordinate their actions for an improved overall safety.

While this sophisticated vision is supported by many current research works [2], most of them neglect the more immediate challenges arising from the transition period from legacy traffic to highly automated vehicular traffic that will take several decades. Due to the step-wise development and deployment of automated vehicles, systems with different levels of autonomy will coexist within the same road network in the *hybrid vehicular traffic* [3]. In this complex scenario, the following challenges for traffic flow optimization arise:

- **Novel traffic phenomena**s as automated vehicles are expected to react defensively on human misbehavior, potentially resulting in traffic flow reductions, while the overall traffic congestion grows, e.g., due to an increased amount of empty runnings and on-demand traffic.
- **Heterogeneous communication technologies** limit the

capabilities for mutual coordination and data provisioning.

- **Resource competition** as the radio spectrum is shared between the different technologies and human cell users as well as automated systems.
- Traffic management systems have to work with **incomplete information**, since established sensors systems only provide local coverage and often break during their operational time.

Existing approaches for modeling these traffic systems have not yet reached the required complexity that allows the derivation of novel insights for hybrid traffic flow optimization. In [4], the authors propose an approach to model different automation levels depending on their resulting safety distance. Another approach [5] proposes an extension of the well-established Intelligent Driver Model (IDM) to consider cooperative behaviors for lane-merging. While forecast and optimization of vehicular traffic flow has traditionally been isolatedly addressed by the traffic physics community, the arising challenges of the hybrid traffic scenario require the joint consideration of communication systems and data analysis methods [6]. In this paper, we present an overall system-of-systems model that brings together the three mentioned disciplines for modeling and analyzing hybrid vehicular traffic with the goal of data-driven traffic flow optimization. The

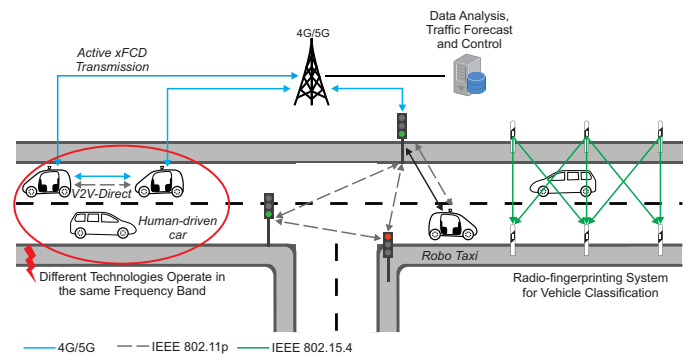


Fig. 1. Hybrid vehicular traffic: Coexistence of different vehicle types and communication technologies.

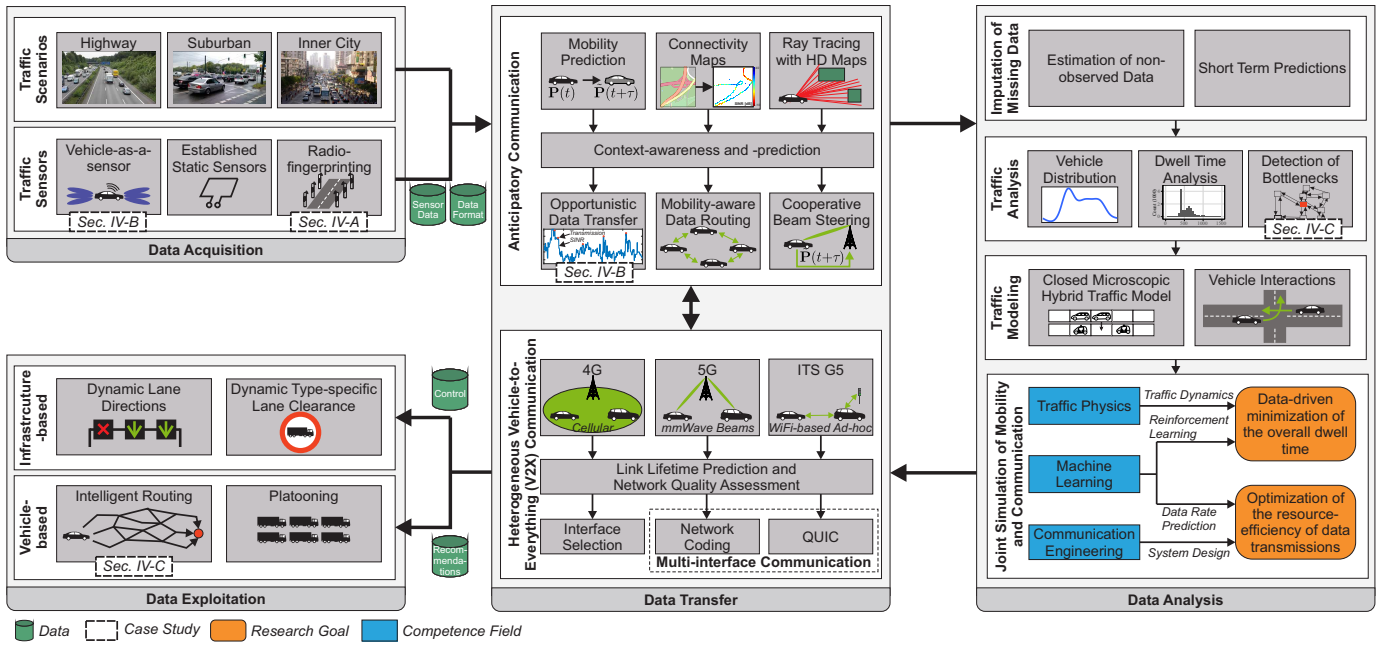


Fig. 2. Overall system architecture model for data-driven optimization of hybrid traffic consisting of data acquisition, data transfer, data analysis and data optimization. Real world sensor data is exploited to generate new traffic models, which are evaluated and optimized using a simulation setup. The derived insights are then utilized for traffic flow optimization in real world traffic management systems.

concepts are substantiated by preliminary results of different case studies that focus on individual aspects such as resource-efficient data transfer and traffic bottleneck detection. The proposed approach leverages ubiquitous machine learning to derive novel insights within all of its main dimensions, which are defined as data acquisition, data transfer, data analysis and data exploitation.

Fig. 1 shows an example scenario and illustrates the different kinds of heterogeneity within hybrid traffic scenarios. While highly automated vehicles (so-called *robo taxis*) are exploited as moving sensor nodes that provide sensor measurements as extended Floating Car Data (xFCD) via different kinds of Vehicle-to-everything (V2X) connections on their own, legacy vehicles are detected and classified based on roadside infrastructure systems.

The structure of the paper is as follows. After presenting the challenges and discussing relevant state-of-the-art research, we present the architecture model of the proposed overall system. The different submodules are explained in details and set into relation to the overall context and existing approaches. Finally, multiple case-studies that provide groundwork for the different components of the overall system are presented and selected results are discussed.

## II. SYSTEM MODEL ARCHITECTURE

In this section, the system architecture and the features of the individual components are explained. An illustration of the overall system architecture model is shown in Fig. II. For the proposed data-driven approach, traffic data is acquired in the real world, transferred to the analysis centers, utilized

for model generation and simulative optimization and finally exploited for improving the real world traffic flow. In the overall scenario, communication systems and machine learning-based data analysis are of great importance. On the one hand, communication is used as a mediator between data acquisition, data analysis and data exploitation, on the other hand, it is subject to machine learning-based optimization itself (cf. Sec. III-B). The following paragraphs provide further information about exemplary groundwork for the key system stages *data acquisition*, *data transfer*, *data analysis* and *data exploitation*. In addition, groundwork that addresses the related challenges is introduced and discussed.

### A. Data Acquisition

The derivation of new models and the following traffic flow optimization requires real world traffic data as a solid foundation. Therefore, the accurate acquisition of traffic data is of tremendous importance for the overall system. Depending on the considered road network type, the degrees of freedom for the vehicular mobility and the resulting challenges for the traffic flow analysis can be significantly different. Consequently, traffic is analyzed isolatedly for highway, suburban and inner city scenarios as well as in an overall model, which considers all traffic types. For the acquisition of actual sensor data, a combination of different real world sensor systems is applied, which includes existing traditional sensors as well as experimental approaches that offer novel acquisition methods and are capable of opening up new data categories. Traditional sensor systems (e.g., induction loops) are intended to acquire traffic parameters such as traffic flow and density. Due to

the required road work, sensor deployments are cost-intensive, which results in limitations for the number of deployments and the resulting sensor coverage. However, since many existing traffic monitoring systems rely on these approaches, they are integrated into the system model in order to benefit from the existing infrastructure. In addition to these established sensor systems, novel methods allow the acquisition of additional traffic parameters (e.g., vehicle type distributions), which enables new approaches for traffic flow optimization.

1) *Vehicle-as-a-sensor*: Vehicular crowdsensing [7] marks a paradigm shift from static sensor installations to exploiting the vehicles themselves as moving sensor nodes, resulting in massive improvements for coverage and data freshness. Using a multitude of specialized sensors, vehicles can be utilized to provide meaningful traffic parameters as well as environmental information, e.g., to detect traffic obstructions or free parking slots. Apart from the considered traffic management scenario, this approach has been demonstrated to significantly improve distributed weather prognosis [2], road condition sensing [8] and inner city air quality monitoring.

Vehicular crowdsensing is closely connected to the *data transfer* topic, as it relies on connectivity and the presence of a data uplink. Moreover, large-scale deployments of crowdsensing services lead to massive increases in the amount of transmitted data, which imply novel challenges on the network infrastructure side (see Sec. III-B).

2) *Novel Roadside Sensors*: In order to enable promising traffic optimization methods like type-specific lane clearance, traffic management systems need to obtain information about the distributions of vehicle types. A multitude of different sensor types for vehicle classification has been proposed, which ranges from vision-based approaches using cameras, radars and lasers to inertial measuring units. A comprehensive overview about existing sensor technologies is presented in [9]. However, none of the existing technologies is able to fulfill all requirements for accuracy, privacy preservation, weather robustness while still being cost-efficiently enough to allow citywide large-scale deployments.

In Sec. III-A, we present a case study on radio-fingerprinting-based vehicles classification. The proposed system can be deployed cost-efficiently in an ad-hoc manner and is able to provide a high classification accuracy even at challenging environmental conditions.

## B. Data Transfer

Alongside with the coexistence of the different autonomy levels comes the coexistence of different communication technologies, that compete over the available resources in the shared radio medium. Within the proposed system for hybrid vehicular traffic analysis, data needs to be exchanged from the cars to the cloud-based data analytics center and between individual vehicles for mutual coordination. The challenges arising from highly-dynamic network topologies, heterogeneous communication technologies and challenging environments are addressed by the combined usage of two

main methodological approaches: *Anticipatory communication* and *heterogeneous V2X*.

1) *Anticipatory Communication*: The anticipatory networking paradigm [10] aims to integrate the communication context into the decision processes within communication systems themselves. Instead of extending the network capacity physically on the infrastructure side, the existing network itself is used in a more efficient way through a context-aware medium access. In the considered vehicular scenario, the dynamics of the communication medium are closely connected to the vehicles mobility behavior. Therefore, for forecasting the network quality a car is going to encounter in the future, trajectory prediction is a basic requirement [11].

For improving the coexistence of different resource-consuming entities, the resource-efficiency of their individual data transmissions needs to be optimized. Resource-efficient communication has several dimensions: Data transmissions need to be short (high data rates) in order to free the occupied resources as soon as possible, they have to be reliable (low amount of retransmissions) and they have to satisfy the data freshness requirements of the intended application. Although different indicators exist for assessing the current network quality of a given communication technology, the validity of the indicators is related to the channel coherence time and the transmission patterns of the application itself. Machine learning has been demonstrated [12] to be able to efficiently consider hidden interdependencies between those aspects, resulting in an improved context-awareness. In Sec. III-B, we present a case study for resource-efficient sensor data transmission based on machine learning and mobility prediction.

In addition to optimizing the resource-efficiency of data transmissions, anticipatory communication can serve as a method to establish robust communication paths within vehicular mesh networks. Here, the consideration of the mobility characteristics of the vehicles allows to derive link-lifetime estimations that enables the predictive optimization of multi-hop routes [12]. With the deployments of 5G communication networks and the millimeter Wave (mmWave) technology, context-aware networking becomes a native aspect of mobile networking, since the extremely high directed radio beams need to be precisely steered towards their moving receiver.

2) *Heterogeneous V2X*: As it is expected that vehicles will make use of a combination of different technologies for specific use-cases [13] – e.g., low-latency ad-hoc communication for safety-related messaging and high data rate cellular communication for sensor data transfer – the unique technological properties can be exploited for dynamic technology selection. Moreover, for each data transfer, the currently used communication technology is chosen in a context-aware manner, e.g., available Road Side Units (RSUs) are opportunistically exploited as cost-efficient internet gateways. Anticipatory communication methods are applied to choose the respective technology in a predictive manner with respect to the application requirements, e.g., through prediction of link availability times. In addition, multi-interface communication using network coding [14] and Quick UDP Internet Connec-

tions (QUIC) is applied in order to utilize multiple links in parallel.

### C. Data Analysis

In the context of the proposed overall system, data analysis is utilized for simulative traffic flow optimization as well as for the generation of hybrid traffic models themselves. Before the required traffic properties are analyzed, the acquired sensor data is utilized to derive estimations about the traffic situation at non-observed measurement locations in order to obtain an improved awareness of the overall system dynamics.

1) *Imputation of Missing Data:* Traffic volume estimation is a fundamental task in macroscopic street-based traffic analysis systems and has important applications, e.g., quality-of-service evaluation, location evaluation or risk analysis. Nowadays, intelligent transportation systems rely on stationary sensors, which provide traffic volume measurements at predefined locations. However, imputation of the unobserved traffic flow values and short-term predictions are highly important research topics.

Existing literature distinguishes between average daily traffic (ADT) estimation and average annual daily traffic flow (AADT or AADF) estimation. Whereas AADF focuses on estimation of a traffic volume depending on the day of the year, ADT estimation provides an average for a particular day. In contrast to the ADT and AADF problem, also short-term traffic prognosis problems exist. In AADT, we concentrate on more extended temporal resolutions, where microscopic influences, e.g., signals have no impact on the traffic flow. For short-term prediction problems, we point the interested reader on related works, e.g., Cellular Automaton [15], Poisson Dependency Networks [16], or Convolutional Neural Networks [17].

A naive approach for AADF estimation is the utilization of ordinary linear regression (OLR) [18]. Street segment attributes (e.g. the number of lanes or function classes) are multiplied by weights, which are subject for least squares regression. Improvements of this technique were achieved by respecting the geographical space by usage of geographically weighted regression (GWR) [18] and by application of k-nearest neighbor approaches (kNN) [19]. In [20] the AADF prediction of kNN for a particular location is improved by weighting measurements by their temporal distance to the prediction time. This approach showed better results than the application of Gaussian maximum likelihood (GML) approaches for weighting of the historical data points.

Being a spatial regression problem, usage of the Kriging method seems to be a natural choice to tackle the AADF problem. This was successfully carried on at University of Texas [21], [22]. To machine learning persons the Kriging method is better known as Gaussian Process Regression, which allows a better understanding of the underlying spatial correlation model by reformulation with a kernel matrix. Application of Gaussian Process Regression is an appealing state-of-the-art method that outperforms recent methods [23]. The method bases on a covariance matrix that denotes the correlations among the traffic flux values at various locations. The work

in [23] tested various covariance matrices among them some that incorporate spatial layout of the sensor locations or even the topology of the street network. However, the performance did not change much for these different correlation models. However, due to the computational complexity of Gaussian Process Regression, application to urban areas was restricted either to small sites or a sample of locations [24].

2) *Traffic Modeling and Analysis:* Although a wide range of different macro- and microscopical traffic models exist, a closed model for hybrid vehicular traffic is still missing. Therefore, the insights of the data acquisition and the data transfer stages shall be brought together to extend existing cellular automaton models [25] to consider different levels of automation.

3) *Integrated Joint Simulation of Mobility and Communication:* As pointed out in the previous paragraphs, modeling hybrid vehicular traffic requires the consideration of traffic modeling, data analysis and communication. For the simulative evaluation and optimization, a common simulation setup that is able to consider all of the individual aspects needs to be created. Current approaches for simulating vehicular networks often rely on coupling specialized simulation frameworks based on an interprocess-communication approach [26]. As a result, the access to information of a specific domain (e.g., using mobility information from the traffic simulator for decisions in the network simulator) is complicated, as specified interfaces and communication protocols are used for the data exchange. In contrast to that, the proposed system model relies on an integrated approach based on Lightweight ICT-centric Mobility Simulation (LIMoSim) [27], which makes use of a shared codebase approach for all different modules.

### D. Data Exploitation

Although a naive approach to increase the road capacity is to construct additional road infrastructure, it is highly cost-intensive and often not even applicable due to space limitation, especially in inner city scenarios. Therefore, a more preferable way is to utilize the existing road network in a more efficient way. In the proposed overall system, this approach is addressed by *infrastructure- and vehicle based methods*, which exploit the results of the previous data analysis stage.

1) *Infrastructure-based Methods:* Infrastructure-based methods are centralized approaches, therefore the traffic management system is able to send control commands directly to the corresponding road infrastructure, which then executes the intended action. Assigning the directionality of lanes with respect to the capacity requirements provides a method for better coping with temporal-limited phenomena (e.g., commuter traffic) of traffic systems. Another optimization approach is to use lanes exclusively for certain vehicle types. Based on the obtained knowledge about the distribution of vehicle types, lanes can be dynamically assigned, e.g., for dividing passenger cars and heavy vehicles. Moreover, the same approach can be applied for dividing non-automated and highly-automated vehicles, which reduces the need for cross-automation-level coordination and enables traffic

systems to benefit more from the advantages of automated vehicles.

2) *Vehicle-based Methods*: Vehicle-based methods are decentralized approaches, therefore traffic management systems are only able to send recommendations to the vehicles, which then may react accordingly or not. Moreover, in the considered hybrid traffic scenario not all vehicles are equipped with communication technology.

Autonomous cars can also leverage the vast amount of data in our scenario resulting from sensors, imputations and predictions and apply these data for short-term navigation decisions, i.e., which direction to chose at a particular junction for not to generate or participate in a traffic congestion. This routing optimization approach was tested with bandit learning in [28] and achieved increased traffic network performance. In [29], multiple trip planing methods, which aim to better distribute the traffic flow over the system are investigated. The case study in Sec. III-C provides a methodological summary of the mentioned paper.

Another promising method for traffic flow optimization, which is especially interesting for freight traffic, is the usage of platooning, where vehicles coordinate their velocity characteristics to achieve minimal safety distances. Additional benefits are expected for combining platooning methods with exclusive lanes for platoons.

### III. GROUNDWORK AND CASE STUDIES

In this section, we present multiple case studies that provide groundwork for isolated research fields in the considered overall system model.

#### A. Data Acquisition: Radio-fingerprinting-based Vehicle Classification

In this case study, we present example results of our research work on radio-fingerprinting-based vehicle classification. The integration of class-specific traffic analysis allows to utilize promising traffic flow optimization techniques such as dynamic type-specific lane clearance. The proposed system aims to provide a highly-accurate low-cost alternative to existing sensor systems. Within the considered hybrid traffic scenario, the system closes the gap between connected vehicles, which are able to provide class information on their own, and non-connected vehicles, which have to be detected using roadside infrastructure systems. In [30], we have proposed initial work for a radio-fingerprinting system, which leverages the idea of *radio tomography* for vehicle classification and can be cost-efficiently deployed in an ad-hoc manner without requiring additional roadwork. The proposed system consists of low-cost communicating sensor nodes that are installed into delineator posts. Three transmitters continuously send data to three receivers on the other road side, which monitor the Received Signal Strength Indicator (RSSI) of all links. Vehicles passing the sensor installation attenuate the signals, whereas the attenuation level and the duration of the attenuation phase is related to the vehicle shape. Therefore, the time series values of all nine links are utilized as a vehicle class-specific

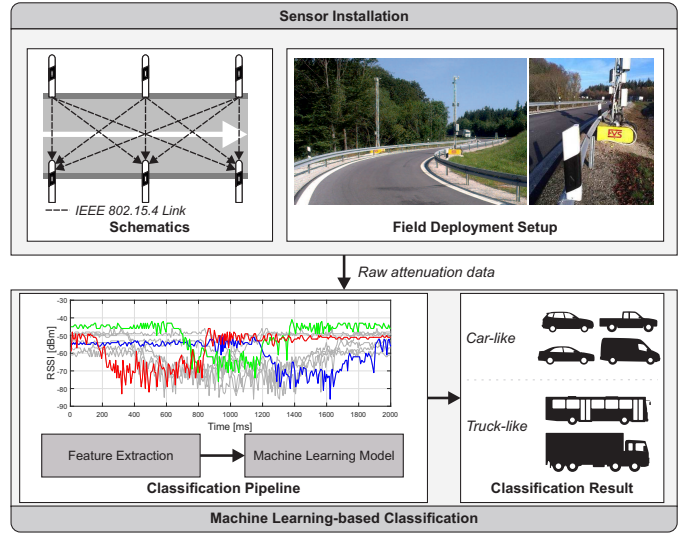


Fig. 3. System model for radio-fingerprinting-based vehicle classification.

C	0.98	0.01
	0.02	0.99
	C	T

**Raw Data**

C	0.96	0.01
	0.04	0.99
	C	T

**Reduced Feature Set**

Fig. 4. Confusion matrices for L1/L2 SVM-based binary vehicle classification. (C: Car-like, T: Truck Like)

*radio fingerprint*. The attenuation pattern itself is utilized for machine learning-based vehicle classification. Fig. 3 shows the system model for the radio-fingerprinting-based vehicle classification system. For the data acquisition, an experimental deployment of the proposed system at the entrance of a rest area on the German highway A9 within an official test field by the German Federal Ministry of Transport and Digital Infrastructure is used. During the training phase, the measured nine-dimensional time series traces are labeled manually based on camera images. The classification can be performed directly on the raw values as well as on a reduced feature set, which uses analytical properties of the different signals. In [30], a detailed comparison of different machine learning models is provided. Fig. 4 shows the empirical results for L1/L2 SVM-based approach for a data set consisting of more than 2500 real world measurements. For the binary classification task between *car-like* and *truck-like* vehicles, a total classification accuracy of 98.2 % is achieved.

With respect to the context of the overall hybrid traffic system, it is assumed that sensor sites exploit their obtained data locally, e.g., for parking space accounting, and contribute themselves to the global traffic knowledge data base through a network connection.



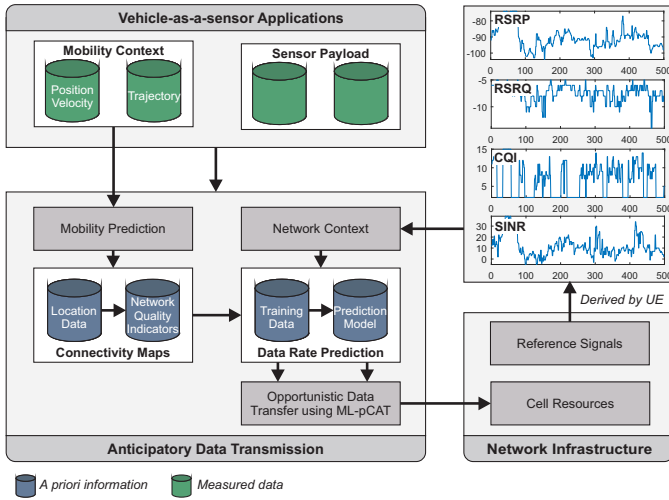


Fig. 5. System model for machine learning-based context-predictive data transmission with ML-pCAT.

### B. Data Transfer: Machine Learning-based Opportunistic Transmission

In this case study, we present groundwork for resource-efficient transmission of vehicular sensor data for optimizing the coexistence of different resource-consuming cell users. The general approach is to transmit the acquired data in a context-aware manner by exploiting knowledge about the properties of the radio channel. During good channel periods, packet loss is less probable and the resulting data rate is high, therefore the limited resources are only occupied for short time periods. For the opportunistic transmission, acquired sensor data is stored in a local buffer until the expected transmission efficiency is considered sufficient for the whole buffered data.

Based on the transmission schemes CAT and pCAT [31], which perform data transmissions with respect to the measured Signal-to-noise-plus-interference Ratio (SINR), we developed Machine Learning-based CAT (ML-CAT) [32] and ML-pCAT [11], which use machine learning-based data rate prediction as a metric for opportunistic data transfer. The prediction leverages the Long Term Evolution (LTE) downlink indicators Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), SINR, Channel Quality Indicator (CQI) as well as mobility parameters and payload size to forecast the currently achievable data rate  $\Phi(t)$ . For the opportunistic data transmission, ML-CAT uses a probabilistic process that computes the transmission probability  $p(t)$  based on the currently predicted data rate and its value range defined by  $\Phi_{min}$  and  $\Phi_{max}$ . The weighting exponent  $\alpha$  is used to define the preference of high metric values.

$$p(t) = \left( \frac{\Phi(t) - \Phi_{min}}{\Phi_{max} - \Phi_{min}} \right)^\alpha \quad (1)$$

With ML-pCAT [11], ML-CAT is extended to integrate to anticipated future network quality states into the transmission process. Mobility prediction is applied to forecast the future vehicle position and a-priori connectivity map information is exploited to estimate the future network quality based on crowdsensing data.

Fig. III-B show the resulting mean data rate and power consumption of ML-CAT and ML-pCAT in a real world evaluation campaign using the public cellular network. Data transmissions are performed from a moving vehicle to a cloud-based server, whereas the overall driven distance is more than 2000 km. Further details about the setup and the parametrization are provided in [32] and [11]. The proposed ML-CAT scheme is compared to naive periodic data transfer (transmission interval 30 s) and the SINR-based CAT and

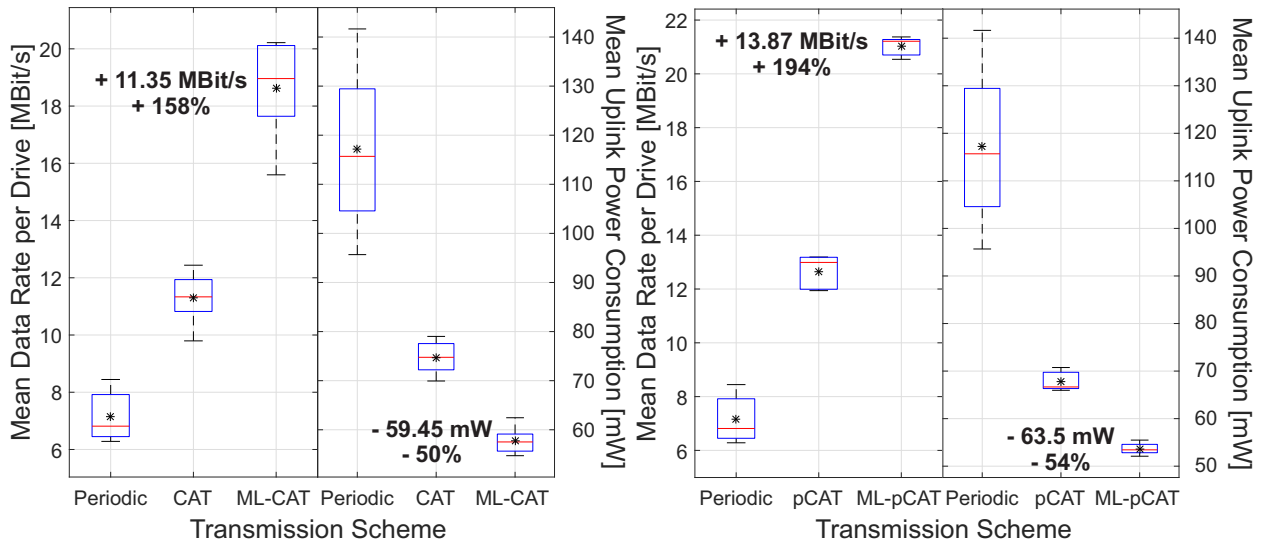


Fig. 6. Results of real world data transmissions for context-aware (CAT-based) and context-predictive (pCAT-based) approaches. Through integration of machine learning-based data rate prediction both approaches can massively increase the resulting data rate while simultaneously reducing the power consumption of the mobile UE.

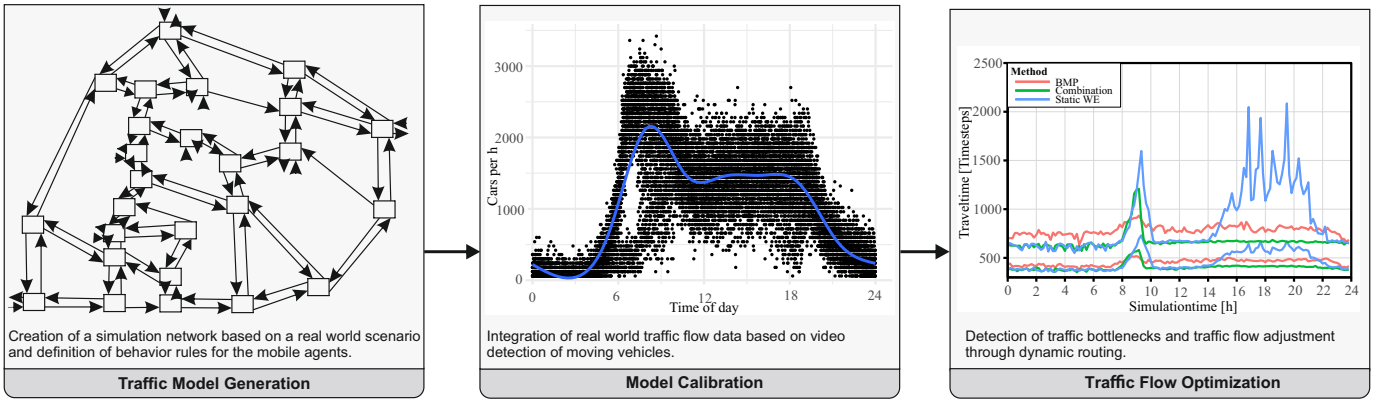


Fig. 7. Methodological process for modeling and optimization of realistic traffic scenarios based on real world traffic data. Road network topology information is utilized to generate a simulation setup, which is calibrated using sensor data.

pCAT schemes. It can be seen that the integration of machine learning for network quality assessment leads to massive increases in the end-to-end data rate. The proposed approach is able to implicitly consider complex interdependencies between the different network quality indicators and the transmitted data itself. The achieved benefits are even increased by considering the future vehicle context along the vehicle's trajectory with ML-pCAT. Additionally, the average power consumption of the mobile device is significantly reduced. On the one hand, data transmission are more likely performed at good channel conditions with a low transmission power and a low packet loss probability, on the other hand the transmissions themselves are executed faster due to the higher data rate. Since the applied transmission power has a dominant impact on the overall power consumption [33], the device battery is used in a more efficient way.

The results of the considered case study show that anticipatory communication can serve as a powerful mechanism to boost the resource-efficiency of data transmissions from the UE side without requiring cost-intensive extensions of the network infrastructure or the spectrum resources.

### C. Data Analysis and Data Exploitation: Modeling and Optimization of Inner City Vehicular Traffic

In the final case study, we present a process for modeling and optimization of realistic traffic scenarios based on real world sensor measurements. Although the considered scenario, which is described in further details in [29], is focusing on non-automated traffic, it provides valuable methodological groundwork for the proposed overall system model as it involves real world data acquisition, model generation, data analysis for bottleneck detection and data exploitation for traffic flow optimization using dynamic vehicle routing.

Fig. III-C shows the system model for the considered simulation scenario. In the first step, a model for the road network is derived based on real world map knowledge. Real world data of a video-detection system located around and within the German city Düsseldorf is used to calibrate the

simulation-parameters of the Break-light-CA-model [25] and the traffic-inflow.

For minimizing the global dwell time of the vehicles within the traffic system, the optimization methods Breakdown Minimization Principle (BMP) [34], Wardrops User Equilibrium (WE) [35], and a newly introduced method, which combines both approaches, are compared. The BMP aims to minimize the entropy by maximization of the distance to the critical traffic flow for each detected bottleneck. Although this approach works well for medium to high traffic flows, it causes larger travel times than required for low traffic flow scenarios, resulting in an increased global travel time. The combined approach is able to avoid these situations through integrating the travel times within the model. In the considered scenario, BMP is able to reduce the average travel time by 23 % and the combined model further achieves a reduction by 10 %.

The results show that dynamic routing methods can successfully utilized for reducing the overall dwell time of vehicles within the traffic system. For the hybrid vehicular traffic, the capabilities for propagating routing information to individual vehicles are limited and depend on the availability of communication technology. Therefore, it needs to be investigated how the established models perform if only a fraction of the vehicles implements the optimization principles.

## IV. CONCLUSION

In this paper, we presented a system-of-systems approach for modeling, analysis and optimization of hybrid vehicular traffic system with a focus on data-driven traffic analysis. On the long way towards fully-autonomous traffic, the coexistence of systems with heterogeneous automation levels leads to novel challenges for traffic management systems, which need to be addressed by new methods of data acquisition, data transfer, data analysis and data exploitation. Therefore, proposed system-of-systems model relies on close interdisciplinary work of traffic physics, data science and communication engineering. Within multiple case studies, we presented groundwork for all main systems aspects and illustrated ways of using

the existing road and network infrastructure in more efficient ways. In future work, we will continue to transfer the different individual components into the proposed overall system model. For analysis and optimization, simulation models based on measured system properties will be derived. After the simulative optimization phase, adjustments will be fed back into the real world systems.

#### 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] W. Brenner and A. Herrmann, *An overview of technology, benefits and impact of automated and autonomous driving on the automotive industry*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 427–442.
- [2] C. T. Calafate, K. Cicenía, O. Alvear, J. C. Cano, and P. Manzoni, "Estimating rainfall intensity by using vehicles as sensors," in *2017 Wireless Days*, March 2017, pp. 21–26.
- [3] SAE International, "Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems (Standard No. J3016)," Tech. Rep., 2014.
- [4] L. Ye and T. Yamamoto, "Modeling connected and autonomous vehicles in heterogeneous traffic flow," *Physica A: Statistical Mechanics and its Applications*, vol. 490, pp. 269 – 277, 2018.
- [5] M. Zhou, X. Qu, and S. Jin, "On the impact of cooperative autonomous vehicles in improving freeway merging: A modified intelligent driver model-based approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1422–1428, June 2017.
- [6] C. Chen, T. H. Luan, X. Guan, N. Lu, and Y. Liu, "Connected vehicular transportation: Data analytics and traffic-dependent networking," *IEEE Vehicular Technology Magazine*, vol. 12, no. 3, pp. 42–54, Sept 2017.
- [7] X. Wang, X. Zheng, Q. Zhang, T. Wang, and D. Shen, "Crowdsourcing in ITS: The state of the work and the networking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1596–1605, June 2016.
- [8] S. Wang, S. Kodagoda, L. Shi, and H. Wang, "Road-terrain classification for land vehicles: Employing an acceleration-based approach," *IEEE Vehicular Technology Magazine*, vol. 12, no. 3, pp. 34–41, Sept 2017.
- [9] J. Guerrero-Ibáñez, S. Zeadally, and J. Contreras-Castillo, "Sensor technologies for intelligent transportation systems," *Sensors*, vol. 18, no. 4, 2018.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] K. Liu, J. K. Ng, J. Wang, V. C. S. Lee, W. Wu, and S. H. Son, "Network-coding-assisted data dissemination via cooperative vehicle-to-vehicle/-infrastructure communications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 6, pp. 1509–1520, June 2016.
- [15] K. Nagel and M. Schreckenberg, "A cellular automaton model for freeway traffic," *Journal de physique I*, vol. 2, no. 12, pp. 2221–2229, 1992.
- [16] L. Habel, A. Molina, T. Zaksek, K. Kersting, and M. Schreckenberg, "Traffic simulations with empirical data: How to replace missing traffic flows?" in *Traffic and Granular Flow '15*, V. L. Knoop and W. Daamen, Eds. Cham: Springer International Publishing, 2016, pp. 491–498.
- [17] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, p. 818, 2017.
- [18] F. Zhao and N. Park, "Using geographically weighted regression models to estimate annual average daily traffic," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1879, pp. 99–107, 2004.
- [19] X. Gong and F. Wang, "Three improvements on KNN-NPR for traffic flow forecasting," in *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*. IEEE, 2002, pp. 736–740.
- [20] W. H. Lam, Y. Tang, and M.-L. Tam, "Comparison of two non-parametric models for daily traffic forecasting in hong kong," *Journal of Forecasting*, vol. 25, no. 3, pp. 173–192, 2006.
- [21] B. F. Selby, "Spatial prediction of AADT in unmeasured locations by universal kriging and microsimulation of vehicle holdings and car-market pricing dynamics," Ph.D. dissertation, University of Texas, Austin, May 2011.
- [22] B. Selby and K. Kockelman, "Spatial prediction of AADT in unmeasured locations by universal kriging," Tech. Rep., 2011.
- [23] T. Liebig, Z. Xu, and M. May, "Incorporating mobility patterns in pedestrian quantity estimation and sensor placement," in *Citizen in Sensor Networks*, 2013, pp. 67–80.
- [24] A. Artikis, M. Weidlich, F. Schnitzler, I. Boutsis, T. Liebig, N. Piatkowski, C. Bockermann, K. Morik, V. Kalogeraki, J. Marecek, A. Gal, S. Mannor, D. Gunopulos, and D. Kinane, "Heterogeneous stream processing and crowdsourcing for urban traffic management," in *Proc. 17th International Conference on Extending Database Technology (EDBT), Athens, Greece, March 24-28, 2014*. OpenProceedings.org, 2014, pp. 712–723.
- [25] W. Knospe, L. Santen, A. Schadschneider, and M. Schreckenberg, "Towards a realistic microscopic description of highway traffic," *Journal of Physics A: Mathematical and General*, vol. 33, no. 48, p. L477, 2000.
- [26] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved IVC analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, Jan 2011.
- [27] B. Sliwa, J. Pillmann, F. Eckermann, L. Habel, M. Schreckenberg, and C. Wietfeld, "Lightweight joint simulation of vehicular mobility and communication with LIMOsim," in *IEEE Vehicular Networking Conference (VNC)*, Torino, Italy, Nov 2017.
- [28] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, "Dynamic route planning with real-time traffic predictions," *Inf. Syst.*, vol. 64, no. C, pp. 258–265, Mar 2017.
- [29] T. Vranken, B. Sliwa, C. Wietfeld, and M. Schreckenberg, "Performance comparison of dynamic vehicle routing methods for minimizing the global dwell time in upcoming smart cities," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, USA, Aug 2018.
- [30] B. Sliwa, N. Piatkowski, M. Haferkamp, D. Dorn, and C. Wietfeld, "Leveraging the channel as a sensor: Real-time vehicle classification using multidimensional radio-fingerprinting," in *2018 IEEE 21st International Conference on Intelligent Transportation Systems (ITSC)*, Maui, Hawaii, USA, Nov 2018.
- [31] C. Ide, B. Dusza, and C. Wietfeld, "Client-based control of the interdependence between LTE MTC and human data traffic in vehicular environments," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 1856–1871, May 2015.
- [32] B. Sliwa, T. Liebig, R. Falkenberg, J. Pillmann, and C. Wietfeld, "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.
- [33] R. Falkenberg, B. Sliwa, N. Piatkowski, and C. Wietfeld, "Machine learning based uplink transmission power prediction for LTE and upcoming 5G networks using passive downlink indicators," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*, Chicago, USA, Aug 2018.
- [34] B. S. Kerner, "Optimum principle for a vehicular traffic network: Minimum probability of congestion," *J. Phys. A: Math. Theor.*, vol. 44, 092001, 2011.



- [35] J. Wardrop, "Some theoretical aspects of road traffic research," in *Proceedings of the Institute of Civil Engineers*, vol. I, no. 1, 1952, pp. 325–378.