This is an unofficial collation of papers edited by Stefan Dietzel and Björn Scheuermann: "Proceedings of the 4th GI/ITG KuVS Fachgespräch Inter-Vehicle Communication (FG-IVC 2016)," Humboldt-Universität zu Berlin, April 2016.

You can find the official proceedings on the edoc server of HU Berlin.

Stefan Dietzel and Björn Scheuermann (Eds.)

Proceedings of the

4th GI/ITG KuVS Fachgespräch Inter-Vehicle Communication (FG-IVC 2016)

31 March and 1 April 2016
Humboldt-Universität zu Berlin

Towards a Multi-Protocol Microscopic IVC Simulation Environment for ADASs

Julian Timpner, Martin Wegner, Hendrik-Jörn Günther and Lars Wolf
Institute of Operating Systems and Computer Networks
Technische Universität Braunschweig
Braunschweig, Germany
Email: (timpner|wegner|guenther|wolf)@ibr.cs.tu-bs.de

***Abstract*—Most of today's ADASs rely on the data of local perception sensors. With the introduction of Vehicle-to-X (V2X) communication, the perception range of next generation vehicles will be extended even further. In addition to the introduction of local perception sensors in Veins, we present *ArteryLTE*, a holistic simulation environment for using the vehicle position data from the microscopic traffic simulator SUMO in combination with ETSI ITS G5 and LTE enabled vehicles, as well as a backend. We use the proposed simulation framework to present preliminary results for detecting lane-level traffic phenomena.**

I. Introduction

Until a few years ago, simulations within the field of automotive research have been focusing on very specific problems regarding the development of components of vehicles. For the analysis of parameters regarding the powertrain, for example, dedicated simulation tools that provide very detailed models of Internal Combustion (IC) engines are used. Engineers focusing on the development of the chassis use different types of simulation tools in order to determine the stresses and forces acting on the vehicle's structural elements during operation. These approaches have in common that they can be simulated and analyzed separately as their interactions can be somewhat simplified. However, these simulation models always try to model the behavior of a single vehicle or components within the vehicle with respect to external influences. When it comes to the simulation of actively intervening Advanced Driver Assistance System (ADAS) such as an Adaptive Cruise Control (ACC), at least one more vehicle within the vicinity of the simulated vehicle needs to be taken into account.

New assistance systems under development today will be based on V2X communication and will gain major momentum within the next years. As outlined above, most of today's tools for the development of ADASs are unable of providing a framework for the analysis of communication-enabled assistance systems. Next to the already existing factors influencing the working principle of today's ADASs, V2X enabled ADASs are also affected by the properties of the employed communication technologies. Therefore, the development of ADASs also advances into the area of different communication technologies. As a consequence, new development tools are required that provide both, the option of simulating the Inter-Vehicle Communication (IVC) as well as the ADAS applications.

With the option of actively exchanging messages between vehicles, cooperation between road participants may be realized. Therefore, ADASs based on V2X communication do not only have a value for the customer, but may also improve traffic flows. Hence, the simulation framework needs to provide the option to analyze both: the interaction of the algorithms on the vehicles on a sub-microscopic level, i.e., on the vehicle level, as well as on the microscopic level to analyze their effect on the traffic efficiency.

This paper focuses on the realization of a holistic simulation framework capable of simulating different communication technologies as well as realizing different implementations of ADAS applications. Section II provides a short overview of the related work regarding simulation tools and their applications. Furthermore, the section introduces our contribution towards the consolidation of different simulation approaches within a common framework, with a strong focus on the integration of local perception sensors. In Section III we also present a particular application of our work, by introducing a backend to our simulation framework. Section IV summarizes our future research.

II. Simulation Framework

As described in Section I we propose a holistic simulation framework called *ArteryLTE* which is capable of simulating ADAS applications and IVC networks on both, the sub-microscopic and microscopic level.

A. Microscopic vehicular network simulation

Depending on the research question, either the network between the vehicles has to be simulated—therefore taking into account that the applications running on the vehicles are not considered—or, in the other case, the focus lies on the simulation of the ADAS application, therefore neglecting the inter-vehicle network. Regardless of the research question, the inherent simplifications make it challenging to analyze their mutual effects. On the one hand, in the case of the stand-alone network simulation, the movement of the network nodes may be random or based on recorded traces without the option of changing the movement patterns due to the interaction of an ADAS application. On the other hand, when developing a novel ADAS application, the limitations of the inter-vehicle network may not be considered. A summary regarding the possible approaches of combining both requirements is provided by Sommer et al. [1].

One approach for joining both perspectives is proposed by the popular open-source *Vehicles in Network Simulation (Veins)* framework [2] for which several extensions are publicly

available. The *Veins* project[1] introduces the combination of the dedicated network simulator OMNeT++ with the dedicated traffic simulator Simulation of Urban Mobility (SUMO). *Veins* implements SUMO's control protocol Traffic Command Interface (TraCI) in OMNeT++ and therefore realizes the online import and manipulation of the vehicles simulated by SUMO through functionalities implemented in OMNeT++. What is more, Veins also provides an implementation of the US Wireless Access in Vehicular Environments (WAVE) Dedicated Short Range Communication (DSRC) communication stack based on IEEE 802.11p.

B. ADAS application development

Riebl et al. [3] present an extension for the Veins framework, called *Artery*[2], which focuses on the implementation of applications (so-called Artery services) for the vehicles within the simulation. The modular architecture of Artery enables heterogeneous vehicle capabilities, by dynamically configuring both, the penetration rate of a communication technology as well as the applications the vehicles are capable of. Furthermore, Artery introduces *Vanetza*, an implementation of the European Telecommunications Standards Institute (ETSI) Intelligent Transport System (ITS) G5 protocol stack alongside the WAVE stack provided by Veins. As part of this extension, Artery also includes a service for disseminating Cooperative Awareness (CA) messages according to the standard [4].

C. LTE support

One project, focusing on the introduction of another communication technology, is *VeinsLTE* [5]. The extension introduces the capability of heterogeneous communication technologies on mobile network nodes. Next to Veins' original WAVE stack, *VeinsLTE* employs *SimuLTE* [6], a complete representation of an Long Term Evolution (LTE) stack within OMNeT++. *VeinsLTE*, however, lacks the option of simulating a different set of applications per vehicle as well as the ETSI ITS G5 protocol stack for direct V2X communication. As part of our research, we combined *Artery* and *VeinsLTE*, therefore circumventing these shortcomings. Instead of the *decision maker* provided by *VeinsLTE*, *Artery*'s middleware is extended by the option of choosing either the ITS G5 or the LTE stack for communication. Upon message generation, the *Artery* services provide information to the middleware in order to choose the appropriate communication technology. We named the combination of *Artery*'s flexible application layer and the LTE communication stack *ArteryLTE*, which will be publicly available[3].

D. Backend Support

As we intend to embed a backend service into our simulation, a static network node is introduced to the network which is connected to the eNodeBs of the LTE network. The overall architecture of our simulation framework is depicted in Figure 1. Located within the presented cell of the eNodeB are two vehicles 1 and 2 which are equipped with both, an LTE and an ITS G5 stack. The location and dynamic

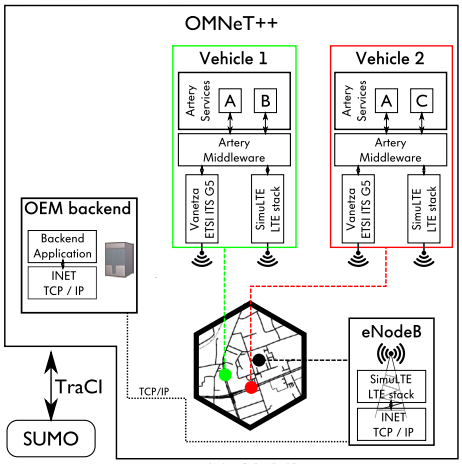


Figure 1: Architecture of the *ArteryLTE* simulation environment.

status of the vehicles are extracted from SUMO via the TraCI protocol. Each vehicle can be provided with different Artery services A, B or C. The *Artery* middleware on the vehicles is responsible for selecting the appropriate communication stack. When transmitting data via LTE to the backend of an Original Equipment Manufacturer (OEM), a Transfer Control Protocol (TCP) connection between the eNodeB and the backend is used.

E. Local perception sensors

When developing novel ADAS applications based on V2X communication, the key difference to today's applications is the vehicle's capability of perceiving objects that are located outside of the perception range of its local perception sensors. However, the information received by V2X communication can be enriched by fusing these information with the data obtained by the vehicle's local perception sensors. Moreover, locally perceived objects may be shared with other vehicles, to provide a more detailed description of the vehicles' environment, as introduced by Günther et al. [7].

The basis for the local perception system within *Artery* is the Global Environment Model (GEM), which acts as a global database for all objects within the simulation and therefore resembles a map of all objects within the simulation on a global scale. Whenever a node is introduced, a Global Environment Model Object (GEMO) is added to that database. A GEMO mainly consists of the pointer to the mobility model of the vehicle within the simulation as well as some further information required to describe the object, such as its geometric dimensions, attachment points for local perception sensors as well as a list of vehicles that have knowledge about that particular GEMO. Whenever a vehicle changes its dynamic properties due to a new simulation step from SUMO, the corresponding GEMO is also updated. The GEM acts as the backbone to the perception system within the simulation. The determination of the presence of an object within the perception range of a sensor is performed by the GEM.

Whereas only one instance of the GEM exists within the simulation, every vehicle equipped with a local perception

[1] http://veins.car2x.org/
[2] https://github.com/riebl/artery
[3] https://github.com/ibr-cm/artery-lte

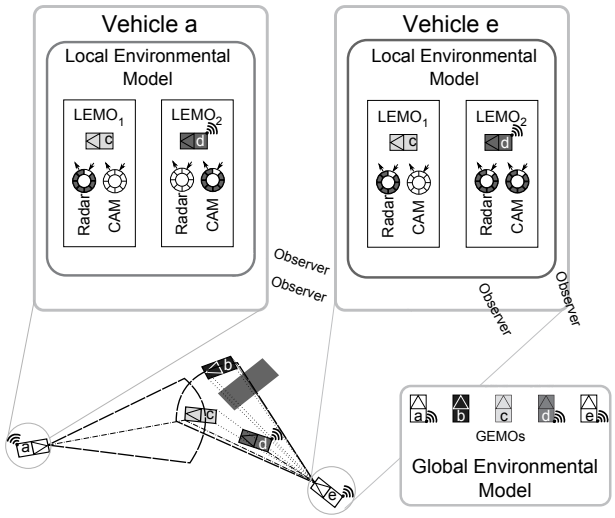


Figure 2: Creation of Local Environment Model Objects (LEMOs) for observed vehicles (indicates a V2X enabled vehicle)

sensor creates its own instance of a Local Environment Model (LEM), as depicted in Figure 2. As every vehicle maintains its own LEM, it acts as a database for all objects that are known to the specific vehicle only. The LEM is part of the Facilities offered by the *Artery* framework and can therefore be accessed by any *Artery* service. Whenever a measurement is performed by the sensor, the objects within its perception range are added as Local Environment Model Objects (LEMOs) to the database within the LEM, as depicted in Figure 2 for vehicles a and e. In analogy to the GEMO, each LEMO also knows about the pointer to the mobility model that belongs to the observed vehicle. Whenever a vehicle is first measured by a perception sensor, i. e., the vehicle has not been sensed by the measuring vehicle before, a new LEMO is created for that vehicle.

Next to the pointer to the mobility model of the described vehicle, the LEMO also consists of several circular buffers, each assigned to a perception sensor of a vehicle, as depicted in Figure 2. The buffers are responsible for storing the measurements of a perception sensor to the observed vehicle at the time of measurement. This feature allows for the creation of a history of measurements for a LEMO, whereas the length of the history, i. e., the number of measurements stored for each object, is variable. The history for each object may be used by a vehicle in order to transmit aggregated information to an OEM backend via an LTE connection.

Every LEMO within the LEM database exists for a limited amount of time only. Whenever the vehicle has not been perceived again by any sensor of the perceiving vehicle within this limited amount of time, it is removed from the LEM. This allows for the continued consideration of vehicles within the algorithms of Artery services, even if the observed vehicle temporarily is not within any line-of-sight of the sensor, i. e., when obstructed by a crossing vehicle. Each sensor is defined as a separate class, by deriving from a base sensor class. The base class provides virtual functions which have to be defined by the sensor according to its properties, such as the variables that can be measured by the sensor. A radio detection and ranging (radar) sensor, for example, will return the relative velocity and distance to the observed object, whereas a camera will also return the orientation of the observed object as well as its geometric dimensions.

III. Simulation Results

As an example application for the proposed *ArteryLTE* framework, we run a high-resolution telemetry service on the backend in order to detect lane-level traffic phenomena.

A. Setup

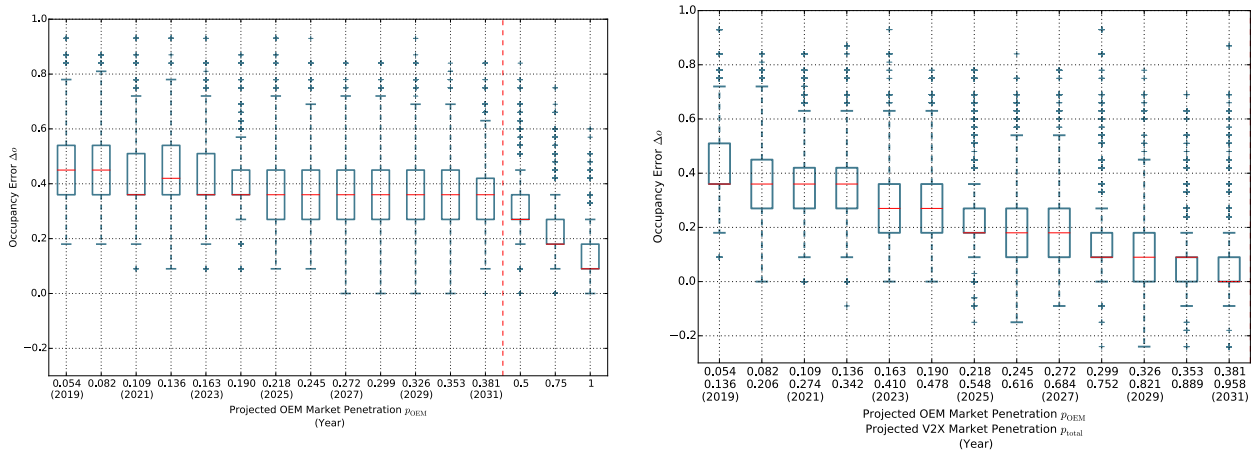
The vehicles in the simulation can be equipped to either communicate (a) with a backend via LTE, (b) with other vehicles via ETSI ITS G5, or (c) both. Vehicles of Type (a) send their dynamic state (i. e., their current position, speed, etc.) with a given frequency f_{lte} to the OEM backend only. Type (b) vehicles broadcast CA messages according to the generation rules specified in the standard [4] and therefore represent those vehicles from other OEMs. Just like the vehicles of Type (a), Type (c) vehicles, in addition to broadcasting CA messages, also send a status update to the OEM backend at a rate of f_{lte}. In addition to their own dynamic state, the status updates include an aggregation of all received CA messages of all V2X vehicles in their vicinity since their last communication attempt with the backend.

To achieve a higher than lane-level resolution of local traffic phenomena, every lane is subdivided into lane sections of 50 m each. As each vehicle should at least report its status once per section, the minimum update frequency should be $f_{\text{lte}} = 1/3.6\,\text{Hz}$, due to a nominal speed limit of 50 km/h. As the driver imperfection in the simulation might lead to vehicles exceeding the speed limit, they alternatively report their status every 50 m.

In our simulations, we aim to determine the required market penetration rates of LTE and ITS G5 communication technologies to achieve a sufficient accuracy in the description of the current traffic situation estimated on an OEM's backend. For this purpose, we conducted two simulation stages and performed the analysis from the perspective of a backend:

Stage 1: The purpose of this stage is to establish the minimum required market penetration rate of the LTE backend communication in order to enable the envisioned service. We conducted 16 simulation runs, varying the penetration rates p_{oem} for the vehicles equipped with the LTE backend communication capabilities according to extrapolated market penetration rates of the Volkswagen (VW) group. Hence, the backend will only receive information transmitted by vehicles of the VW group.

Stage 2: In order to determine the additional impact of Vehicle-to-Vehicle (V2V) communication, the second stage adds Type (b) vehicles according to the assumed total market penetration rate p_{total}. These vehicles transmit CA messages within their local communication range. Consequently, the VW group vehicles are now of Type (c) and therefore transmit the collected CA messages to the OEM's backend as well. This causes a virtual increase of the number of vehicle positions in the backend, therefore yielding a larger database for estimating the current traffic situation.



(a) Stage 1

(b) Stage 2

Figure 3: Comparison of occupancy errors Δo per simulation stage.

B. Data Analysis

To assess the accuracy of the telemetry service, we opted for the occupancy of lane sections as a relevant evaluation metric. The occupancy o of a section s is defined as the overall space occupied by vehicles on s relative to the length of s. o_{backend} is based on the positions of vehicles reported to the backend by OEM vehicles (Type (a) or Type (c)). o_{TraCI} is based on the actual position of every vehicle as obtained via TraCI. As o_{TraCI} represents the ground truth, we determine the error Δo of the telemetry service as the difference between the two values: $\Delta o = o_{\text{TraCI}} - o_{\text{backend}}$

C. Findings

Figure 3 shows the distribution of the occupancy error Δo for increasing market penetration rates p_{OEM} (Figure 3a) and p_{total} (Figure 3b). Figure 3a shows that with an increasing market penetration rate p_{OEM} of Type (a) vehicles, the occupancy error Δo decreases. However, it becomes clear that with the OEM's maximum penetration rate $p_{\text{OEM}} = 38.1\,\%$ (achieved by the end of 2031), Δo is not significantly lower than at the time of market introduction in the year 2018. Even with a (very large) theoretical market share of $p_{\text{OEM}} = 75\,\%$ a median occupancy error Δo of about 20 % occurs. This shows that using the fleet-data from one OEM alone will not be sufficient for realizing the envisioned telemetry service.

Figure 3b shows the leverage of introducing V2V communication. An OEM market share of $p_{\text{OEM}} = 38.1\,\%$ now corresponds to a total V2V market penetration of $p_{\text{total}} = 95.8\,\%$. As shown in Figure 3a, increasing p_{OEM} alone has no significant impact on the occupancy error Δo. However, under the assumption of a parallel introduction of V2V technology to the market, the potential for a high-resolution Floating Car Data (FCD) aggregation can be fully leveraged.

IV. Next Steps

As part of this paper, we gave an overview of the existing approaches for combining network simulation and traffic simulation for the purpose of analyzing V2X communication based ADAS applications. Additionally, we present *ArteryLTE*, a holistic simulation environment for using the vehicle position data from the microscopic traffic simulator SUMO in combination with ETSI ITS G5 and LTE enabled vehicles based on Veins. As an extension to the framework, we focus on the sustainable introduction of local perception sensors for the vehicles—a feature not yet existing within the Veins community. Preliminary simulation results show the effectiveness of a dual simulation stack within the vehicles with respect to future backend applications. Our future work will focus on employing the local perception sensors to enrich the database on the OEM's backend.

References

[1] C. Sommer and F. Dressler. "Progressing Toward Realistic Mobility Models in VANET Simulations". In: *IEEE Communications Magazine* 46.11 (Nov. 2008), pp. 132–137.

[2] C. Sommer, R. German, and F. Dressler. "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis". In: *IEEE Trans. Mobile Comput.* 10.1 (Jan. 2011), pp. 3–15.

[3] R. Riebl et al. "Artery - Extendig Veins for VANET applications". In: *Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. 2015.

[4] *ETSI EN 302 637-2 V1.3.1 - Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service*. ETSI, Sept. 2014.

[5] F. Hagenauer, F. Dressler, and C. Sommer. "Poster: A simulator for heterogeneous vehicular networks". In: *Proc. Vehicular Networking Conference (VNC)*. IEEE, Dec. 2014, pp. 185–186.

[6] A. Virdis, G. Stea, and G. Nardini. "SimuLTE-A modular system-level simulator for LTE/LTE-A networks based on OMNeT++". In: *Proc. SIMULTECH*. IEEE. 2014, pp. 59–70.

[7] H.-J. Günther, O. Trauer, and L. Wolf. "The potential of collective perception in vehicular ad-hoc networks". In: *Proc. ITS Telecommunications (ITST)*. Dec. 2015, pp. 1–5.

Using Searchable Encryption to Protect Privacy in Connected Cars

Matthias Matousek, Christoph Bösch and Frank Kargl
Institute of Distributed Systems
Ulm University
{matthias.matousek, christoph.boesch, frank.kargl}@uni-ulm.de

***Abstract*—Providing vehicles with extended connectivity introduces new opportunities for services, and also security applications such as misbehavior detection. However, for many applications, personal data needs to be processed by the system providers, which impairs the privacy of the vehicle users. While focusing our research on new possibilities of connected car security, we follow *privacy by design* principles. We explore the utilisation of various privacy-enhancing technologies (PET) in order to provide advanced connected car applications, while preserving the personal data of the vehicle users. Specifically, we aim to develop practical schemes that utilise Searchable Encryption to provide a framework for secure and privacy-preserving connected car applications.**

I. Introduction

Connected cars are a growing topic for car manufacturers. Most automotive brands now offer services that provide control or information querying of vehicles via web interface or smart phone app. As soon as a specific vehicle is paired with a user account, the owner can use services such as the localisation of his car, querying information like travelled distance or gas usage, and even remote control some functions such as heating or unlocking and locking doors.

As the trend of connected cars leads to more and more data about vehicles being saved and processed in back end systems, the security of theses systems needs to be considered. Additional connectivity and services may increase the attack surface of modern vehicles, but it also increases the potential for detection and prevention of misbehavior and malfunction. Thus, our research focuses on new methods to utilize connected car data in order to perform misbehavior detection and security event management on a fleet-bases.

While security monitoring systems for connected cars are a promising approach, it also raises the question of user privacy. For such a security system, large amounts of personal data need to be processed. A personal vehicle often accompanies its users wherever these travel, thereby collecting large amounts of information on their whereabouts, behaviour, and possibly even lifestyle.

Research has shown that having access to specific car data allows adversaries to deduce further information, such as identifying behaviour or the identity of the corresponding individuals.

Thus, the privacy of car users requires protection. The users' trust in a provider or vendor should not be violated. Further, even when the manufacturer shows only good intentions, data should also be secure when it is leaked due to circumstances such as hacking attacks on the provider back end. Moreover, legal subpoenas that require businesses to hand over their users' data, as well as other attempts to access private data such as surveillance programs, unsettle customers. The usage of a vehicle often accurately reflects its owners' behavior, and thus the generated data needs to be protected.

II. User Privacy through Searchable Encryption

Searchable Encryption (SE) [1] is a promising technology in an approach to provide fine-grained access control for encrypted data. The idea of SE is to enable search algorithms to work on ciphertexts without the need of prior decryption, and without even the need to have knowledge of the corresponding secret key.

A. Searchable Encryption

Generally, two approaches for SE schemes exist. One approach is to use a specialised encryption scheme that allows the ciphertext to be searched directly (e.g. Song et al. [2]). Index-based schemes, however, have better search performance and allow for arbitrary encryption ciphers. Thus, we focus on the latter approach.

While the client (the entity that generates and encrypts the data) is always able to access his data, he can also generate a so-called trapdoor to enable another party to perform a search on the encrypted index. The trapdoor is bound to the specific keyword that it was created for. Thus, the client can limit access to particular data.

In general, a SE scheme consists of the following four algorithms:

- $K_C(\lambda)$: This algorithm is run by the client C, takes a security parameter λ as input, and outputs a secret key K_C.
- $I(K_C, D)$: This algorithm is run by the client C, takes a key K_C and data items D as input, and outputs an encrypted index I, which allows to search D for specific keywords (using a trapdoor).
- $T_s(K_C, s)$: This algorithm is run by the client C, takes a key K_C and a search keyword s as input, and outputs a trapdoor T_s.
- $X(T_s, I)$: This algorithm takes a trapdoor T_s for search keyword s and an encrypted index I as input, and outputs

the query result X. This might be a handle to the (encrypted) data entry or the data item itself

In the vehicular context, data that is generated and collected within a vehicle should be sent to and stored on e.g. a back end server or shared with third-party providers. In order to protect it from unauthorized access, all data is encrypted before it leaves the car. Using SE, the user may generate trapdoors that allows the back end provider to access specific data, and thus to process it without breaching the users' privacy. Similarly, SE can be utilised to manage access for several parties, thereby allowing third-party services.

Depending on the use case scenario, the server requires more or less access to data. In some cases it might suffice to learn whether a given keyword is present in the ciphertext. Often, however, additional knowledge is required. Different SE schemes can provide further access. Schemes that allow for range queries can be used for data that is within certain limits [3], and decryptable SE can even give the trapdoor-holder the ability to access the plaintext of the search result [4].

B. Application Scenarios

Many applications could benefit from the privacy protection that can be achieved with SE, while at the same time being able to function normally.

Pay-as-you-drive insurance policies (PAYD) are a recent trend that bases the costs of an insurance policy on driving behavior. It has the potential to be a fairer alternative to traditional blanket coverage, but introduces severe privacy issues. SE could be used to ensure that only necessary data is accessed. E.g. range queries could be employed to determine whether acceleration regularly exceeds a certain threshold, or the vehicle is driven at night time.

Connected car services and the security of connected vehicles are our primary interest for the application of SE. Similarly to the PAYD scenario, other services could be limited to their required data with SE schemes. We are specifically interested in determining whether SE can be used to enable a back-end-located misbehavior detection that is privacy-preserving.

C. Discussion

For applications within connected vehicles other cryptographic primitives could be thought of, that might be applicable to the presented use cases. This raises the question whether SE is better suited.

Secure multi-party computation has the goal to let several parties compute functions, while keeping their respective input data private [5]. While this is fitting for the envisioned use cases, it would require frequent collaboration of the vehicle with the back end servers. This is impractical for vehicles in deployment. The back end needs to handle potentially large amounts of data quickly. It thus cannot rely on communication with all the vehicles, which might not even be reachable all the time.

Functional Encryption (FE) is a related technique that can be used to perform computations on encrypted data and gain access to the computed result in cleartext without requiring the secret key to the ciphertext [6]. However, most implementations of FE only achieve low performance. SE is currently superior in this regard.

While SE is relatively efficient in regard to computing complexity, it has several limitations. It remains to be evaluated whether it applies to all of our use cases, and whether it is flexible enough to be used in the automotive context that constantly progresses and introduces new applications.

III. Conclusion and Future Work

We proposed the application of Searchable Encryption (SE) in order to provide fine-grained access control to vehicular data of connected cars. Due to its good performance, it is suited to provide privacy protection even in scenarios where large amounts of data are processed. In addition to merely identifying present keywords in a ciphertext, SE can also provide further access to the encrypted data—such as ranges of numeric values, or even the cleartext of search results.

Future work will consist of the identification of the required data processing capabilities, and whether SE can be applied to the given scenarios. We are specifically interested in performing misbehavior detection over an entire fleet in the back end.

In addition to the assessment of suitable schemes, we are planning to implement a novel protocol for privacy-preserving data sharing using SE in the automotive context. A subsequent evaluation of the system is expected to provide us with insights on its usability, applicability and performance.

The eventual goal is to provide a framework that allows for data sharing with strong privacy protection for different applications.

References

[1] C. Bösch, P. Hartel, W. Jonker, and A. Peter, "A survey of provably secure searchable encryption," *ACM Comput. Surv.*, vol. 47, no. 2, pp. 18:1–18:51, Aug. 2014.

[2] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Security and Privacy, 2000. S P 2000. Proceedings. 2000 IEEE Symposium on*, 2000, pp. 44–55.

[3] D. Boneh and B. Waters, *Theory of Cryptography: 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, The Netherlands, February 21-24, 2007. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, ch. Conjunctive, Subset, and Range Queries on Encrypted Data, pp. 535–554.

[4] T. Fuhr and P. Paillier, *Provable Security: First International Conference, ProvSec 2007, Wollongong, Australia, November 1-2, 2007. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, ch. Decryptable Searchable Encryption, pp. 228–236.

[5] O. Goldreich, "Secure multi-party computation," *Manuscript. Preliminary version*, pp. 86–97, 1998.

[6] D. Boneh, A. Sahai, and B. Waters, *Theory of Cryptography: 8th Theory of Cryptography Conference, TCC 2011, Providence, RI, USA, March 28-30, 2011. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, ch. Functional Encryption: Definitions and Challenges, pp. 253–273.

Security Evolution in Vehicular Systems

Dominik Lang, Christopher Corbett, Frank Kargl
Institute of Distributed Systems, Ulm University, Germany
Email: {dominik.lang,christopher.corbett,frank.kargl}@uni-ulm.de

***Abstract*—Modern vehicles contain a complex network of computer systems, which makes security considerations a necessary part of the design process. Due to a vehicle's long deployment phase, static security solutions become obsolete and ineffective over time. Research has mostly focused on how to improve security in vehicles, however, not addressing the need to keep security solutions effective during the entire lifetime of a vehicle. In order to address the changing environment there is a need to create strategies for architectural components, security mechanisms, and update processes that together enable the evolution of the security mechanisms themselves. Our approach is to analyze security mechanisms for how they can fail and what this means from a security evolution perspective. Based on this analysis we can then create solutions in order to evolve deployed security mechanisms over time.**

I. Introduction

Vehicles have evolved into complex computer systems containing up to 100 different Electronic Control Units (ECUs) of which many are interconnected. These ECUs provide functionality for different types of tasks that have different requirements, e. g. low latency or high bandwidth. As a result, the internal network infrastructure in vehicles has evolved to contain several different bus systems that provide different properties for different use cases. In addition, modern automobiles also provide wireless and wired interfaces to the outside world including Bluetooth, WLAN, cellular networking, on-board diagnostic ports (OBDs), USB, and many more. These technologies enable useful services and functionalities, but on the other hand also create new attack surfaces for potential attackers. Especially the wireless interfaces to the outside world open up possibilities for remote attacks.

With these new attack surfaces, it is consensus that security mechanisms are an important aspect to add to the on-board infrastructure, and literature knows a vast number of proposals on how to enhance automotive security (e. g. [1], [2]). However, automotive security is challenged by the very long life cycles of vehicles and also by their safety requirements, which mandate a conservative approach to making changes to deployed vehicles. Thus, for every modification it needs to be ensured that no regressions or flaws are introduced into the safety-critical components, similar to other safety-critical fields such as industrial control systems (ICS) [3]. This conservative approach to automotive IT systems engineering conflicts with the typical approach in IT security, where security mechanisms age, need to be enhanced or replaced, and fast reaction to new attacks is required.

This contradiction is what we address in our work on *security evolution*. Within this work we have split our aims into two separate categories:

1) the identification and categorization of security mechanisms and problems in an automotive systems' life cycle, and
2) the proposal of flexible ways for security mechanisms and architectures to evolve during the life cycle of a vehicle in order to always maintain the required level of security.

The process of security evolution needs to keep security mechanisms in a vehicle updateable and secure at all layers, from hardware to software, and also involves external components, such as public key infrastructures (PKIs). In this paper we present a categorization of domains that security evolution needs to address.

The remainder of this paper is structured as follows: Section II gives a brief overview of security in modern vehicles, Section III categorizes security mechanisms and potential failures, and finally Section IV concludes this paper.

II. Security in Modern Vehicles

Security in vehicles can be categorized into on-board security and V2X communication security (vehicle-to-vehicle and vehicle-to-infrastructure). This work is applicable to both categories, as it focuses on security systems, mechanisms, and components deployed inside a vehicle, which are used to secure the entire infrastructure, i. e. both on-board and V2X security.

ECUs in modern vehicles can be categorized into powertrain, chassis, body and comfort, and driver assistance and safety. As part of the comfort category, a vehicle can contain systems for infotainment (e. g. navigation systems and radio) and telematics units, which are connected to backend servers via a cellular network (e. g. GM's OnStar). Future extensions may introduce other V2X capabilities such as vehicular ad hoc networks (VANETs), for example using dedicated short range radio communication (DSRC) for cooperative safety applications.

The infotainment, telematics, and in general V2X systems provide attack surfaces for remote attacks. Especially telematics units have been targeted by previous work to gain access to the on-board vehicular system via cellular communication [4], [5], [6] and thus allowing the attacker to remain at a safe distance.

On the other hand, the on-board system provides an attack surface for controlling various aspects of the vehicle, including

safety-critical functionalities, such as the brakes and throttle. Attacks have focused on the ability to extract arbitrary data from ECUs by reading their memory and flashing ECUs with malicious code. Attacks have also exploited the ability to replay messages and arbitrarily spoof messages on the CAN bus [6], [7].

As VANETs are not yet deployed, there are no real attacks, but research and development has addressed security (and privacy) from the start. The IEEE 1609.2 standard [8] specifies the use of a PKI in order to equip vehicles with certificates that allow to sign messages between the communicating partners in order to ensure integrity.

In order to protect against malicious modifications of ECU software and unauthorized spoofing of messages, the most important aspects of security for on-board vehicular systems is to provide integrity, authentication, authorization, and availability. Therefore, research has proposed how to secure in-vehicle networks, both in securing the network itself, such as CAN, and in creating a secure architecture for on-board systems [9], [10].

In summary, while risk levels may differ substantially depending on the field of application, security mechanisms ensuring especially integrity, authentication, authorization, and availability (and to a lesser extent confidentiality) require evolutionary capabilities independent of the compartment where they are used.

III. Security Evolution

To approach security evolution in a meaningful way, one first needs to assess and categorize security mechanisms in vehicles and V2X systems. Therefore, we performed a systematic security and risk analysis of security technologies and systems used in vehicles at all layers, e. g. hardware, software, cryptography, architecture, protocols, and network technologies.

Based on this analysis, we examine possible causes that result in a necessary evolution of security mechanisms. We categorize the involved security mechanisms and potential security failures into:

a) ***Configuration****:* One of the biggest problems for security is complexity. Unfortunately security mechanisms often are very complex and difficult to configure. It is easy to make mistakes, for example when configuring firewall and intrusion detection rules, or access control lists. This problem is also evident on the OWASP Top 10 list [11] with ”Security Misconfiguration” being the fifth most critical web application security flaw. One mistake often allows an attacker to completely circumvent the entire security mechanism. Common misconfigurations are: displaying error handling messages back to the user (e. g. SQL errors), enabled directory listings in web servers, running production software in debug mode, using default key material and passwords, or misconfiguring firewall rules.

b) ***Software implementation****:* Software implementations of security mechanisms in ECUs may have design flaws or bugs, which can allow attackers to circumvent a security mechanism. Classical buffer overflows are one example, which were also used in the prominent Heartbleed attack on OpenSSL. Another implementation bug in OpenSSL was used by the FREAK attack, which allowed a man-in-the-middle attacker to enforce the usage of weak RSA keys, which the attacker could then crack.

c) ***Security protocols****:* Security protocols are secure versions of communication protocols, i. e. they protect the interaction between communicating agents; in the case of vehicular systems this applies to both on-board and V2X communication. Communication protocols are subject to various attacks, such as replay, impersonation, and man-in-the-middle attacks. In order to protect against these attacks, security protocols can become very complex, and it is easy to accidentally overlook an aspect or define false assumptions, which then results in possible attacks. These can be subtle mistakes that are only discovered years later. A well-known example is WEP: after it was standardized and deployed, the first attacks were found, which were based on wrong usage of the RC4 cipher [12]. Another well-known example is the Needham-Schroeder protocol [13], where Denning and Sacco pointed out an attack a few years later [14].

d) ***System security****:* System security mechanisms deployed on the ECU level may be insufficient. For example, Address Space Layout Randomization (ASLR) may fail, if not properly implemented or if more sophisticated attacks become available.

e) ***Symmetric cryptography (including hash functions and random number generation)****:* As recently seen with SHA-1 [15], attacks against cryptographic symmetric ciphers and hash functions may become more sophisticated or powerful, and mechanisms considered secure five years ago may not be nowadays. Beyond, even if the algorithm itself is still secure, key lengths may not be appropriate after some some years from now. As an extreme example, many implementations use AES with 128-bit keys; in case of the successful construction of large quantum computers, it is necessary to change the key length to 256-bits due to Grover’s algorithm [16], which (from a brute-force search point of view) cuts the number of bits in half, i. e. a 128-bit key can be recovered in 2^{64} steps.

f) ***Asymmetric cryptography****:* The same that applies to symmetric cryptography applies also to asymmetric cryptography. We discuss it separately as the mechanisms are often fundamentally different and key lengths substantially longer. A prominent example in this category is the Logjam attack on TLS [17]: this attack used the fact that many servers were using a single 512-bit group for the Diffie-Hellman key exchange, which the researchers then precomputed. They

were then able to calculate the key from the key exchange. The solution to this problem is to use 2048-bit or larger primes, thus creating the need to update the key material in the field.

*g) **Hardware security modules and functions**:* For performance and cost reasons, some of the mechanisms listed above are cast in (immutable) hardware; in this case replacement is not straightforward. Even if flexible hardware like FPGAs are used, their performance limitations may hinder deployment of more powerful mechanisms, and cost constraints prevent proactive deployment of hardware with spare performance.

*h) **Backend security functions and trusted third parties**:* Security functions embedded into (web-based) backend systems form another part of the security architecture, most notably PKIs, but also authentication and authorization mechanisms. These backend functions are easier to update in case of compromise. However, retaining interoperability with the deployed fleet is one challenge, as well as the failure of a root of trust (e. g. a compromised root CA). The latter case removes the possibility of trustworthy remote interaction with deployed vehicles.

As this list shows, security mechanisms pervade all parts of vehicular architectures. Security evolution capabilities must also become part of all these components in order to be able maintain the state of security. However, the examples given above show that implementing security evolution may not be straightforward in many cases and require further research.

Our next step is to focus on the categories one by one, analyze their requirements, technologies, risk structures, design strategies, methods, and architectures with focus on the evolutionary process of the specific security mechanisms.

Often, security evolution cannot be retrofitted but needs to be an initial capability and part of the system architecture. For example, if modification of keys is required, one needs to identify ways to do this without compromising the integrity of the key store. On the software level, implementations need to be structured in a flexible manner without hard coding assumptions on key size, cipher mode, and the cipher itself.

Security evolution strategies can be split into two different categories: *proactive architectural components* and *secure update mechanisms*. Proactive architectural components are put into place in order to allow the evolution of security mechanisms without compromising their security. They are based on foreseeable security problems and provide proactive means for flexibility, for example by replacing a cipher with a stronger one. On the other hand, secure update mechanisms provide generic strategies on how to perform updates in a deployed system. These strategies entail the requirement for secure and verified updates, and thus especially involve authentication, authorization, and integrity considerations. Considering the compromise of root CAs makes this a non-trivial task.

IV. Conclusion

In this paper we discussed that the modern vehicle is in need of security mechanisms for both on-board and V2X systems. However, the long life cycle of automotive systems contradicts with the ageing of security mechanisms. As a result, deployed security mechanisms cannot guarantee security properties in the long run. Consequently, security solutions need to incorporate solutions on how to evolve deployed security mechanisms in order to keep them up-to-date. This is especially difficult in the domain of safety-critical automotive systems, which require a conservative engineering approach. Moreover, we discussed that the introduction of these security evolution mechanisms are not straightforward and require further research. With this contribution, we highlight the need for security evolution and establish it as a future line of research in automotive security.

References

[1] A. Groll and C. Ruland, "Secure and authentic communication on existing in-vehicle networks," in *2009 IEEE Intelligent Vehicles Symposium*.

[2] M. Wolf, A. Weimerskirch, and C. Paar, "Security in automotive bus systems," in *Proceedings of the workshop on embedded security in cars (escar)04*.

[3] F. Kargl, R. van der Heijden, H. Konig, A. Valdes, and M. Dacier, "Insights on the Security and Dependability of Industrial Control Systems," *IEEE Security Privacy*, vol. 12, no. 6, Nov. 2014.

[4] I. Foster, A. Prudhomme, K. Koscher, and S. Savage, "Fast and Vulnerable: A Story of Telematic Failures," 2015.

[5] S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno, "Comprehensive Experimental Analyses of Automotive Attack Surfaces," Tech. Rep., Aug. 2011.

[6] C. Miller and C. Valasek, "Remote Exploitation of an Unaltered Passenger Vehicle," Tech. Rep., Aug. 2015. [Online]. Available: http://illmatics.com/Remote Car Hacking.pdf

[7] K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage, "Experimental Security Analysis of a Modern Automobile," in *2010 IEEE Symposium on Security and Privacy (SP)*, May 2010.

[8] *1609.2-2013 - IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications and Management Messages*, 2013.

[9] P. Kleberger, N. Nowdehi, and T. Olovsson, "Towards designing secure in-vehicle network architectures using community detection algorithms," in *2014 IEEE Vehicular Networking Conference (VNC)*, Dec. 2014.

[10] M. Wolf, A. Weimerskirch, and T. Wollinger, "State of the Art: Embedding Security in Vehicles," *EURASIP Journal on Embedded Systems*, vol. 2007, no. 1, Jun. 2007.

[11] "OWASP Top 10 - 2013." [Online]. Available: https://www.owasp.org

[12] N. Borisov, I. Goldberg, and D. Wagner, "Intercepting mobile communications: The insecurity of 802.11," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '01. ACM.

[13] R. M. Needham and M. D. Schroeder, "Using encryption for authentication in large networks of computers," vol. 21, no. 12.

[14] D. E. Denning and G. M. Sacco, "Timestamps in key distribution protocols," vol. 24, no. 8.

[15] M. Stevens, P. Karpman, and T. Peyrin, "Freestart collision for full SHA-1."

[16] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, ser. STOC '96. ACM.

[17] D. Adrian, K. Bhargavan, Z. Durumeric, P. Gaudry, M. Green, J. A. Halderman, N. Heninger, D. Springall, E. Thom, L. Valenta, B. VanderSloot, E. Wustrow, S. Zanella-Bguelin, and P. Zimmermann, "Imperfect forward secrecy: How diffie-hellman fails in practice," in *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '15. ACM.

Feasibility of Verify-on-Demand in VANETs

Sebastian Bittl
Independent
Munich, Germany
sebastian.bittl@mytum.de

Karsten Roscher
Fraunhofer ESK
Munich, Germany
karsten.roscher@esk.fraunhofer.de

***Abstract*—Wireless ad hoc networks are an important topic in the automotive domain. Thereby, strict security requirements lead to high effort for verification of digital signatures used to secure message exchange. A popular approach to limit such effort is to apply verify-on-demand schemes. However, we show that verify-on-demand requires much more cross layer dependencies than identified before. Moreover, a massive denial of service weakness of this kind of verification mechanism is found. Thus, we recommend to prefer verify-all schemes over their verify-on-demand counterparts.**

***Index Terms*—Verify-on-Demand, VANET, Security;**

I. Introduction

Vehicular ad hoc networks (VANETs) are a topic of high interest, as they promise to increase future safety of driving. Wireless data exchange and realtime requirements of safety critical use cases require strong but yet efficient security mechanisms for VANETs. This even holds for the so called Day 1 use cases, for which mass deployment is planned within upcoming years [1]–[3].

To provide authenticity and integrity of exchanged messages, these are typically signed using primitives of asymmetric cryptography and elliptic curve cryptography (ECC). However, high numbers of received broadcast messages pose performance problems for signature verification at receivers. Realization of hardware verification modules capable of verifying all received messages under high node density is still challenging [4], [5]. Thus, mechanisms for reducing the number of required verifications have been looked at.

Verify-on-demand (VoD) performs a relevance check for the message, which is only verified after is was identified to be relevant, i.e., its contained information is to be used by the receiver. [6] suggests to derive the relevance check directly from the decision making process of applications. This is based on the design of the US WAVE (Wireless Access in Vehicular Environments) standardization framework and extension of this scheme to the ETSI ITS (Intelligent Transport System) standards has not been looked at in detail [7]. Moreover, it is assumed that a relevance decision only affects a single message [6].

We show that the need for authenticated data sets does not only arise from within the application layer, but also from other protocol layers. Moreover, introduction of VoD leads to a complex set of dependencies between functionality on different protocol layers. Furthermore, we find that introduction of VoD significantly decreases the burden for an attacker to perform a denial of service (DOS) attack on a VANET. Additionally, the discovered DOS weakness can be well targeted to dedicated nodes without significant influence on other nodes, which differentiates the attack from other common DOS attacks in VANETs.

Related work is looked at in Section II. General security considerations about VoD are given in Section III. Sections IV and V study cross influence between VoD and routing functionality as well as certificate handling, respectively. Afterwards, Section VI proposes a dedicated DOS attack on VoD. Finally, Section VII provides a conclusion about this work.

II. Related Work and Attacker Model

The VoD concept is introduced in [6] as an alternative to a verify-all scheme. Thereby, a verify-all scheme is characterized by verification of all received messages before these are handed over to further message parsing. In contrast, VoD only verifies messages after a demand for usage of their content has been identified. The impact of VoD on overall security in VANETs is only studied very briefly in [6]. Some background about the VoD concept is given in [7].

VoD was initially proposed for WAVE, i.e., in connection with information dissemination via BSMs (basic safety messages). In contrast, ETSI ITS uses CAMs (cooperative awareness messages) for cyclic data distribution as well as DENMs (decentralized environment notification messages) for on demand information dissemination [3]. Moreover, WAVE does not describe a way to store the content of received messages, while ETSI ITS uses an LDM (local dynamic map) to store dedicated VANET messages.

[8] proposes an implementation of the VoD concept for ETSI ITS by storing signatures alongside with corresponding messages in an extended LDM. Moreover, the LDM also hands over messages to the security entity for verification and stores the corresponding result to avoid multiple verification of a single message. However, our analysis in Section IV shows that this approach leads to massive cross layer dependencies.

Dedicated network layer protocols for VANETs have been developed. Thereby, a major difference between WAVE and ETSI ITS is the support for multi-hop communication, which is only available in ETSI ITS. Within ETSI ITS the term GeoNetworking is commonly used, as network layer routing in VANETs is often based on geographic properties of the vehicular environment. Two major concepts for forwarder selection exit. Firstly, the sender of a message can select the next hop, e.g., by greedy forwarding [9]. Secondly, the sender can be selected in a decentralized manner, e.g., by contention

based forwarding (CBF) [10]. The influence of VoD on multi-hop communication has not been regarded in prior work.

DOS attacks on VANETs are a well known issue. Typically, jamming or misuse of certificate dissemination features is used to raise the channel load in a dedicated area in a way to cause ordinary communication to become at least unreliable or even brake down completely [11], [12]. [6] argues that DOS attack surfaces of the VoD scheme do not increase the vulnerability of the VANET, as VANET security could be always attacked by this kind of attacks. In contrast, we show that VoD significantly reduces the burdens of an attacker to perform a DOS attack. Moreover, the attacker does not even need to violate channel usage regulations and dedicated targeting of single vehicles is possible, which is not the case for prior found DOS attacks.

The attacker is assumed as an active advisory using a single communication unit. This means he can receive, store, modify, create and send arbitrary messages at a single point in the network. However, the attacker has no access to valid credentials to sign his messages.

III. General Security Considerations of Data Handling under a VoD scheme

In general, it is recommended to keep the interface exposed to an attacker as small as possible. For example, [13] argues that the format of the security envelope of ETSI ITS should be adjusted to avoid parsing of its content before signature verification takes place. With VoD, the whole message gets parsed on all protocol layers before the decision on whether to verity the message is done [6]. In contrast, verify-all only exposes low level data processing interfaces up to the network layer security entity. Thus, the surface for an attack on data parsing and usage is increased massively by VoD in comparison to a verify-all scheme.

Within ETSI ITS the data sets on various protocol layers use much more complex encoding in comparison to WAVE, for which VoD was proposed. Thereby, protocol layers above the MAC layer use variable length data sets and deeply nested data types [14]–[16]. On the facility layer ASN.1 encoding, e.g., UPER (unaligned packed encoding rules) for CAM and DENM, is used. Parsing of data sets with such complex encoding schemes requires complex implementations leading to many possibilities for security problems. Even for much simpler ASN.1 schemes, like BER (basic encoding rules), many security problems have been found in the past, e.g., the BERserk vulnerability [17]–[20].

Thus, the effort for secure implementation of all functionality handling received data is significantly increased by using VoD in comparison to a verify-all scheme. This puts the VoD concept into question from a system design perspective.

IV. Interaction of Verify-on-Demand and Network Layer Protocols

In the following we separate the discussion of VoD's influence on the network layer into the impact on VANET specific GeoNetworking and IPv6 over GeoNetworking.

A. GeoNetworking

VoD influences greedy forwarding and CBF for multi-hop communication. Such forwarding algorithms need to keep track of locations of nodes in their surrounding [9]. Thus, each received message leads to an update of a neighborhood table. For greedy forwarding members of the neighborhood table are possible forwarders. In case a forwarder gets selected, the message containing the last known position of this node needs to be verified. Otherwise, an attacker could cause forwarding to non existent bogus nodes (neighborhood table poisoning). This would clearly harm further dissemination of the message's content. This affects received messages which should be forwarded as well as multi-hop messages generated by the node itself (e.g., DENMs).

In case of CBF, the neighborhood table is used to determine whether the own node is a possible forwarder of the received message, i.e., forwarding by the own node causes progress towards the destination. This is required as the prior sending node's position is not contained in a multi-hop message [15]. Thus, its position is determined from the neighborhood table using its MAC address. This requires a verification similar to the case of greedy forwarding. Otherwise, an attacker could foil the CBF algorithm by either causing incorrect forwarding or causing failure to forward by the attacked node.

A significant problem of neighborhood table keeping in connection with VoD is the possibility to cause bogus updates, which replace valid entries in the table. Thereby, an attacker uses the ID of a valid node for its own faked messages. In the worst case, a neighborhood table contains no valid entries at all, due to such an attack. To avoid such an attack, mainly three countermeasures can be thought of.

1) One could verify all messages before the neighborhood table update. However, this disables VoD completely, as every message gets verified.
2) Instead of replacing entries in the table, one could keep prior entries, too. Old entries are only removed after a later update got verified. However, this significantly increases memory requirements, due to an expected low number of verifications.
3) One could only store entries in the neighborhood table after the corresponding message got verified. However, low numbers of verifications will cause neighborhood tables to be (very) sparse. Thus, it can be expected that routing will suffer significantly from such an approach.

Thus, usage of a combination of multi-hop communication and VoD is not recommended.

Furthermore, all received messages, which a node wants to forward, have to be verified in advance to forwarding, independent from the used forwarding strategy. This is done to avoid creation of bogus channel load by an attacker [15], [21]. ETSI ITS does not change the signature of a forwarded message. For VoD two main cases have to be distinguished.

1) Verification of a node's position data for forwarding needs to verify a prior received and stored message. Only a central (i.e., cross layer) storage of full messages can avoid multiple verification of the same message by stor-

ing the verification result. However, such an architecture introduces an extra dependency of the network layer and the message storage.

2) Verification of a received and to be forwarded message splits again into two cases depending on relevance of the received message for the receiver.
   a) A receiver outside the message's relevance area only forwards it. Thus, the message is not stored in the LDM, as it never reaches the facility layer. It is only handled by lower layers up to the network layer. Hence, the network layer needs to cause verification by the security entity.
   b) A receiver inside the message's relevance area hands the message over to higher layers and forwards it. Thus, it gets stored in the LDM and the network layer causes its verification, as in case 2a. In case an application finds a message's data to be relevant, it can be used without further delay, as it has already been verified.

Using the LDM as the messages' storage, as suggested in [8], causes a cross layer dependency of network and facility layer as well as interactions of both entities with the security entity. Hence, separation of layers and uniqueness of responsibilities within the protocol stack suffers from such a design.

Instead we recommend a message storage within the cross layer security entity. It can provide a common interface for message verification for all protocol layers.

B. IPv6 over GeoNetworking

IPv6 over GeoNetworking is used to support IPv6 based communication with arbitrary higher level protocols over the dedicated VANET network layer. Such protocols use meta data, whose usage has to be preceded by message verification. However, the core aim of IPv6 over GeoNetworking is to use unchanged standard internet protocols. Thus, the VANET network layer has to ensure verification of all messages passed to an IPv6 interface. Hence, VoD is inappropriate for this kind of communication as every single message has to be verified to avoid attacks on higher level protocols or applications.

V. Interaction of Verify-on-Demand and Certificate Handling

Validation of a message is not limited to checking only its own signature. Moreover, the certificate (chain) used to secure the message needs to be verified, too. Within ETSI ITS there are at most two levels of unverified certificates, which are the pseudonym certificate (PSC) of the sending node and the authorization authority certificate (AAC) which is used to secure the PSC. The AAC is signed using a root certificate known to all nodes in the VANET.

The PSC is individual per node. Thus, a rapidly changing vehicular environment leads to reception of many different PSCs. Hence, a high number of verifications is required for a verify-all approach for PSCs. In contrast, the number of AACs can be expected to be small (see also Section VI). To avoid verification of certificates for messages which are never verified, VoD should be extended to certificates as well. Otherwise, the massive reduction of required verification capabilities as outlined in [6] cannot be reached, due to verification of many certificates.

Unfortunately, even certificates with valid format can become quite large [16], [22]. With a verify-all strategy only validated certificates are stored, but in case of VoD all unverified ones have to be stored, too. Thus, one has to take care that memory for storing PSCs does not become a system bottleneck in case of an attack. Moreover, the verification status has to be stored for each certificate to avoid multiple verifications.

A separate storage for unverified certificates is recommended in case the LDM design from [8] is used. Otherwise, the LDM would also need to keep track of inter-message dependencies of included parts of certificate chains, due to sporadic and on-demand inclusion of certificates [11], [12]. This would add a lot of complexity to the LDM, apart from increasing the interdependency of LDM and security entity.

One should note that this issue does not only affect ETSI ITS, but WAVE as well. For WAVE the situation is even worse, because the corresponding security standard does not limit the amount of hierarchy levels of the PKI system [23].

VI. Efficient Denial of Service Attack

Prior to the actual attack, the attacker stores valid PSCs, which he extracts from received messages. To carry out the attack, a stored PSC is added to the security envelope of a message generated by the attacker. Thereby, the content of the message is chosen in a way to be always regarded as relevant for the attacked vehicle. Relevance criteria can be easily obtained from the definitions of basic safety critical use cases [3]. Moreover, the attacker uses a different PSC, and thereby also different identifiers on all protocol layers, for each message. The messages' signatures consist of random data, as the private keys for the PSCs are unknown to the attacker.

The described attack, enables an attacker to achieve multiple goals at once. These include that for each sent message,

- the receiver regards the message as relevant for itself, which leads to
- message verification including
  1) verification of the formerly unknown PSC, which will succeed and lead to
  2) verification of the signature, which will fail.

Thus, each message sent by the attacker will lead to two computationally expensive verifications. If the attacker can send enough messages to supersede verification capabilities of receivers, he can block or at least delay verification of valid messages. This leads to a successful DOS attack on applications depending on data updates from received messages.

VoD schemes aim to massively limit the requirements for verification capabilities at receivers [6]. Thus, even a quite low number of faked messages, e.g., 10 per second, will exceed the provided capabilities. Thus, the attacker does not need dedicated equipment to jam the wireless channel or increase the channel load by misusing protocol features like described in [11], [12] to perform a DOS attack. Furthermore, the attacker may be able carry out the attack without a need

to violate legal regulations on usage patterns of the wireless channel reserved for VANET communication.

Moreover, the faked messages can be targeted to a dedicated node by using unicast communication at the network layer. Thereby, the attack will go unnoticed by the rest of the network. Both properties reduce the risk of the attacker to be detected and punished.

The attacker can use CAMs and/or DENMs for his attack. Thereby, usage of DENMs enables the attacker to attack all vehicles in the (freely selectable!) relevance area of the DENM. Verification before forwarding (see Section IV) limits the impact to the communication range of the attacker.

The changing identifiers used by the attacker disable simple countermeasures, like blocking of messages from senders after reception of multiple invalid messages. Disabling entire classes of messages, like DENMs containing a dedicated warning type, can only limit the attack if the attacker uses just the blocked dedicated type(s) of messages. However, the attacker could just send a mix of all possible DENMs. Thus, blocking of attacked message types would yield blocking all messages, which leads to a successful DOS attack, too.

The amount of AACs can be assumed to be highly limited. Otherwise, the attacker could send a valid certificate chain (e.g., PSC and AAC) with all elements being unknown to nodes. Thus, three verifications would be required for each message. Within WAVE the length of the certificate chain is not limited. Thus, the number of verifications required to validate a single message can be even higher. However, it can be assumed that the number of higher level certificates will be small in practice. Thus, an attacker cannot provide enough of them to enforce more than two verifications per message.

The described attack resembles the worst case scenario for a VoD scheme, as no verification can be spared. To counter the described attack, a system using VoD would need to have verification capabilities equal to a verify-all scheme. However, this clearly violates the objectives of the VoD design. Thus, the found DOS weakness puts the VoD design into question from a system robustness perspective.

VII. Conclusions and Future Work

Secured communication within VANETs is an important, but yet challenging issue. Reduction of the signature verification load within receivers by verify-on-demand (VoD) schemes is a popular method to limit performance requirements.

The provided analysis shows that VoD leads to a significant number of extra cross layer dependencies. Thus, overall complexity of VANET protocol stacks is increased and separation of dedicated communication layers is reduced. This holds especially for approaches which store to be verified messages within the facility layer LDM, e.g., [8]. Thus, we propose to instead use storage within the cross layer security entity.

Moreover, the amount of interfaces which have to be protected against malformed input from wireless attacks is massively increased by VoD. Finally, the effort for performing a successful DOS attack against dedicated nodes or groups of nodes is significantly reduced by introduction of VoD.

Our findings lead to the conclusion that usage of VoD in the currently proposed form is not recommended for VANETs. Instead approaches of verify-all schemes, like in [5], should be preferred. Moreover, future work could look for more computationally efficient algorithms for securing VANET messages.

References

[1] J. Harding, G. R. Powell, R. F. Yoon et al., "Vehicle-to-Vehicle Communications: Readiness of V2V Technology for Application," Washington, DC: National Highway Traffic Safety Administration, Tech. Rep. DOT HS 812 014, Aug. 2014.

[2] "Memorandum of Understanding for OEMs within the CAR 2 CAR Communication Consortium on Deployment Strategy for cooperative ITS in Europe," June 2011, V 4.0102.

[3] C. Campolo, A. Molinaro, and R. Scopigno, Eds., *Vehicular ad hoc Networks - Standards, Solutions, and Research*. Springer, Dec. 2015.

[4] T. Schütze, "Automotive Security: Cryptography for Car2X Communication," in *Embedded World Conference*, Mar. 2011, pp. 1–16.

[5] M. Knezevic, V. Nikov, and P. Rombouts, "Low-Latency ECDSA Signature Verification - A Road Towards Safer Traffic -," *IACR Cryptology ePrint Archive*, pp. 862 – 877, Oct. 2014.

[6] H. Krishnan and A. Weimerskirch, "Verify-on-Demand - a pracitcal and scalable Approach for Broadcast Authentication in Vehicle-to-Vehicle Communication," *SAE International Journal of Passenger Cars - Mechanical Systems*, vol. 4, no. 1, pp. 536–546, 2011.

[7] A. Weimerskirch, "V2X Security & Privacy: The Current State and Its Future," in *Proceedings 18th ITS World Congress*, Oct. 2011.

[8] E. Koenders, D. Oort, and K. Rozema, "An open Local Dynamic Map," in *Proceedings 10th ITS European Congress*, June 2014.

[9] C. Sommer and F. Dressler, *Vehiclular Networking*. Cambridge University Press, 2015.

[10] H. Füßler, J. Widmer, M. Käsemann, M. Mauve, and H. Hartenstein, "Contention-Based Forwarding for Mobile Ad Hoc Networks," *Elsevier's Ad Hoc Networks*, vol. 1, no. 4, pp. 351–369, Nov. 2003.

[11] S. Bittl, B. Aydinli, and K. Roscher, "Effective Certificate Distribution in ETSI ITS VANETs using Implicit and Explicit Requests," in *8th International Workshop Nets4Cars/Nets4Trains/Nets4Aircraft*, ser. LNCS 9066, M. Kassab et al., Ed., May 2015, pp. 72–83.

[12] S. Bittl and K. Roscher, "Efficient Authorization Authority Certificate Distribution in VANETs," in *2nd International Conference on Information Systems Security and Privacy*, Feb. 2016, pp. 85–96.

[13] N. Nowdehi and T. Olovsson, "Experiences from Implementing the ETSI ITS Secured Message Service," in *IEEE Intelligent Vehicles Symposium*, 2014, pp. 1055–1060.

[14] *Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service*, ETSI ES 302 637-2, Rev. V1.3.2, Nov. 2014.

[15] *Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 4: Geographical Addressing and Forwarding for Point-to-Point and Point-to-Multipoint Communications; Sub-part 1: Media-Independent Functionality*, ETSI ES 302 636-4-1, Rev. V1.2.1, July 2014.

[16] *Intelligent Transport Systems (ITS); Security; Security header and certificate formats*, ETSI TS 103 097, Rev. V1.2.1, June 2015.

[17] E. Whelan, "SNMP and Potential ASN.1 Vulnerabilities," CISSP, Tech. Rep., Dec. 2002.

[18] N. Cottin, "ASN.1 security issues," online: http://powerasn.ncottin.net/download/ASN1_SecurityIssues.pdf, Oct. 2007.

[19] Intel, "BERserk Vulnerability - Part 1: RSA signature forgery attack due to incorrect parsing of ASN.1 encoded DigestInfo in PKCS#1 v1.5," Intel Security: Advanced Threat Research, Tech. Rep., Sept. 2014.

[20] ——, "BERserk Vulnerability - Part 2: Certificate Forgery in Mozilla NSS," Intel Security: Advanced Threat Research, Tech. Rep., Oct. 2014.

[21] Buburuzan, T. et al., "Draft C2C-CC Standards System Profile," CAR 2 CAR Communication Consortium, Tech. Rep., Jan. 2014, V1.0.4.

[22] S. Bittl, K. Roscher, and A. A. Gonzalez, "Security Overhead and its Impact in VANETs," in *8th IFIP Wireless Mobile Networking Conference*, Oct. 2015.

[23] *IEEE Standard for Wireless Access in Vehicular Environments - Security Services for Applications and Management Messages*, Intelligent Transportation Systems Commitee of the IEEE Vehicular Technology Society Std. P1609.2, Rev. 2013, Apr. 2013.

Physical Layer-Based Message Authentication in VANETs

Ala'a Al-Momani
Institute of Distributed Systems
Ulm University, Germany
Email: alaa.al-momani@uni-ulm.de

Frank Kargl
Institute of Distributed Systems
Ulm University, Germany
Email: frank.kargl@uni-ulm.de

Christian Waldschmidt
Institute of Microwave Techniques
Ulm University, Germany
Email: christian.waldschmidt@uni-ulm.de

***Abstract*—Authenticating legitimate nodes is a major concern of the envisioned vehicular networks. To achieve this, standards and literature propose to use asymmetric cryptographic mechanisms which generate significant overheads in terms of time and power consumption. In this paper, we address this problem and we propose a novel idea of exploiting physical layer characteristics to rely on them for re-authenticating future beacons after verifying the first one cryptographically. Despite the challenges in such high mobility networks, possible concrete approaches to start the evaluation of our scheme are presented. Our approaches are inspired by the vehicular channel related work conclusions which give signs of future success to our scheme in this critical field.**

I. INTRODUCTION

As pointed out in [1], in contradiction with other networks, securing vehicular networks at the physical layer has been disregarded. The reason of this is the nature of such networks by means of the high mobility of vehicles, which makes investigating physical layer and exploiting its information for the sake of security a challenging task. In standards, the foreseen authentication process states that digital signatures in addition to certificates have to be generated and attached to messages by senders. The digital signatures are based on Elliptic Curve Digital Signature Algorithm (ECDSA) [2] where each vehicle is equipped with a public and a secret key; the secret key is used to sign all outgoing messages while the public key is amended with some other attributes forming the certificate of the vehicle. Receivers then check signatures and certificates of messages for correctness and decide whether they are originated from legitimate senders or would-be intruders, e.g. roadside attackers.

In order to inform neighbouring vehicles about the current state, vehicles have to keep sending their states periodically with a frequency of 1-10 Hz including the position, the speed, the heading and other similar information in messages called "beacons". For each beacon, the procedure for generating and attaching signatures needs to be executed at the sender side. In addition, receivers have to keep verifying newly received beacons even if they are from the same sender. This whole procedure to ensure that only legitimate vehicles are able to exchange messages among each others in the network creates significant overheads due to the complex cryptographic calculations resulting in major drawbacks that should be highlighted. The following gives a summary of these shortcomings::

- The decrease of bandwidth utilization due to the increase in message size that is necessary because of the must-included signature.
- The increase in packet collisions due to the increase of the number of packets per message in an already congested channel.
- The increase in the end-to-end delay because of:
  - The increase in the required time to generate signatures.
  - The increase of the transmission delay due to the need for transmitting additional bits.
  - The increase in the required time due to the verification process of signatures at the receivers.

In this paper, we restrict our concerns to the last issue that points out the long time needed to verify the ECDSA-based signatures. We propose the idea of a novel scheme for re-authenticating the periodic messages, i.e. beacons, in vehicular networks. By exploiting the unique physical layer features between a specific sender and a specific receiver, we aim to eliminate the drawbacks of the classical authentication mechanisms proposed in the standards of vehicular networks while maintaining a reasonable degree of security.

We also provide concrete approaches to start the evaluation of our scheme inspired by the outcome from propagation studies in realistic vehicular environments which gives signs of success to our proposed scheme. Moreover, we propose the idea of multi-factor authentication with the use of subjective-logic [3].

II. RELATED WORK

The previous authentication process takes place in the upper layers of the OSI model. Researchers have realized its drawbacks and pointed them out, for example, [4; 5]. In order to mitigate them, some researchers [5] suggested to use hardware secure modules, but equipping vehicles with sophisticated processing units adds additional cost. Others suggested to omit signatures and certificates to reduce the introduced latency in order to achieve reasonable efficiency for the critical applications [6], but this will lead to insecure networks where unauthenticated nodes are able to spoof and inject messages into networks.

A. Security at the Physical Layer

The previous shortcomings in the upper layers′ security motivate researchers in other wireless communication networks to look for other solutions. They suggested to integrate the physical layer into the authentication process in static and low mobility networks [7; 8; 9; 10; 11; 12].

Mainly, the work in exploiting physical layer features for the use of security and authentication can be divided into two classes:

- Class 1: Extracting the secret key from the common wireless channel between the transmitter and the receiver, e.g. [9; 10; 11].
- Class 2: Fingerprinting the wireless channel established between the transmitter and the receiver, e.g. [7; 8; 12].

Both of them are based on the fact that the wireless channel established between a specific transmitter and a specific receiver is unique and only known to both of them. The first class uses the unique channel variation to establish the secret key. This approach is considered secured such that only the transmitter and the receiver are able to construct the key. However, the verification process still has to be applied in this case, which means that the long verification time still exists. On the other hand, the second class does not require any key extraction or verification. It relies on the uniqueness of the frequency response for each transmitter-receiver pair, and hence a receiver identifies a transmitter based on the history it has for that transmitter. This way, the need of the signature verification process vanishes with its drawbacks while maintaining a reasonable degree of security.

At 5 GHz, over a span o 10 MHz, with indoor stationary user terminals, Xiao et al. [7; 8] proposed ways to exploit spatial variability of the frequency response. They found that variations are strongly correlated in time while very weakly correlated in space giving a positive impact on performance in such a static scenario. In addition, they concluded that channel time variations can improve performance whereas frequency correlation degrades it. Their results show that it is possible to distinguish between legitimate nodes and other illegitimate nodes based on the corresponding physical layer characteristics to each one of them.

B. Vehicular Channel Propagation Models

The distinct features of vehicular networks arise from the nature of the rapidly changing topology due to vehicles′ rapid movement. This results in a significant uniqueness in the statistical characteristics of the multipath propagation in V2V communication compared to other indoor or even cellular communication. The investigation of vehicular channel characterization is fairly young research topic [13]. It gained researchers′ interest when WAVE initiative and other vehicular applications raised concerns regarding vehicular communication. Before 2006, V2V channel characteristics were rarely investigated, e.g. [14] [15]. Since 2006, there has been a lot of research, e.g. [16; 17; 18; 19; 20; 21; 22; 23; 24], addressing the vehicular channel propagation models based on measurement campaigns considering different frequency bands.

One of the earlier works on V2V channel investigation at the 5.9 GHz band is [25], where the Tapped Delay Line (TDL) approach was used to model the channel. They stated that these kind of channels are doubly selective, in other words, they are both time- and frequency-selective channels. Also at 5.9 GHz, Cheng et al. [21] conducted a measurement study on V2V narrow-band channel; they presented a single- and dual-slope large-scale path loss model with Nakagami distribution to describe the small-scale fading. In addition, they introduced the Speed-Separation (S-S) diagram, which is a new tool for understanding and estimating Doppler spread and coherence time.. Kunisch and Pamp did a measurement experiment in [24] at 5.9 GHz over a span of 20 MHz, in which they extracted the scattering function which contains information about Doppler spread and path delays. They also provided a good explanation of each propagation path scenario.

Other measurement studies were conducted at different bands. In [16], Karedal et al. were able to track individual propagation paths at 5.2 GHz. Paier et al. [17] investigated pathloss, Power Delay Profiles (PDPs), and Delay-Doppler spectra from a highway measurement over 240 MHz at 5.2GHz where the transmitter and the receiver vehicles travelled in opposite direction. Examples of other studies at different bands are [26] at 5.3 GHz, and [27] at 5.6 GHz.

III. Physical Layer-Based Message Authentication

A. Requirements

For our scheme to work, physical layer characteristics have to meet some requirements in order to be able to rely on them for future verification of beacons without checking the signatures. Requirements for these characteristics include:

- Stability: The characteristics should show stability over at least two consecutive beacons. In other words, the time correlation of this specific physical layer characteristic should be high enough to ensure its stability within the reception of two consecutive beacons.
- Uniqueness: to allow discrimination among several transmitters in the vehicular network, the uniqueness of the physical layer characteristic among them has to be ensured. This means that the characteristic the receiver relies on has to be spatially uncorrelated to be able to distinguish between several transmitters at different locations at the same time. However, due to vehicles movement, this characteristic should allow a degree of spatial correlation such that the measured value of the characteristic when vehicle A at location X is correlated with the measured value when the same vehicle is at location X'. A noteworthy point in this regard is that vehicles do not move in random paths, but in deterministic tracks where prediction of movement could be easily employed.
- Measurable: receivers need to be able to observe and measure the specific characteristics.
- Unspoofable: attackers should not be able to spoof the characteristics, luring receivers into accepting false messages.

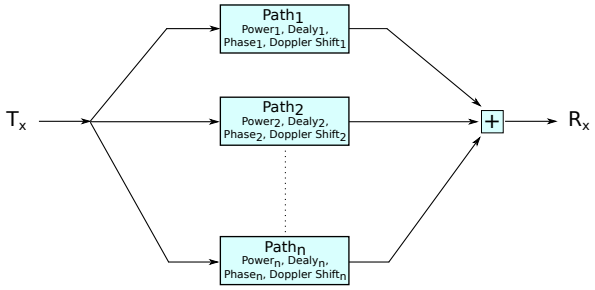


Fig. 1. Multipath Model

B. The Proposed Scheme

We take a slightly different approach than the classical fingerprint approach discussed earlier. Our new approach considers (re)authentication of earlier communication partners by characteristics of the communication channel. It is based on the observation that a radio channel between transmitter and receiver has a characteristic with a unique signature (for example defined by multi-path propagation, Doppler shifts...), which is hard for an attacker to guess or manipulate. Hence, instead of looking at the frequency response, as it may be challenging in high mobility scenarios, looking at the individual multipath components and extracting their characteristics will provide more robustness. Such a case is shown in Figure 1 where each contribution has its own delay, power, phase, and Doppler shift.

The conducted measurement studies showed that each average PDP consists of several identifiable contributions and that they are presented over several consecutive time instants (typically in order of seconds) [27]. This was the observation in [26] and [16] as well. Hence, periodic beacons could be authenticated based on this channel signature. For this purpose, a first beacon would be authenticated by means of classical cryptography, establishing an initial trust anchor.

As long as the channel charachteristic remains sufficiently stable between this and a consecutive beacon, all subsequent beacons could now be authenticated by the means of their channel signature associated with the original transmitter. Costly cryptographic verification processes may potentially be skipped for some beacons.

The process is exemplified in Figure 2. A transmitter T sends periodic messages. The first message has to be cryptographically verified in any case in order to produce an initial trust anchor. Thereafter, messages are only verified cryptographically if the receiver trust at A or B in the message being delivered over the same channel falls below a certain threshold. Receiver A needs to cryptographically verify the third beacon while beacons 3 and 4 need to be verified for receiver B′s case.

C. Multi-factor Authentication and Possible Enhancement

Multi-factor Authentication can play a role in the authentication process where the receiver can rely on different

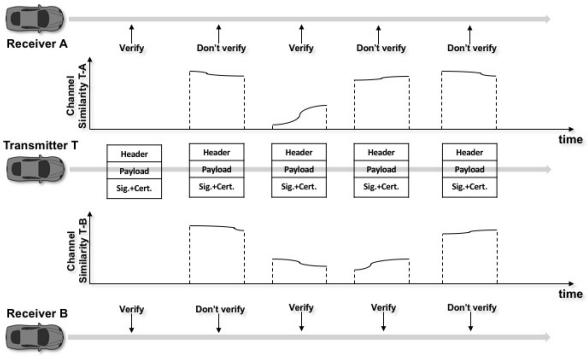


Fig. 2. Physical Layer Based Message Authentication

observations to form his opinion about the received beacon whether it is originated from a legitimate node or not. In order to avoid resorting to the cryptographic verification to rebuild the overall confidence in the latter beacons, the receiver may combine the output from our proposed scheme with another lightweight authentication mechanism. This requires a sufficient degree of flexibility of the outputs to be combined together and allow multi-factor authentication process. Hence, we here foresee the use of subjective logic to form a holistic framework for such authentication mechanisms where each factor (e.g. multipath characteristics, lightweight authentication mechanism, etc...) gives an opinion about the received beacon whether it is generated from the same transmitter as the previously received beacons or not. Subjective logic extends the classical logic theory by introducing a degree of certainty to each opinion. It has a wide set of operators allowing the fusion of individual factor opinions into one opinion taking the degree of certainty of each output into account. It has been deployed in VANETs in [28] to form a misbehaviour detection framework which our proposed scheme could be integrated into.

IV. CONCLUSION

This paper addresses the shortcomings of using ECDSA-based signatures in the envisioned V2V communication to achieve a proper authentication of the periodic beacons. The main drawback of such signatures is the long verification time needed to verify the signature by the receiver. We proposed a novel idea of integrating physical layer into the authentication process aiming at eliminating the shortcomings of the upper layer authentication mechanisms. Inspired by related work in vehicular channel propagation models outcome, which gives signs of future success for our scheme, we proposed concrete approaches to start the evaluation of the proposed scheme, in addition to proposing possible ways of enhancements including multi-factor authentication using subjective logic.

REFERENCES

[1] A. Al-Momani, F. Kargl, C. Waldschmidt, S. Moser, and F. Slomka, "Wireless channel-based message authentication," in *Vehicular Networking Conference (VNC), 2015 IEEE*. IEEE, 2015, pp. 271–274.

[2] ANSI, "Public key cryptography for the financial services industry: The elliptic curve digital signature algorithm," no. ANSI X9.62, 1998.

[3] A. Josang, "Artificial reasoning with subjective logic," in *the second australian workshop on Commonsense Reasoning*, vol. 48, 1997, p. 34.

[4] M. Raya and J.-P. Hubaux, "Securing vehicular ad hoc networks," *Journal of Computer Security*, vol. 15, no. 1, pp. 39–68, 2007.

[5] P. Papadimitratos, L. Buttyan, T. S. Holczer, E. Schoch, J. Freudiger, M. Raya, Z. Ma, F. Kargl, A. Kung, and J.-P. Hubaux, "Secure vehicular communication systems: design and architecture," *Communications Magazine, IEEE*, vol. 46, no. 11, pp. 100–109, 2008.

[6] E. Schoch and F. Kargl, "On the efficiency of secure beaconing in vanets," in *Proceedings of the third ACM conference on Wireless network security*. ACM, 2010, pp. 111–116.

[7] L. Xiao, L. J. Greenstein, N. B. Mandayam, and W. Trappe, "Using the physical layer for wireless authentication in time-variant channels," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 7, pp. 2571–2579, 2008.

[8] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Fingerprints in the ether: Using the physical layer for wireless authentication," in *Communications, 2007. ICC'07. IEEE International Conference on*. IEEE, 2007, pp. 4646–4651.

[9] S. Jana, S. N. Premnath, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "On the effectiveness of secret key extraction from wireless signal strength in real environments," in *Proceedings of the 15th annual international conference on Mobile computing and networking*. ACM, 2009, pp. 321–332.

[10] C. Chen and M. A. Jensen, "Secret key establishment using temporally and spatially correlated wireless channel coefficients," *Mobile Computing, IEEE Transactions on*, vol. 10, no. 2, pp. 205–215, 2011.

[11] N. Patwari, J. Croft, S. Jana, and S. K. Kasera, "High-rate uncorrelated bit extraction for shared secret key generation from channel measurements," *Mobile Computing, IEEE Transactions on*, vol. 9, no. 1, pp. 17–30, 2010.

[12] I. O. Kennedy and A. M. Kuzminskiy, "Rf fingerprint detection in a wireless multipath channel," in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*. IEEE, 2010, pp. 820–823.

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

[14] A. S. Akki and F. Haber, "A statistical model of mobile-to-mobile land communication channel," *Vehicular Technology, IEEE Transactions on*, vol. 35, no. 1, pp. 2–7, 1986.

[15] J. Maurer, T. Fugen, and W. Wiesbeck, "Narrow-band measurement and analysis of the inter-vehicle transmission channel at 5.2 ghz," in *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, vol. 3. IEEE, 2002, pp. 1274–1278.

[16] J. Karedal, F. Tufvesson, N. Czink, A. Paier, C. Dumard, T. Zemen, C. F. Mecklenbräuker, and A. F. Molisch, "A geometry-based stochastic mimo model for vehicle-to-vehicle communications," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 7, pp. 3646–3657, 2009.

[17] A. Paier, J. Karedal, N. Czink, C. Dumard, T. Zemen, F. Tufvesson, A. F. Molisch, and C. F. Mecklenbräuker, "Characterization of vehicle-to-vehicle radio channels from measurements at 5.2 ghz," *Wireless personal communications*, vol. 50, no. 1, pp. 19–32, 2009.

[18] R. Meireles, M. Boban, P. Steenkiste, O. Tonguz, and J. Barros, "Experimental study on the impact of vehicular obstructions in vanets," in *Vehicular Networking Conference (VNC), 2010 IEEE*. IEEE, 2010, pp. 338–345.

[19] S. A. H. Tabatabaei, M. Fleury, N. N. Qadri, and M. Ghanbari, "Improving propagation modeling in urban environments for vehicular ad hoc networks," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 3, pp. 705–716, 2011.

[20] Y. Jeong, J. W. Chong, H. Shin, and M. Z. Win, "Intervehicle communication: Cox-fox modeling," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 9, pp. 418–433, 2013.

[21] L. Cheng, B. E. Henty, D. D. Stancil, F. Bai, and P. Mudalige, "Mobile vehicle-to-vehicle narrow-band channel measurement and characterization of the 5.9 ghz dedicated short range communication (dsrc) frequency band," *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 8, pp. 1501–1516, 2007.

[22] N. Akhtar, S. C. Ergen, and O. Ozkasap, "Vehicle mobility and communication channel models for realistic and efficient highway vanet simulation," *Vehicular Technology, IEEE Transactions on*, vol. 64, no. 1, pp. 248–262, 2015.

[23] I. Sen and D. W. Matolak, "Vehicle–vehicle channel models for the 5-ghz band," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 9, no. 2, pp. 235–245, 2008.

[24] J. Kunisch and J. Pamp, "Wideband car-to-car radio channel measurements and model at 5.9 ghz," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*. IEEE, 2008, pp. 1–5.

[25] G. Acosta-Marum and M. Ingram, "Doubly selective vehicle-to-vehicle channel measurements and modeling at 5.9 ghz," in *Proc. Int. Symp. Wireless Personal Multimedia Commun*. Citeseer, 2006.

[26] O. Renaudin, V.-M. Kolmonen, P. Vainikainen, and C. Oestges, "Wideband mimo car-to-car radio channel measurements at 5.3 ghz," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*. IEEE, 2008, pp. 1–5.

[27] J. Karedal, F. Tufvesson, T. Abbas, O. Klemp, A. Paier, L. Bernadó, and A. F. Molisch, "Radio channel measurements at street intersections for vehicle-to-vehicle safety applications," in *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. IEEE, 2010, pp. 1–5.

[28] R. van der Heijden, S. Dietzel, and F. Kargl, "Misbehavior Detection in Vehicular Ad-hoc Networks," in *1st GI/ITG KuVS Fachgespräch Inter-Vehicle Communication (FG-IVC 2013)*, Innsbruck, Austria, February 2013.

Quantifying the Influence of Route Choice Optimization on Emissions and Fuel Consumption

Daniel Cagara*, Björn Scheuermann†
Humboldt University of Berlin, Germany
Email: *cagarada@informatik.hu-berlin.de, †scheuerb@informatik.hu-berlin.de

***Abstract*—Genetic algorithm-based route choice optimization techniques have recently demonstrated their potential to reduce the travel costs on the system level. In this context, the costs are reduced according to exactly one metric such as, e. g., the sum of all individual travel times. In the present study, we will investigate the tradeoff between the saved travel time and the additional costs in terms of higher emissions and a higher fuel consumption.**

I. Introduction

Intelligent transportation system technologies (ITS) [6] play a key role in the improvement of the efficiency of road networks in terms of their utilization. Typically used are in-vehicle systems that use either vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) communication [4] in consort with traffic information systems [9], [7], [8] to collaboratively optimize the route choices in the network. These intelligent solutions can help to balance the traffic across the network in order to avoid critical conditions such as traffic jams. But aside from showing that these approaches in fact do improve the road network efficiency, it can be worth to take a look at the environmental impacts they come with. In this paper, we focus on the evaluation and quantification of these impacts in terms of CO_2 (carbon dioxide) emissions as well as the fuel consumption. To this end, we consider the recently proposed concept of route choice optimization using distributed Genetic Algorithms [3].

II. Evaluation

Realistic vehicle movement plays a very important role in the evaluation of systems related to vehicular road traffic. For the evaluation in this paper, we therefore use a real world scenario from the city of Bologna [2] which represents an observed peak-hour traffic demand (8:00 am – 9:00 am). During this time period, a total of $11,000$ vehicles is inserted into the road network. A warm-up period of $15\mathrm{min}$ is bypassed to allow traffic load in the road network to stabilize before the evaluation starts. Also the last $15\mathrm{min}$ are not accounted for in this evaluation as the traffic load starts decreasing. The evaluation itself is performed in SUMO [1], a microscopic vehicular road traffic simulator.

We compare two sets of *route choices* in the given scenario. A set of route choices in essence describes which route *each* of the $11,000$ vehicles drives from his origin to his destination. On the one hand we have the *unoptimized* route choices: every driver chooses the shortest path (in terms of the free-flow travel time) from his origin to his destination. We compare this to *optimized* route choices on the other hand. These are optimized using the techniques described in our recent publication [3]. It aims at improving the costs for the system as a whole. In this context, the sum of the vehicles' travel times is used as the target function for a distributed on-line optimization process.

SUMO comes equipped with the ability to record emissions and fuel consumption according to the HBEFA v2.1 model [5]. In this context, the emissions are recorded in $\frac{\mathrm{mg}}{\mathrm{s}}$ and the fuel consumption in $\frac{\mathrm{ml}}{\mathrm{s}}$ in each timestep for each car. Additionally, the cars' total emission and fuel consumption is recorded after they have finished their journey. The two metrics are closely related, but not fully equivalent, because different classes of vehicles use different types of fuel, which in turn correspond to different amounts of CO_2 per litre of fuel [5].

It can be observed that the route choice optimization increases the total emissions of CO_2 from $2605.43\,\mathrm{kg}$ to $3000.96\,\mathrm{kg}$ ($+15.18$ %) and the total fuel consumption from $1038.74\,\mathrm{l}$ to $1196.36\,\mathrm{l}$ ($+15.17$ %), while the total driven distance increases from $13\,075.66\,\mathrm{km}$ to $15\,977.07\,\mathrm{km}$ ($+22.18$ %). However, due to the optimization, a different set of vehicles will reach their destination during the simulated time frame. More precisely, 8556 vehicles arrive at their destination before the route choice optimization is applied. Using the optimized route choices this number grows to 9369. Therefore, the absolute fuel consumption or CO_2 emission do not constitute suitable metrics: they do not refer to the resources spent for achieving the same goal or delivering the same amount of service.

We therefore compare the additional expenses of the optimization in terms of higher CO_2 emissions and a higher fuel consumption per kilometer of the *unoptimized* route. That is, in a sense, we use the unoptimized (i. e., shortest path) route length as a measure for the "amount" of transportation service delivered to the driver of the respective car. Let r_c be the route of car c as it was planned in the unoptimized scenario and l_{r_c} the length of this route. Let $\mathrm{CO_2}(r_c)$ and $\mathrm{FUEL}(r_c)$ denote the CO_2 emission and the fuel consumption along that route, respectively. Then, given the CO_2 emission $\mathrm{CO_2}(\widehat{r_c})$ and the fuel consumption $\mathrm{FUEL}(\widehat{r_c})$ in the optimized scenario, we normalize these values to the length of the unoptimized route l_{r_c}. Furthermore, we are interested the in the ratios $\frac{\mathrm{FUEL}(\widehat{r_c})}{\mathrm{FUEL}(r_c)}$ and $\frac{\mathrm{CO_2}(\widehat{r_c})}{\mathrm{CO_2}(r_c)}$ for all cars that finish their journey in both the unoptimized and the optimized case. These ratios show by

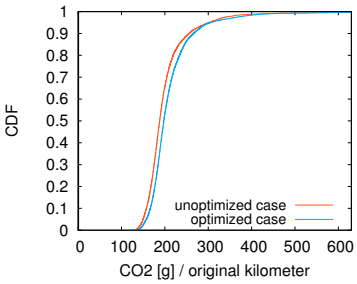


Fig. 1. Comparison of CO_2 emissions.

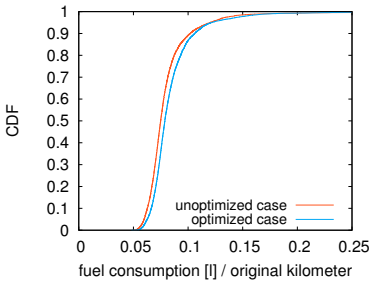


Fig. 2. Comparison of the fuel consumption.

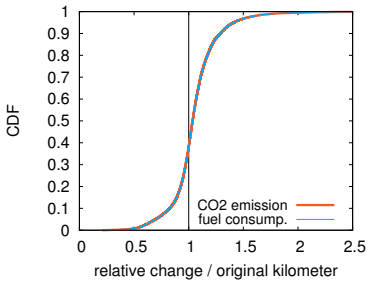


Fig. 3. Relative change after optimization.

which factor the CO_2 emission and the fuel consumption of a vehicle driving from the same origin to the same destination have changed.

Figure 1 shows the cumulative distribution function (CDF) of the CO_2 emissions per kilometer before and after the optimization of the route choices, as discussed before in both cases in relation to the length of the unoptimized, shortest-path route. It can be seen that the values do not change much: only a small subset of vehicles experience noticeably higher emissions and a higher fuel consumption. Also, only a small number of vehicles can lower these values to a non-negligible extent. The same pattern can be observed when looking at the comparison of the absolute fuel consumption in Figure 2.

While the CDFs of the absolute values show that the distribution of emissions and fuel consumption do not change much, it does not show what that means from the drivers' point of view. Figure 3 shows the CDF of the relative changes in the absolute CO_2 emissions and fuel consumption per original kilometer after the optimization has been applied. Here, it can be seen that, around $40\,\%$ of all vehicles can reduce their CO_2 emission and fuel consumption noticeably while the other $60\,\%$ experience increased values. This indicates that disadvantages (in terms of a non-optimal emission / fuel consumption) are reallocated among the cars: we still have cars that cause much pollution, its just that these are different cars after the optimization of the route choices. In a next step, we take a look at the same values, but now accumulated over all routes. It can be seen that the total emissions of CO_2 have increased from $199.25\,\mathrm{g}$ to $207.33\,\mathrm{g}$ and the fuel consumption has increased from $79.44\,\mathrm{ml}$ to $82.65\,\mathrm{ml}$ per kilometer. This corresponds to a $4.05\,\%$ increase of CO_2 emission and an increase of $4.04\,\%$ in the fuel consumption for the entire system. The observed increase in emissions and fuel consumption after applying our proposed route choice optimization methodology is relatively small given the fact that the costs for the system as a whole (in terms of the total travel time) can be reduced by over $20\,\%$ [3].

III. Conclusion

In this abstract, we focused on the quantification of the impact on both the CO_2 (carbon dioxide) emissions as well as the fuel consumption which is caused by the route choice optimization using the on-line route optimization techniques described in [3]. In this context we evaluated the increase in emissions and fuel consumptions for the new (optimized) route choices per kilometer of the route as it was driven in the scenario with unoptimized route choices. It could be shown that the system as a whole experiences an $4.05\,\%$ increase of CO_2 emission and an increase of $4.06\,\%$ in the fuel consumption per original kilometer. At the same time, the cost (in terms of the total travel time) for the entire system was reduced by over $20\,\%$. While from the ecological point of view the route choice optimization leads to poorer results, the environmental overhead is reasonably low compared to the benefit that can be achieved by the drivers in terms of a lower travel time.

References

[1] M. Behrisch, L. Bieker, J. Erdmann, and D. Krajzewicz. SUMO – simulation of urban mobility: An overview. In *SIMUL '11: Proceedings of the Third International Conference on Advances in System Simulation*, Oct. 2011.

[2] L. Bieker, D. Krajzewicz, A. Morra, C. Michelacci, and F. Cartolano. Traffic simulation for all: a real world traffic scenario from the city of Bologna. In *SUMO '14: Proceedings of the Conference on Modeling Mobility with Open Data*. May 2014.

[3] D. Cagara, B. Scheuermann, and A. L. Bazzan. Traffic optimization on islands. In *VNC '15: Proceedings of the IEEE Vehicular Networking Conference 2015*, Dec. 2015.

[4] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil. Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions. *Communications Surveys & Tutorials, IEEE*, 13(4), 2011.

[5] M. Keller and P. de Haan. Dokumentation hbefa 2.1. *Bern/Essen/Graz/Heidelberg, INFRAS*, 2004.

[6] P. Papadimitratos, A. La Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza. Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation. *Communications Magazine, IEEE*, 47(11), 2009.

[7] J. Rybicki, B. Scheuermann, W. Kiess, C. Lochert, P. Fallahi, and M. Mauve. Challenge: Peers on wheels – a road to new traffic information systems. In *MobiCom '07: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, pages 215–221, Sept. 2007.

[8] J. Rybicki, B. Scheuermann, M. Koegel, and M. Mauve. Peertis: A peer-to-peer traffic information system. In *VANET '09: Proceedings of the 6th ACM International Workshop on VehiculAr Inter-NETworking*, pages 23–32, Sept. 2009.

[9] L. Wischoff, A. Ebner, H. Rohling, M. Lott, and R. Halfmann. Sotis-a self-organizing traffic information system. In *VTC '03-Spring: Proceedings of the 57th IEEE Vehicular Technology Conference*, pages 2442–2446, Apr. 2003.

Byzantine Consensus in Vehicle Platooning via Inter-Vehicle Communication

Martin Wegner, Wenbo Xu, Rüdiger Kapitza, Lars Wolf
Institute of Operating Systems and Computer Networks
Technische Universität Braunschweig, Germany
Email: (wegner|wxu|kapitza|wolf)@ibr.cs.tu-bs.de

***Abstract*—Cooperative driving and platooning have gained a growing focus recently. Letting vehicles to reach a consensus and make a joint decision is necessary for some applications. To address this problem, we propose a novel consensus protocol named BFT-ARM that fits real sensor values and can tolerate $t(< n/3)$ Byzantine nodes out of n. BFT-ARM guarantees that the decision is close to the median of all good nodes. We also present the simulation framework ArteryLTE to evaluate our protocol.** [1]

***Keywords*—Cooperative driving, Byzantine consensus, Inter-Vehicle Communication**

I. Introduction

In recent years, an increasing amount of Advanced Driver Assistance Systems (ADASs) can be seen in automobiles. This also includes the development of cooperative driving functions. Many of these applications can be improved by exchanging data via Inter-Vehicle Communication (IVC) to improve safety, resource usage, energy efficiency and driving experience [1]. One example of cooperative driving is *platooning*, where a group of vehicles can follow each other automatically and keep the optimal distance among each other [2, 3]. Apart from following passively, there are also a number of useful applications which first need to agree on a common value for a cooperative decision. For example, in certain application scenarios vehicles will need to detect the traffic condition or weather condition in their surroundings to adjust their operations to, or to calculate the best route according each vehicle's own navigation device, or to set a preferred speed for the cruise control, etc. The common properties of such applications are: 1) The value is required to be agreed among all vehicles. 2) The value to be agreed upon can be measured individually by each vehicle. 3) Even if some faulty vehicles do not follow the common decision, the safety of others is not violated.

We also consider the existence of faulty or malicious nodes in the group. Faulty behaviours include not only crash faults but also arbitrary faults like bit flip or providing inaccurate, inconsistent and even malicious values. All these faults are referred to as Byzantine faults [4]. The Byzantine consensus protocol is aiming at achieving consensus among all correct participants, despite of a limited number of faulty nodes. Examples of such protocols can be found in [5, 6].

Most Byzantine consensus problems assume the value domain is discrete and limited, for instance the binary consensus [5] or multi-value consensus [7]. However, in automobile applications, especially those involving sensor values, the values can be continuous and "smooth" e. g. speed, distance, temperature, etc.

In this work, we designed a new consensus protocol for the continuous value domain in vehicle platooning named BFT-ARM (Byzantine Fault Tolerant and Asynchronous Real-value consensus with Median validity). We also built a simulation framework based on our previous work, and will use it to evaluate the consensus protocol.

The paper is organized as follows. Section II discusses some related work. Section III defines the system model and problem. Section IV presents the design of BFT-ARM and section V introduces the evaluation framework. Section VI concludes the paper.

II. Related Work

In the *Byzantine consensus problem*, each node has an input value and try to make a consensus on one of them, and there are a limited number of Byzantine nodes [5]. An important aspect of Byzantine consensus problem is how to define the validity of the agreed value. There are different opinions from different viewpoints. E.g., Neiger distinguishes the *Validity* and *Strong Validity* [8]. The former requires that if all correct nodes have the same input value, they will decide on that value, but does not guarantee anything when the input values are different. And Strong Validity requires that the decided value comes from a correct node. Then Neiger proves that achieving Strong Validity requires at least $t \cdot |\mathcal{D}|$ nodes, where t is the maximum tolerable Byzantine nodes and $\mathcal{D}$ is the domain of the input values. This means a tremendous number of nodes are necessary when the input value domain is large.

A recent work of Stolz and Wattenhofer proposes a weaker requirement compared to the strong validity, called *median validity* [9]. It only requires the agreed value is close to the median of all good nodes. This is especially useful

[1]This work is part of the DFG Research Unit *Controlling Concurrent Change*, funding number FOR 1800.

with a continuous value domain. However, their protocol assumes a synchronous communication where message transmission time has a known upper bound. This assumption is not applicable in IVC scenario. So we designed the new BFT-ARM protocol to achieve the median validity.

There are also some other work from the viewpoint of control theory to manage platooning via consensus approach [10]. This is useful for another family of applications like instant speed and distance control, which is an orthogonal direction to our work.

III. System Model and Problem Statement

System Model: The platooning consists of n vehicles, or nodes as abstraction: $\{p_1, p_2, \ldots, p_n\}$. Every node has an input value $x_i \in \mathbb{R}$ [2], e.g. from its sensor or configuration. A node is called *correct* if 1) its input value is correct and 2) it exactly follows the protocol. Among all the nodes up to t ($< n/3$) nodes can be *faulty*, meaning that they can behave arbitrarily such as take an incorrect value from a malfunctioning sensor or not follow the protocol.

Consensus problem: BFT-ARM achieves consensus on a value $v \in \mathbb{R}$ satisfying the following conditions:

- *Agreement*: No two correct nodes decide differently.
- *Termination*: Every correct node eventually decide.
- *Validity*: The decided value of correct nodes v is *valid* (see definition below).

Inspired by the work of [9], the validity is defined in the following way. Assuming there are actually f ($\leqslant t$) faulty nodes during runtime (not known by the consensus algorithm). Let SG be the sorted array of input values of all good nodes (the index starts from 0). Then $SG[\lceil \frac{n-f}{2} \rceil - 1]$ represents the median value of SG.

Definition 1. *Validity: a decision v is valid, if*

$$SG[\lceil \frac{n-f}{2} \rceil - 1 - t] \leqslant v \leqslant SG[\lceil \frac{n-f}{2} \rceil - 1 + t] \quad (1)$$

In other words, a valid value is the one within the range of the middle $(2t+1)$ correct nodes.

Network: Nodes communicate via messages. The network is asynchronous. That means messages can experience an unbounded delay, get lost, duplicated or corrupted. However an eventually synchronous connection is required to overcome the FLP impossibility [11]. BFT-ARM does not rely on the synchrony to achieve agreement. So termination needs an eventually synchronous connection.

Digital Signature and Trusted Subsystem: The messages are signed with digital signatures. A message m signed by a node i is notated as $\langle m \rangle_{\sigma_i}$. We also assume that every node possesses a trusted subsystem for message authentication and verification with a monotonic counter. The value of the counter can be used to authenticate some special messages, and increases by 1 after that. The partner who receives one of these messages can also verify that this message has the valid counter value without any gaps to the previous ones. Examples of such subsystem applied in Byzantine fault tolerant systems can be found in [12, 13]. We assume that faulty nodes cannot break the digital signature mechanism nor the trusted subsystem.

[2]In practice, the value space is still a finite set limited by the platform. We do not discuss Turing uncomputable numbers or real computation here.

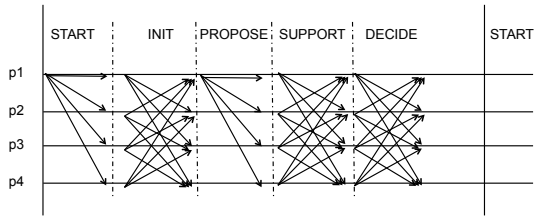


Figure 1. BFT-ARM in normal case.

IV. BFT-ARM Design

A. Normal case operation

The normal case protocol is illustrated in Figure 1.

It can be divided into 6 steps:

1) The leader p_i periodically activates a consensus request with a broadcast $\langle \text{START}, seq, p_i \rangle_{\sigma_i}$. seq is a sequence number generated by the trusted counter.
2) Upon received START message, each node p_j firstly verifies the sequence number. If it is a valid sequence number, it broadcasts (including to itself) with its input value in $\langle \text{INIT}, seq, p_j, x_j \rangle_{\sigma_j}$.
3) Upon the leader received $(n-t)$ values (including itself), it sorts the received values and picks the median value v_{med}. Then it proposes v_{med} together with the $(n-t)$ original signed INIT messages attached as a certificate $\vec{cm}$. Namely: $\langle \text{PROPOSE}, seq, p_i, v_{med}, \vec{cm} \rangle_{\sigma_i}$.
4) Upon a node p_j received the PROPOSE message, it verifies that v_{med} is really the median of all the values in $\vec{cm}$. If so, it broadcasts $\langle \text{SUPPORT}, seq, p_j, v_{med} \rangle_{\sigma_j}$.
5) Upon a node p_j received $\lceil (n+t+1)/2 \rceil$ SUPPORT for the same v_{med}, it broadcasts $\langle \text{DECIDE}, seq, p_j, v_{med} \rangle_{\sigma_j}$.
6) Upon a node p_j received $\lceil (n+t+1)/2 \rceil$ DECIDE for the same v_{med}, it decides v_{med}.

From step 3 on, BFT-ARM is similar to the PBFT protocol [6]. So if the leader proposes the correct v_{med} matching the certificate, all correct nodes will decide v_{med}.

Now we prove v_{med} is valid according to Definition 1.

Theorem 1. *Let SA be the sorted array of the input values of any $(n-t)$ nodes. The median of SA is denoted as $v = SA[\lceil \frac{n-t}{2} \rceil - 1]$. Then v is valid.*

Proof. According to the definition of median, there are at least $(\lceil \frac{n-t}{2} \rceil - 1)$ nodes whose value is no greater than v. Among them there are at least $(\lceil \frac{n-t}{2} \rceil - 1 - f)$ good nodes.

And because $f \leqslant t < n/3$, we have $\lceil \frac{n-t}{2} \rceil - 1 - f \geqslant \lceil \frac{n-f}{2} \rceil - 1 - t \geqslant 0$. So $v \geqslant SG[\lceil \frac{n-t}{2} \rceil - 1 - f] \geqslant SG[\lceil \frac{n-f}{2} \rceil - 1 - t]$. Similarly, we can prove that $v \leqslant SG[\lceil \frac{n-f}{2} \rceil - 1 + t]$. Because of the Definition 1, v is valid. □

Thus the validity of the proposal can be confirmed by comparing with the certificate of $(n - t)$ values in step 4.

We use the trusted counter to generate a sequence number for every START message from the leader. The sequence number is monotonically increasing by one every time, so there is exactly one sequence number assigned to every consensus period. In this way, faulty nodes cannot provide an outdated value (replay attack). If a node detects that the sequence number does not belong to this period, it will discard the message. A synchronized clock is not required here, but the interval of the period is known to everyone. From the first time a node receives the sequence number from the leader, it can determine the correspondence between the sequence number and period.

B. Suspect leader protocol

When the leader is faulty or disconnected from the group, leading to a fail of consensus within a predefined timeout, the other nodes will initiate a suspect leader protocol, basically similar to the PBFT view change protocol (without considering about the history). When a node p_j suspects the leader p_{cur}, it broadcasts a $\langle \text{SUSPECT}, p_j, p_{cur}, p_{new} \rangle_{\sigma_j}$, where p_{new} is the next leader according to a deterministic rule, e.g., based on the position information of the platoon to choose the one behind the current leader until the end and then change the direction forwards.

When p_{new} receives $(\lceil (n + t + 1)/2 \rceil - 1)$ messages suspecting current leader, it takes over the leader role and broadcasts $\langle \text{NEWLEADER}, p_{new}, seq_{new} \rangle_{\sigma_{new}}$ with its own sequence number, and operates as in the normal case.

V. Evaluation

To evaluate BFT-ARM in platooning environments, we intend to use an extended version of the *ArteryLTE*[3] simulation framework, which is detailed in [14].

A. Simulation Framework

ArteryLTE is based on the renowned open-source *Vehicles in Network Simulation (Veins)* framework [15]. The *Veins* project[4] combines the dedicated network simulator OMNeT++ with the microscopic traffic simulator *Simulation of Urban Mobility* (SUMO). In addition, Veins also provides an implementation of the US Wireless Access in Vehicular Environments (WAVE) Dedicated Short Range Communication (DSRC) stack based on IEEE 802.11p.

ArteryLTE integrates several extensions to Veins:

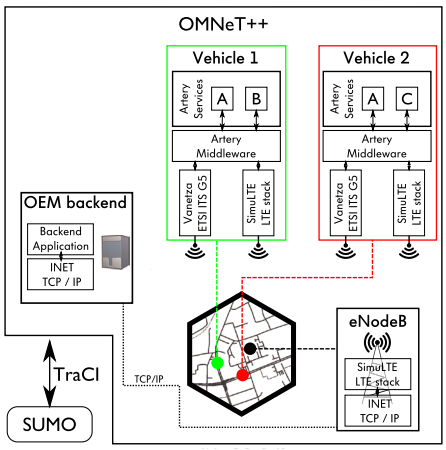


Figure 2. Architecture of the *ArteryLTE* simulation framework.

First, a modular middleware for Veins called *Artery*[5] [16] is used to implement heterogeneous vehicle capabilities. Multiple applications (so-called Artery services) can be implemented and dynamically configured for vehicles per market penetration rates. Furthermore via *Vanetza*[6], the European equivalent to the WAVE stack, the European Telecommunications Standards Institute (ETSI) Intelligent Transport System (ITS) G5 protocol stack, is brought in and used to disseminate Cooperative Awareness (CA) messages [17].

Second, *ArteryLTE* integrates Long Term Evolution (LTE) support for vehicles as introduced to Veins by the *VeinsLTE* [18] project, thus enabling heterogeneous communication technologies on the network nodes. VeinsLTE's *decision maker* is replaced by an option in *Artery*'s middleware that allows Artery services to choose between either the ITS G5 or the LTE stack for communication.

Third, *ArteryLTE* includes support for backend-based applications. A backend is represented by a static network node in the network which is connected to the eNodeBs of the LTE network.

The overall architecture of the *ArteryLTE* simulation framework is depicted in Figure 2. In the presented cell of the eNodeB two vehicles are shown, both equipped with an LTE and an ITS G5 stack. Different Artery services (A,B and C) are deployed on the vehicles respectively. Data transmitted via LTE by the vehicles is forwarded using a Transfer Control Protocol (TCP) connection between the eNodeB and the backend.

Furthermore, local perception sensors for advanced ADASs are the latest addition to *ArteryLTE* [14].

[3] https://github.com/ibr-cm/artery-lte
[4] http://veins.car2x.org/
[5] https://github.com/riebl/artery
[6] https://github.com/riebl/vanetza

B. Extension of the Framework

We are bringing yet another extension into the *ArteryLTE* framework: The *Plexe* extension [19] to Veins enables the simulation of vehicle platoons with corresponding control algorithms, such as for cruise control, and the implementation of cooperative driving applications. We are in the process of porting the changes made by *Plexe* to Veins to *ArteryLTE*'s codebase so that ArteryLTE is able to interact with Plexe's SUMO version via the Traffic Command Interface (TraCI) protocol. We will use the platooning examples and the included control algorithms of Plexe as the basic scenario for our application. Vehicles of the platoon will—in a first step—be equipped with IEEE 802.11p for local communication to run the presented consensus protocol.

VI. Conclusion and Future Work

As soon as the basic setup of BFT-ARM is implemented, in a first step we evaluate the characteristics of the consensus protocol among vehicles via IVC. We then intend to use the whole potential of our communication environment to improve the consensus process as well as to introduce further features. For example, to take advantage of the available heterogeneous networks, we envisage the ability to fall back to cellular communication in cases where local communication of a group is disrupted. Furthermore, in the case of ADASs that are tightly coupled to an Original Equipment Manufacturer (OEM) backend, running an agreement might be assisted by this backend as, e. g., the backend may initiate a consensus, or a leader change based on data available to the backend such as network metrics or local traffic data [14].

References

[1] Theodore Willke, Patcharinee Tientrakool, and Nicholas Maxemchuk. "A survey of inter-vehicle communication protocols and their applications". In: *IEEE Communications Surveys & Tutorials* 11.2 (2009), pp. 3–20.

[2] Pedro Fernandes and Urbano Nunes. "Platooning With IVC-Enabled Autonomous Vehicles: Strategies to Mitigate Communication Delays, Improve Safety and Traffic Flow". English. In: *IEEE Transactions on Intelligent Transportation Systems* 13.1 (Mar. 2012), pp. 91–106.

[3] Michele Segata et al. "Supporting platooning maneuvers through IVC: An initial protocol analysis for the JOIN maneuver". English. In: *2014 11th Annual Conference on Wireless On-demand Network Systems and Services (WONS)*. IEEE, Apr. 2014, pp. 130–137.

[4] Leslie Lamport, Robert Shostak, and Marshall Pease. "The Byzantine Generals Problem". In: *ACM Trans. Program. Lang. Syst.* 4.3 (July 1982), pp. 382–401.

[5] Gabriel Bracha. "Asynchronous Byzantine agreement protocols". In: *Information and Computation* 75.2 (1987), pp. 130–143.

[6] Miguel Castro and Barbara Liskov. "Practical Byzantine Fault Tolerance". In: *Proceedings of the Third Symposium on Operating Systems Design and Implementation*. OSDI '99. New Orleans, Louisiana, USA: USENIX Association, 1999, pp. 173–186.

[7] Kim Potter Kihlstrom, Louise E Moser, and P Michael Melliar-Smith. "Byzantine fault detectors for solving consensus". In: *The Computer Journal* 46.1 (2003), pp. 16–35.

[8] Gil Neiger. "Distributed consensus revisited". In: *Information Processing Letters* 49.4 (1994), pp. 195–201.

[9] David Stolz and Roger Wattenhofer. "Byzantine Agreement with Median Validity". In: *19th International Conference on Priniciples of Distributed Systems (OPODIS), Rennes, France*. 2015.

[10] S Santini et al. "A consensus-based approach for platooning with inter-vehicular communications". In: *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE. 2015, pp. 1158–1166.

[11] Michael J Fischer, Nancy A Lynch, and Michael S Paterson. "Impossibility of distributed consensus with one faulty process". In: *Journal of the ACM (JACM)* 32.2 (1985), pp. 374–382.

[12] Giuliana Santos Veronese et al. "Efficient byzantine fault-tolerance". In: *Computers, IEEE Transactions on* 62.1 (2013), pp. 16–30.

[13] Rüdiger Kapitza et al. "CheapBFT: Resource-efficient Byzantine Fault Tolerance". In: *Proceedings of the EuroSys 2012 Conference (EuroSys '12)*. Ed. by European Chapter of ACM SIGOPS. Bern, Switzerland, 2012, pp. 295–308.

[14] Julian Timpner et al. "Towards a Multi-Protocol Microscopic IVC Simulation Environment for ADASs". Berlin, 2016.

[15] C. Sommer, R. German, and F. Dressler. "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis". In: *IEEE Trans. Mobile Comput.* 10.1 (Jan. 2011), pp. 3–15.

[16] R. Riebl et al. "Artery - Extendig Veins for VANET applications". In: *Models and Technologies for Intelligent Transportation Systems (MT-ITS)*. 2015.

[17] *ETSI EN 302 637-2 V1.3.1 - Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service*. ETSI, Sept. 2014.

[18] F. Hagenauer, F. Dressler, and C. Sommer. "Poster: A simulator for heterogeneous vehicular networks". In: *Proc. Vehicular Networking Conference (VNC)*. IEEE, Dec. 2014, pp. 185–186.

[19] Michele Segata et al. "Plexe: A platooning extension for Veins". In: *2014 IEEE Vehicular Networking Conference (VNC)*. IEEE, Dec. 2014, pp. 53–60.

On Context-Aware Communication Mode Selection in Hybrid Vehicular Networks

Smriti Gopinath, Lars Wischhof, Michael Jaumann

Department of Computer Science and Mathematics
Munich University of Applied Sciences (MUAS), Germany
[smriti.gopinath|wischhof|michael.jaumann]@hm.edu

Christoph Ponikwar, Hans-Joachim Hof
MuSe - Munich IT Security Research Group
Department of Computer Science and Mathematics
Munich University of Applied Sciences (MUAS), Germany
[christoph.ponikwar|hof]@hm.edu

***Abstract*—Future vehicles will most-likely have multiple communication technologies and modes available. After classifying V2X applications in five distinct classes, a context-aware selection of the communication mode is advocated. A suitable architecture is outlined. First simulation results for the example of a DENM-based application indicate that a context-aware selection outperforms a static assignment.**

***Index Terms*—hybrid vehicular networks, requirements, context-indicators.**

I. Introduction

Vehicle-to-vehicle and vehicle-to-infrastructure communication, often summarized as vehicle-to-X (V2X) communication, has been an active research topic for more than a decade. Respective standards have been defined, e.g. in Europe with the ITS-G5 or the US with the WAVE/IEEE 1609 standards. However, until now no OEM has introduced V2X-communication based on these WLAN-like technologies on a larger scale. On the other hand, cellular communication technologies – for voice and data communication – are standard in medium and high end vehicles. Cellular data communication is used to connect to a (OEM-specific) backend system, e.g. in order to perform infotainment functionalities such as online search or to obtain online traffic information.

More recently, the standardization bodies for cellular communications have added variants of direct communication, e.g. the so-called proximity services allow Device-to-Device (D2D) communication as part of the LTE standards. While these extensions where not initially developed for vehicular applications, it is now considered in a recent 3GPP study [1]. Furthermore, the automotive industry is considered an important application sector for the future 5G cellular standards [2], which will most-likely also support a direct communication. The motivation to include direct communication in cellular standards for vehicular use-cases is two-fold: Some vehicular use-case require high message rates (in the order of 10 Hz or more) and low delays which are hard to achieve by indirect cellular communication and would most-likely require a high level of spacial reuse, i.e. high cost of deployment. Additionally, from the mobile operator point of view, direct communication can reduce the traffic (and resulting costs) in the core network, a fact that is also exploited by offloading [3] in other domains.

Therefore, it can be assumed that future connected vehicles will have multiple communication modes available: cellular/indirect communication and direct communication, the latter either provided by the well-known IEEE 802.11p or as part of the cellular standard itself. Depending on the coverage situation of the vehicle, four different situations can be distinguished, as illustrated in Fig. 1. This implies that depending on the local situation of a vehicle, the communication mode must be selected depending on criteria such as QoS requirements of an application (i.e. latency requirements, dissemination area, etc.), status of each communication mode (availability, current load, signal strength, etc.), number of other vehicles in range, and many more.

This context-aware selection of the suitable communication mode is particularly important during the phase of market introduction of the direct communication technology: During the first years of market introduction of the direct communication technology in most situations there will be no communication partner in range for direct communication. However, since cellular communication can rely on an existing, already deployed infrastructure, a vehicle aware of this situation can switch to cellular communication. In contrast, when a high market penetration has been achieved, vehicles should detect situations with highly loaded cellular networks and/or multiple vehicle with direct communication capabilities within range, in order to reduce the load in the core network of the cellular system and to avoid overload conditions.

The process of selecting the appropriate communication mode is non-trivial and – in order to avoid redundant, potentially inconsistent, implementation – from our point of

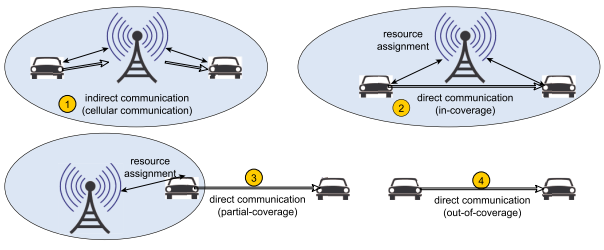


Fig. 1. Communication modes for cellular and direct communication.

view should be implemented in an intermediate communication layer, transparently handling the selection process from the application. The paper first presents a classification of vehicular networking applications along with their requirements. Four diverse use-cases, representative of each class and the applicability of the various wireless technologies, are discussed further. Based on these observations, a novel Hybrid Overlay Protocol (HOP) layer is proposed, which uses context-indicators in order to select the optimal communication mode. The paper concludes with preliminary simulation results for a specific context indicator, the number of vehicles in direct communication range, which illustrate the benefits of the proposed concept.

A. Related Work

In general, the challenge of selecting the appropriate communication mode in cellular networks supporting direct/D2D communication has been considered in several publications, an overview is given in [4]. Criteria such as the distance or the link quality to a direct communication partner are considered. However, these criteria are difficult to apply in C2X communication due to the rapidly changing positions and link qualities. For hybrid vehicular networks, Zheng et. al. in [5] introduce a Hybrid Link Layer (HLL) for load and resource sharing between cellular networks and IEEE 802.11p. In contrast, the focus of our work is not load sharing but selecting the optimal mode and technology on a per-packet basis according to the requirements/class of the generating application.

II. Vehicular Networking Applications – Requirements and Challenges

V2X applications are usually classified into safety and non-safety based categories. Within this article, a more specific classification of applications into five classes is put forward based on specific requirements on the wireless technologies:

Cooperative Sensing (Safety) applications use V2X communication for situation awareness, e.g. to reduce risks of accidents while driving. Vehicles in a local area communicate periodically in order to inform each other about their position, speed, acceleration and path, for example via periodic Cooperative Awareness Messages (CAM). The challenge here is finding low-latency, reliable, and efficient methods for disseminating safety-relevant data among neighbouring vehicles.

Cooperative Sensing (Information/Non-Safety) applications use communication to extend the horizon of perception for driver information systems. While conventional on-board sensors of a vehicle (e.g. camera, radar or lidar) depend on a line-of-sight situation and are limited to a range of approx. 50-200 meters, V2X communication can overcome these limits. The delay and reliability requirements are not as stringent as for safety applications, but the range in which the vehicle needs to be aware of relevant information is large, i.e. in general it exceeds 5 km.

Cooperative Maneuvering applications apply C2X communication for driving automation functions in the levels 3 to 5 as defined in SAE J3016. In order to realize cooperative maneuvers, vehicles use bidirectional communication, e.g. in order to exchange information on planned trajectories and agree on trajectory changes. Low latency ($\leq$ 10ms, [2]), reliable bi-directional communication is a key requirement for this application class.

In-Vehicle Internet Access applications extend the Internet into the vehicle by offering Internet-based applications for the driver and passengers. The acceptable delay as well as the required data rate are similar to those of typical smartphone use-cases.

Mobility Monitoring and Configuration applications involve communicating with a (usually parked) vehicle remotely in order to obtain information on its status. The user interacts with the vehicle via smartphone or using an Internet website.

A. Applicability of Wireless Technologies

The use-case classes in Sec. II differ to a large extent in their requirements and applicable technologies (Tab. I[1]). For safety applications, the stringent low delay requirements cannot always be fulfilled by current cellular technologies. 5G technologies might include a low-latency direct communication mode, e. g. as evaluated in the METIS project. Cooperative Sensing (Non-safety) applications involve a larger area of dissemination and interaction with fewer vehicles which can often be satisfied by cellular communications and by direct communication – if a suitable data dissemination scheme is applied [6].

III. Hybrid Overlay Protocol (HOP)

As motivated in Sec. I, the proposed solution to match the widely varying requirements of the application classes of Sec. II to the capabilities of the wireless communication technologies and modes in a specific context is to introduce an intermediate HOP layer, as illustrated in Fig. 2. This basic idea has already been introduced in a previous article [7], whereas the focus in this article is on the context-aware communication mode selection. Therefore, we will summarize the main aspects relevant for the presented results.

A. Basic Concept

Future V2X communication systems will support communication in at least two modes: a direct/ad-hoc mode and a indirect/cellular mode. Considering the highly dynamic vehicular environment with rapidly changing transmission conditions, the HOP layer adaptively decides the optimal communication mode on a per-packet basis. The selection of the communication mode involves the calculation of values characterizing the current context of the vehicle, termed Context Indicators (CIs) in the following. CIs can be based on vehicle sensors, e.g. vehicle speed, or on (meta-)information received from the

[1]Technologies that cannot completely fulfil all requirements are enclosed in parentheses. For 5G, a suitable direct communication mode is assumed.

TABLE I
APPLICABILITY OF WIRELESS TECHNOLOGIES

Category	Use Case Example	Applicable Wireless Technologies
Cooperative Sensing (Safety)	Intersection Collision Avoidance	802.11p, 5G, (UMTS, LTE, LTE-A)
Cooperative Sensing (Non-safety)	Dynamic Map Information	(802.11p, LTE-A D2D), UMTS, LTE, LTE-A, 5G
Cooperative Maneuvering	Cooperative Lane-Merging	(802.11p, LTE-A), 5G
In-Vehicle Internet Access	Information Retrieval from Internet Websites	UMTS, LTE, LTE-A, 5G
Mobility Monitoring and Configuration	Car Status Information	GPRS/EDGE, UMTS, LTE, LTE-A, 5G

lower communication layers, e.g. average data rate, channel busy time ratio, etc. In addition to the data payload, the applications also specify their requirements in form of requirement indicators (RIs) such as the maximum latency, range of dissemination, etc. The communication mode is then selected by matching the calculated CI's with the target RI's.

1) Context-Aware Communication Mode Selection: For initial simulations, three CIs are considered:

Channel Quality Indicator (CQI) of the LTE downlink. Cellular mode is used only when the CQI value measured is above a threshold value C_{CQI}.

Queue length of LTE in uplink. If it exceeds C_{LTEqueue}, the cellular network is assumed to be highly loaded.

1- & 2-hop Neighbour Count of vehicles capable of direct communication seen in the last T_{neigh} seconds. If the number of vehicles in 2-hop range exceeds $C_{2\mathrm{hop}}$, it is assumed that information can be disseminated in a large area via direct mode.

2) Mode Selection: is performed based on these three CQIs in the following way: In case of poor conditions for cellular networks are indicated (CQI $\leq C_{\mathrm{CQI}}$ or queue length $> C_{\mathrm{LTEqueue}}$), direct mode is selected. In case of good cellular conditions, cellular mode is selected if less than $C_{2\mathrm{hop}}$ neighbours are in two-hop range, or if the cellular network has not been used for a period of T_{cell}. In all other cases, direct mode is selected. The rationale for the latter is to guarantee a minimum rate of messages on the cellular network, to reduce the delay for wide-area dissemination of messages (assuming the RI indicates a large dissemination range). The CQI and neighbour count values are periodically updated at a configurable frequency independent of the mode selection, queue length CI is updated via signal from MAC to HOP layer.

B. Simulation

For simulative evaluation, the proposed architecture is implemented in the discrete event simulator OMNeT++ 5.0b3. The network simulation is based on the INET-framework in version 3.2.1, SimuLTE [8] for simulating the LTE user plane and veins/veinslte [9]. As a first step, an example application of the Cooperative Sensing (Information/Non-Safety) class is investigated which sends Decentralized Environmental Notification Messages (DENMs). Received DENMs are forwarded using Contention-Based Forwarding (CBF) and repeated for a duration of 5 mins. Performance criteria is the number of DENMs with unique actionIDs received *on time to react*. A message is received on-time if its delay is less than the time needed by a fast vehicle (180 km/h) to get to the position where it was sent, i.e. a driver has sufficient time to react.

Due to the limited space, we present only a single 2x2 highway scenario with a traffic density of 8.33 veh/lane/km where 5% of the vehicles use C2X communication. Vehicles generate DENMs with new ActionIDs with normally-distributed inter-event times with a mean of 30 s and a standard deviation of 10s. A single LTE cell using 100 RBs in a ITU-T rural macro-cell scenario is used for cellular communication, for direct communication IEEE 802.11p (with veins default parameter values) is used. Context-aware communication mode selection is based on the CIs described in Sec. III-A1 with parameters $C_{\mathrm{CQI}} = 0$, $C_{\mathrm{LTEqueue}} = 20kB$ and $C_{2\mathrm{hop}} = 10$.

Fig. 3 compares the proposed CI-based scheme with IEEE 802.11p only, LTE-A only. In order to obtain an upper-bound for the performance, we also show the result if all messages are always sent on both media and the one, which is received earliest, is counted. It can be observed that for lower distances, direct communication outperforms cellular communication. For larger distances, LTE outperforms direct communication, as it can be expected in this kind of scenario. The adaptive scheme outperforms both single technologies. However, for medium distances, redundant transmission on both media is outperforming the adaptive scheme – indicating that the adaptive selection is non-optimal in some cases.

C. Conclusions

Following a classification of V2X use-cases in five distinct categories, in this article, a hybrid overlay architecture has been advocated which performs a context-aware selection of communication modes in case multiple modes are available. Initial simulation results for a mode selection scheme based on CQI, cellular queue length and number of neighbors illustrate that an adaptive selection can outperform a static selection. A systematic investigation of CIs and their performance is therefore the next step.

REFERENCES

[1] 3GPP, *TR 22.885 V14.0.0 (2015-12): 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on LTE Support for V2x Services (Release 14)*, Dec. 2015.

[2] 5GPP, ERTICO ITS EUROPE, and European Commission, *5g Automotive Vision*, Oct. 2015.

[3] S. Andreev, O. Galinina, A. Pyattaev, K. Johnsson, and Y. Koucheryavy, "Analyzing Assisted Offloading of Cellular User Sessions onto D2d Links in Unlicensed Bands," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, 2014.

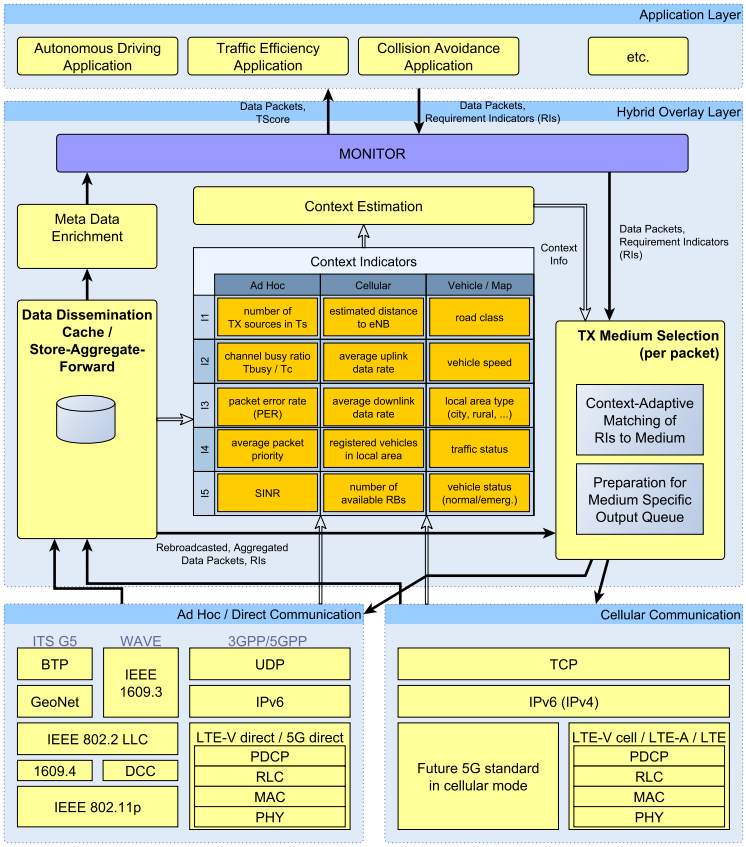


Fig. 2. Architecture of Hybrid Overlay Protocol (HOP) Layer and its proposed integration in the communication stack. Solid lines indicate the flow of the data packets, contoured lines indicate the flow of status/meta-information.

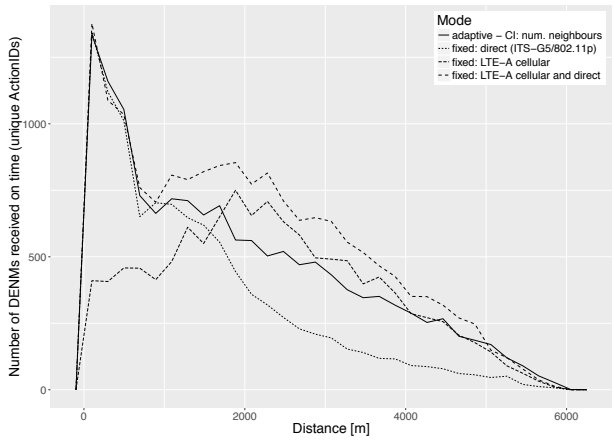


Fig. 3. Comparison of number of DENMs (uniq. action-ID) received on time.

[4] D. Marshall, S. Durrani, J. Guo, and N. Yang, "Performance comparison of device-to-device mode selection schemes," IEEE, Aug. 2015, pp. 1536–1541.

[5] K. Zheng, Q. Zheng, P. Chatzimisios, W. Xiang, and Y. Zhou, "Heterogeneous Vehicular Networking: A Survey on Architecture, Challenges and Solutions," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2015.

[6] L. Wischhof, A. Ebner, and H. Rohling, "Information dissemination in self-organizing intervehicle networks," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 6, no. 1, pp. 90–101, 2005.

[7] S. Gopinath, L. Wischhof, C. Ponikwar, and H.-J. Hof, "Hybrid Solutions for Data Dissemination in Vehicular Networks," in *Proc. 8th International Wireless Days Conference*, Toulouse, France, Mar. 2016.

[8] A. Virdis, G. Stea, and G. Nardini, "Simulating LTE/LTE-Advanced Networks with SimuLTE," in *Simulation and Modeling Methodologies, Technologies and Applications*, M. S. Obaidat, T. Oeren, J. Kacprzyk, and J. Filipe, Eds., vol. 402, Cham: Springer International Publishing, 2015, pp. 83–105.

[9] F. Hagenauer, F. Dressler, and C. Sommer, "Poster: A simulator for heterogeneous vehicular networks," IEEE, Dec. 2014, pp. 185–186.

Impact of a Three Dimensional Environment to Inter-vehicle Connectivity

Lisa Kristiana, Corinna Schmitt, Burkhard Stiller
Communication Systems Group CSG, Department of Informatics IFI, University of Zurich
Binzmühlestrasse 14, CH—8050 Zurich, Switzerland
[kristianalschmittlstiller]@ifi.uzh.ch

***Abstract*—Developing non-safety applications, such as Web surfing and social network, for inter-vehicle networks requires a reliable and stable connectivity among vehicles. One challenge to reach such a reliable and stable connectivity is the road in a large city environment, as it appears in a three dimensional topology (*i.e.* a road with overpasses). These situations lead potentially to restricted connectivity since with respect to propagation vehicles are driven on different road levels, which can well form obstacles such that the connectivity among vehicles is disturbed. This paper addresses specifically the three dimensional topology of roads in terms of a level environment model and investigates the impact of various height of overpass between two communicating vehicles.**

***Index Terms*—Road topology, Inter-vehicle connectivity, Three-dimensional forwarding**

I. Introduction

The inter-vehicle connectivity for non-safety applications, such as Web surfing and social network can be performed during transportation [1]. Therefore, reliable and stable inter-vehicle connectivity is required. The term inter-vehicle connectivity refers to a process of basic communication between two vehicles such as exchanging mutual message and locating position coordinates [2].

Fig. 1. Road Level Topology. (Source: www.media.viva.co.id)

As in large city environment, there are several major factors that influence reliable and stable connectivity. The first factor is propagation [3]. During transmission and reception phase, a signal propagates over the environment components (e.g., buildings, trees, and other vehicles) [4], [5], and the attenuation occurs. The second factor is the mobility of vehicles which leads to frequent network topology changes [2]. Last but not least, three-dimensional topology of road, as shown in Figure 1, affects the signal reception [6]. Looking to the aspect of road level topology, (i.e., an overpass), the signal reception is often weak due to distraction such as reflection, attenuation, and scattering by overpass' shapes [7]. This poor signal reception can cause disconnection of transmission and followed by probability of reconnects and reestablishment of a new connection.

Several proposed approaches do not consider the road level topology. Thus, the major issue that is still being investigated is the decreasing of signal reception when vehicles move in the different road level. In order to analyze the impact of various height of road topology level to inter-vehicle connectivity, this situation leads to the following research questions:

1) How is the impact of overpass to network performance?

2) Will the height of road level influence the inter-vehicle connectivity?

This paper is organized as following: Section II provides related work with respect to the technology, propagation model, and forwarding method. Section III discusses road level forwarding model in details. Section IV provides simulation parameters and evaluation of the addressed road level forwarding model with various road levels. Finally, summary and future work are provided.

II. Related work

The term mobile node represents a vehicle and it is assumed to be equipped with a navigation system as Global Positioning System (GPS) and wireless communication (Wi-Fi/IEEE 802.11), therefore, the term mobile node and vehicle can be used alternately. In inter-vehicle network, Wi-Fi/IEEE 802.11 as a short range radio technology is possible to be used to establish a communication between vehicles [9]. A Wi-Fi ad-hoc mode can support inter-vehicular networking through the ad-hoc broadcast [10]. This communication technology has been enhanced for non-safety application with required modification since it has to support communication between vehicles moving at high speeds [9]. In consideration of two-dimensional area which refers to an euclidean area, therefore, it leads to inaccuracy of three-dimensional modeling [8], [13].

Normally, roads in a large city environment can have contour characteristic. This means that some roads can have various levels topology. This overpass topology leads to two key issues *i.e.*, propagation model and forwarding scheme are discussed as following:

A. Propagation model in large city environment

The characteristic of propagation channel may vary depending on the environment. Propagation characteristic influences both signal transmission and reception [17]. A con-

sideration of propagation model that is influenced by the existence of obstacles is proposed as an approach to obtain an optimum transmission [17]. In case of buildings, composed of a concrete block, signal transmission will be attenuated or even restricted [7], [18]. In case of overpasses, it is assumed that the overpass is not made from material with good conductivity, thus, signal attenuates or restricted along the overpass [8].

By definition, signal transmission that enter the overpass is assumed as signal loss since it will fade, depending on overpass length (*e.g.*, GPS signal for navigation systems and Wi-Fi signals degradation) [7]. The longer the overpass, the signal loss probability rises and the signal reception is decreased. There is a trade-off whether to disconnect the transmission and search for a new connection or to maintain the current and distracted connection. For instance, when a vehicle moves below overpass with high speed, but suddenly decreasing the speed due to traffic condition, the distracted connection expands or the vehicle becomes temporary unreachable [14].

It is important to define the particular environment as a preliminary set up. In a free space environment, (*i.e.* the environment where the electromagnetic wave transmits without any obstructions) the propagation channel is considered as line-of-sight transmission model. This model is only in theoretical case and used as a reference to other models. In case of road hierarchy topology, the propagation channel is modeled as a propagation loss model with overpass as an obstacle [8]. This model takes into account of height of road and it is assumed one electromagnetic wave ray will be received directly, while another ray will reflect on the ground and other objects which is known as nakagami propagation model [17], [24].

B. Forwarding Method

Generally, routing protocols play an important role to ensure all packets are transmitted from sender node to destination node. The main core of routing protocols is a forwarding decision mechanism. This forwarding mechanism decides the best method to transmit the information from sender node (*S*) to the next receiver node (*R*) and finally to the destination node (*D*). During the decision process, *S* has to select the proper *R* of all intermediate nodes (*I*). Intermediate nodes are mobile nodes which has a possibility as a next hop node.

In order to select most appropriate *I*, complete information of all neighboring nodes is collected to provide a valid forwarding decision. Methods of forwarding decision vary with respect to vehicle's complete information, such as planar position information (*i.e.* distance between vehicles) [12], [19], transmission power information (*i.e.* signal power), mobility information (*i.e.* velocity) [13], and non-planar position information (*i.e.* angle) [25].

Most of all forwarding method experiments are applied in two-dimensional area. Thus, several challenges are considered in applying forwarding method in three-dimensional area. These challenges are: (1) The distance of the corresponding vehicles on upper road level and lower road level, can form further transmission range. (2) The various speed of the vehicles can form a frequent topology changing, which can effect transmission disconnection. (3) The direction factor becomes more complex when it is applied in three-dimensional environment than in two-dimensional environment. Based on greedy forwarding method [20], Link State Aware Hierarchical Road (LSHR), takes into account of vehicles which are located on the same road level to forward the packet and avoid the vehicle which are located in different road level [22]. However, the overpass is not considered as obstacle in this work.

III. Road Level Forwarding Model

This work deploys the angle-based propagation scheme to greedy forwarding concept. The idea of deploying the angle-based propagation is to realistically restrict the area of forwarding. Angle-based restricted scheme filters out intermediate node candidates due to the width of road and the height of road. In one hand the horizontal relative angle concept is implemented when vehicles are in the planar area, on the other hand, the vertical relative angle is implemented in the non-planar area. When the area restriction is set, thus, the greedy forwarding concept is implemented.

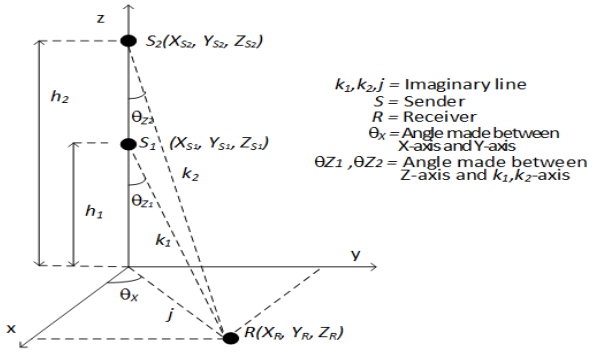


Fig. 2. Two Angles Describe Vehicles' Position on Different Road Layer Topology.

The scenario as illustrated in Figure 2, black dots represent vehicles which are located on both upper road level and lower road level. In three-dimensional environment, it is necessary to considers *z*-coordinate as an important parameter to locate a position accurately [21]. Thus, the location of a vehicle can be represented as coordinates (*x, y, z*).

Distance: It is influenced by speed and direction factors. Distance factor leads to maximum, optimum, and minimum transmission modes. It is based on distance between current and intermediate nodes allowing to define the transmission range mode as follows. Given a current mobile node b_i has geographical coordinates of x_i, y_i and z_i. The potential neighbor node n_d with coordinates of x_d, y_d and z_d. Thus, the Euclidean distance between the two is given in Equation 1:

$$\Delta d = \sqrt{(x_d - x_i)^2 + (y_d - y_i)^2 + (z_d - z_i)^2} \qquad \text{(Eqn. 1)}$$

The closer the distance, the better the connectivity. In addition, the distance between vehicles is correlated with speed. Thus, the speed given a velocity vector of a current node b_i is given in Equation 2:

$$v_i = \left(\sqrt{v_{xi}^2 + v_{yi}^2}\right) \qquad \text{(Eqn. 2)}$$

Relative Angle: It is also necessary to consider various height and width of the road due to signal transmission and reception. A current node and intermediate node on a different road level (*i.e.* vehicles on upper road layer and lower road layer) can create an angle between them as illustrated in Figure 2. Angles in degrees are measured in two ways: First, it is measured between the positive *x*-axis and positive *y*-axis, which results in θ_x while the second angle θ_z is measured

between positive z-axis and the vehicle located on lower layer road. This θ_z angle influences transmission range between vehicles on upper and lower road level. In order to simplify and clearly describe two communicating vehicles, S - a vehicle moving on upper road level - forms an angle θ_z with respect to R - a vehicle moving on lower road layer.

$$\theta_{z1} = atan\left[\frac{\sqrt{x_R^2 + y_R^2}}{h1}\right] \quad \text{(Eqn. 3)}$$

$$\theta_{z2} = atan\left[\frac{\sqrt{x_R^2 + y_R^2}}{h2}\right] \quad \text{(Eqn. 4)}$$

To generalize the complexity of road level topology, it is important to analyze various road levels means that different heights of roads form different angles can be calculated by Equations (3) and (4), where $h_1 = z_{S1} - z_R$, and $h_2 = z_{S2} - z_R$, (cf. Figure 2). It can be assumed that the higher the upper road, the smaller angle measured as following the right-angled triangle formula. In this work, a modified propagation loss model is used with an addition of obstacle aware propagation. The obstacle aware propagation will block the signal within the specific range as briefly described in section II.A.

First assumption that has to be made is the vehicle distribution. Vehicles are located both on upper and lower road level. Sending process are done by vehicles on upper road level and receiving process are done by vehicles on lower one. By default, this scenario uses the overpass height of 10 m to 20 m. Both road levels have two lanes and in the middle of lower road level, an overpass crosses over the lower road level. The angle is measured from S to R. S can be both the origin source or the current sender. The x-axis represents the width of road, y-axis represents the length of road and z-axis represents the height of road. In order to simplify the relative-angle calculation between two nodes (*i.e.*, source and intermediate nodes), the z-axis is predefined.

Second assumption is that the angle is measured when S detects an intermediate node, which is located on the lower road level and in line with S. The intermediate node can be a final destination or the next forwarded mobile node. Thus, the measured angles between S and R (*i.e.* θ_x and θ_z), forms perpendicular intersection of two straight lines.

IV. Simulation

The measurement of link performance assumes end-to-end point connection which means that the connection is evaluated from original sender to final destination.

TABLE I: Parameter Settings

Parameter	Units
Transmission Range IEEE 802.11b/g	140 m
Number of Nodes	10 - 40
Simulation Area	500 m x 500 m
Upper Road Height	10 - 20 m
Average Vehicle Velocity	30-70 km/h
Packet Size	1024 Byte
Simulation Time	500 s
Number of Driving Lanes	2

Table I lists the chosen parameters using NS3 [26] as a simulation tool. The experiment is conducted with the following scenario: The simulation area is set as 500 m x 500 m with the first assumption that the vehicles are moving in a free traffic (*i.e.* no traffic light and no traffic jam). The second assumption is to ensure that vehicles, which are driving on both road levels, experience out of coverage from one vehicle to another. An initial position of each vehicle is located on both road levels with the average speed of vehicles span from 30 to 70 km/h. In this scenario the routing protocols GPRS is used. This routing protocol is selected since it represents a position-based routing protocol which is not using beacon but relying on position of current node. Thus, GPSR is suitable due to rate of change of the topology.

V. Initiate Evaluation

The evaluation results are illustrated in Figures 3-6. The simulation describe the existence of overpass and the various height of road level. These different heights of road level, In the other words, the height of road level creates non-extreme disconnection with due to the occurrence of the out of coverage events.

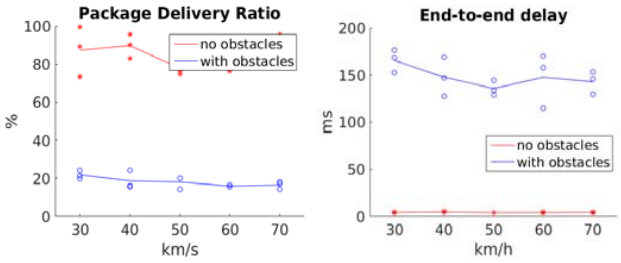


Fig. 3. Impact of Overpass versus Speed

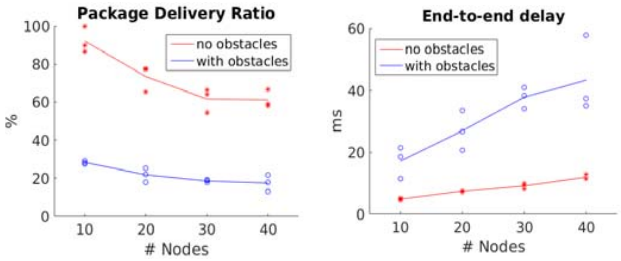


Fig. 4. Impact of Overpass versus Number of Nodes

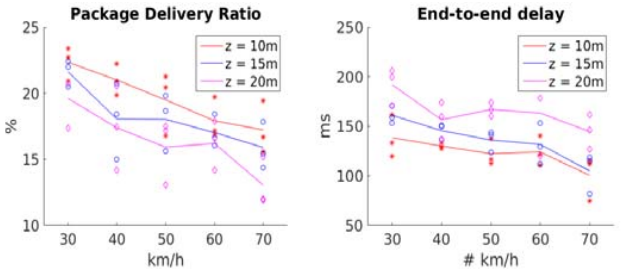


Fig. 5. Impact of Road Height versus Speed

Figure 3 is the result of applying the obstacle propagation model and it is compared with the nakagami propagation model. In this case obstacles are simply blocking the signal

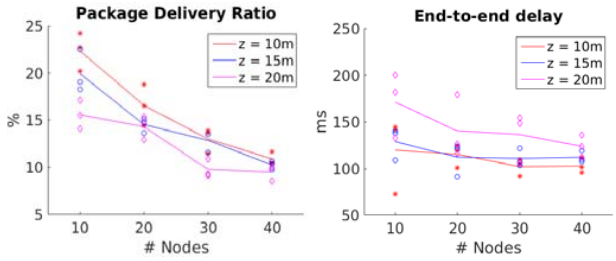


Fig. 6. Impact of Road Height versus Number of Nodes

transmission, thus, both Packet Delivery Ratio (PDR) and End-to-End (E2E) delay have less performance compare to the nakagami propagation. Figure 4, describe the network performance with respect to Figure 3. With the high mobility, high E2E delay also occurs due to obstacle existence. Figure 5 and Figure 6 show network performance of the various height of overpass. In these two cases, overpass is considered as the obstacle with the horizontal position, thus, simply blocking the signal transmission whenever a connection occurs between vehicles which are located in the same x-axis coordinates, which basically means that one vehicle is located right on the top of other vehicle (i.e., on the overpass). Therefore, it is obvious that disconnections occur. This also shows the higher the road topology level leads to the higher chance of disconnection due to the transmission range.

VI. Summary and Future Work

This work discusses the impact of environment to inter-vehicle connectivity by applying obstacle propagation model in order to obtain the realistic three-dimensional environment. The various heights of road topology level have shown the different transmission range which required for a realistic three-dimensional case. The z-axis location coordinate is considered as a additional weight value in order to spot the location on the different altitude, thus, it can not be neglected. The relative angle calculation can be further required in order to locate the precise node position.

For further step, it is important to define the detail of obstacle model, since there is a substantial information due to connection opportunity of obstacle types such as building, trees, and other participant vehicles. In addition, buildings have the various shapes which lead to various propagation models. The further investigation of appropriate channel model will be considered in three-dimensional case.

VII. References

[1] M. Mauve, B. Scheuermann, *VANET Convenience and Efficiency Applications*, VANET Vehicular Applications and Inter-Networking Technology, 2010, pp 81-105.

[2] M. Aoki, H. Fujii, *Inter-vehicle Communication: Technical Issues on Vehicle Control Application*, IEEE Communications Magazine, Vol. 34, No. 10, Aug 2002, pp 90 - 93 .

[3] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, M.S. Palma, *A Survey of Various Propagation Models for Mobile Communication*, IEEE Antennas and Propagation Magazine, Vol. 45, No. 3, June 2003, pp 51-82.

[4] J. Walfisch, H. Bertoni, *Theoretical Model of UHF Propagation in Urban Environments*, IEEE Transactions on Antennas and Propagation, Vol. 36, No. 12, Dec 1988, pp 1788-1796.

[5] M. Boban, T. T. V. Vinhoza, M. Ferreira, J. Barros, O. K. Tonguz, I*mpact of Vehicles as Obstacles in Vehicular Ad-hoc Networks*, IEEE Journal on Selected Areas in Communications, Vol. 29, No. 1, Jan 2011, pp 15-28.

[6] N. P. Vaity, D. V. Thombre, *Road Topology-based Performance Analysis of Distance Vector Routing Protocol in VANET*, International Journal of Advance Computational Engineering and Networking (ISSN), Vol. 1, No. 2, Apr 2013, pp 24-31.

[7] A. Hrovat, G. Kandus, T. Javornik, *A Survey of Radio Propagation Modeling for Tunnels*, IEEE Communications Surveys and Tutorial, Vol. 16, No. 2, Oct 2013, pp 658 - 669.

[8] L. Kristiana, C. Schmitt, B. Stiller, *Investigating a Reliable Inter-vehicle Network in a Three Dimensional Environment,* Fachgespräch (GI-IVG), Ulm, Germany, March 2015.

[9] V. Gonzales, A. L. Santos, C. Pinart, F. Milagro, *Experimental Demonstration of the Viability of IEEE 802.11b Based Inter-vehicle Communication*, International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom), Mar 2008.

[10] L. S. Mojela, M. J. Booysen, *On the Use of WiMAX and Wi-Fi to Provide In-vehicle Connectivity and Media Distribution*, IEEE International Conference on Industrial Technology (ICIT), Feb 2013, pp 1353-1358.

[11] V. Naumov, R. Baumann, T. Gross, *An Evaluation of Inter-Vehicle Ad-hoc Networks Based on Realistic Vehicular Traces*, The 7th ACM International Symposium on Mobile Ad-hoc Networking and Computing (MobiHoc), 2006, pp 108-119.

[12] G. Kao, T. Fevens, J. Opatrny, *Position-Based Routing on 3D Geometric Graphs in Mobile Ad-hoc Networks,* The 17th Canadian Conference on Computational Geometry, (CCCG), 2005, pp 88-91.

[13] S. M. N. Alam, Z. J. Haas, *Coverage and Connectivity in Three-Dimensional Networks*, 12th International Conference on Mobile Computing and Networking (MobiCom), Sep 2006, pp 346-357.

[14] A. F. Molisch, F.Tufvesson, J. Karedal, C. F. Mecklenbräuker, *A Survey on Vehicle-to-Vehicle Propagation Channels*, IEEE Wireless Communication, Vol.16, No. 6, Dec 2009, pp 12-22.

[15] S. Durocher, D. Kirkpatrick, L. Narayanan, *On Routing with Guaranteed Delivery in Three-Dimensional Ad-hoc Wireless Networks*, Wireless Networks, Vol. 16, No. 1 Jan 2010, pp 227-235.

[16] F. Huang, K. Leung, V. Li, *Transmission Radius Control in Wireless Ad-hoc Networks with Smart Antennas,* IEEE Transaction on Communications, Vol. 58, No. 8, Jul 2010, pp 2356-2370.

[17] Y. Yoon, J. Kim, M. Jung, Y. Chong, *Radio Propagation Characteristics in the Large City,* 16th International Conference on Advanced Communication Technology (ICACT), Feb 2014, pp 558-562.

[18] R. K. Schmidt, T. Köllmer, T. Leinmüller, B. Böddeker, G. Schäfer, *Degradation of Transmission Range in VANETs caused by Interference,* Praxis der Informationsverarbeitung und Kommunikation, PIK, Vol. 32, No. 4, Jan 2010, pp 224–234.

[19] C. Lochert, M. Mauve, H. Füssler, *Geographic Routing in City Scenarios,* ACM SIGMOBILE Mobile Computing and Communications, Vol. 9, No. 1, Jan 2005, pp 69-72.

[20] B. Karp, H.T. Kung, *GPSR: Greedy Perimeter Stateless Routing for Wireless Network*, International Conference on Mobile Computing and Networking (MobiCom), 2000, pp 243-254.

[21] A.E. Abdallah, T. Fevens, J. Opatrny, *High Delivery Rate Position-based Routing Algorithms for 3D Ad-Hoc Networks*, Computer Communications, Vol. 31, No. 4, Oct 2007, pp 807-817

[22] Q. Lin, C. Li, X. Wang, L. Zu, *A Three-Dimensional Scenario Oriented Routing Protocol in Vehicular Ad-hoc Networks,* Vehicular Technology Conference (VTC), June 2013, pp 1-5.

[23] T. Fevens, G. Kao, and J. Opatrny, *3-D Localized Position-Based Routing With Nearly Certain Delivery in Mobile Ad-hoc Networks*, International Symposium on Wireless Pervasive Computing (ISWPC), Feb 2007, pp 344-349.

[24] T. Islam, Y. Hu, E. Onur, B. Boltjes, J.F.C.M. de Jongh, *Realistic Simulation of IEEE 802.11p Channel in Mobile Vehicle to Vehicle Communication*, 13th Conference on Microwave Techniques (COMITE), 2013, pp 156-161.

[25] L. Kristiana, C. Schmitt, B. Stiller, *Survey of Angle-based Forwarding Methods in VANET Communications*, IFI Technical Report No. 2, 2016, https://files.ifi.uzh.ch/CSG/staff/schmitt/Extern/publications/IFI-2016.02.pdf

[26] NS3, https://www.nsnam.org/, Last visit August 28, 2015.

On Computation and Application of k Most Locally-Optimal Paths in Road Networks

Holger Döbler and Björn Scheuermann
Humboldt-Universität zu Berlin, Berlin, Germany
Email: {holger.doebler, scheuermann}@informatik.hu-berlin.de

***Abstract*—For some applications, e.g. route planning services, it is desirable to answer a point-to-point shortest path query on a road network with a set of alternative paths. We discuss the general requirements for such sets of paths such as shortness, diversity, etcetera. As a measure to rank reasonable alternatives we propose the local optimality ratio, because it implicitly covers all of these requirements. We present an algorithm that computes the k best alternatives in terms of this measure.**

I. Introduction

In every-day life, car drivers use route planning software, either web-based or as part of a dedicated navigation device, to find the optimal route for a pair of origin and destination locations. The objective function that a user seeks to obtain an optimal path for, may depend on the user as well as the situation. Therefore the creators of route planning software provide the user with a set of objective functions to select from. Typical objective functions are path lengths with respect to an arc weight function such as "travel distance", "travel time", or "travel time with the side condition that certain road classes are avoided". But users are often not satisfied with blindly trusting the automatically computed result and like to be presented with some of the suboptimal alternative paths as well in order to have the freedom to choose from a set of paths.

The problem, which paths should reasonably be presented to a user, is not mathematically well-defined, as opposed to the problem of existence and finding of the optimal path with respect to some given arc weights. First we discuss (non-exhaustively) different answers to this vague question found in literature.

Assuming that the user's desired arc weight function is given and is non-negative, we address the question of reasonable alternatives to the minimal cost paths without considering different arc weight functions. One well-known selection criterion for alternative paths is "local optimality," but existing algorithms [1] perform a Boolean test for local optimality on candidate paths, not considering whether one accepted candidate has a higher degree of local optimality than other, including discarded, candidate paths.

Our contribution is to introduce a quantitative measure for paths, the local optimality ratio (LOR), and to present an algorithmic solution to finding the set of k paths with highest LOR for directed graphs with static non-negative arc weights.

The rest of this paper is organized as follows. In Section I-A we give an overview of related work. We discuss use cases of alternative route sets and the associated objectives to search for in Section II. In Section III we explain the notion of local optimality and introduce related quantities. We present our proposed algorithm in Section IV along with an evaluation that indicates that the algorithm is feasible for online queries on urban area sized networks in practice. In Section V we discuss remaining deficiencies of the algorithm and present approaches to alleviate them as well as to extend the algorithm for use in other classes of road networks.

A. Related work

A set of alternative paths can be obtained by optimizing paths with respect to different arc weights, or combinations thereof, or by computing paths that are Pareto-optimal with respect to different arc weights [2], [3]. Instead we focus on alternative paths with respect to a single arc weight function.

The iterative penalty method [4]–[6] as well as the gateway path method [7] produce sets of paths that are short[1] as well as notably different.

The method of "alternative graphs" [8] can be seen as an extension of the iterative penalty method that abolishes some of its deficiencies and efficiently encodes the resultant set of alternative paths in a reduced sub-graph. In [9] is shown how these ideas can be turned into algorithm suitable for interactive use.

The authors of [10] introduce the concept of plateaus as a quantitative measure to rank single via paths, i.e. paths that can be represented as concatenation of two shortest paths. A plateau is maximal path sub-graph, whose nodes as via-nodes induce the same resultant single via path. It is argued, that paths corresponding to long plateaus are preferable, but the explicit selection criterion applied is not specified.

Abraham et al. [1] also focus on single via paths and propose that a sub-optimal path can be considered "admissible" if every subpath up to a certain length is optimal, if the relative stretch of every subpath is bounded by a given constant, and if the sharing with the shortest path is bounded. They propose an algorithm that approximatively tests single via paths for passing these criteria for threshold values that are parameters to the algorithm.

[1]Hereafter, we shall always understand "short" as "optimal with respect to the selected arc weight function."

II. Applications of alternative paths

We will discuss two use cases of alternative paths: giving freedom of choice to users of route planning software and providing a reduced search space for more complex optimization problems. In the bigger part of this paper we will focus on the former application and discuss the latter application only for the sake of additional motivation.

As explained in the introduction, the need for alternative routes arises when presenting point-to-point query results to a user of route planning software. In particular we consider the following requirements to the set of paths being proposed:

All best paths shall be included.
: A user is likely to be interested in each path that is best with respect to the chosen arc weights.

The paths shall be pairwise notably different.
: A user would anticipate that a tiny variation of a "good" path has similar properties in most respects. Therefore such tiny variations need not to be explicitly presented.

All paths shall be "somewhat short."
: In the end the user wants a path that is short with respect to the given arc weight function. Note that this is contrary to the request for notably different paths, since in most cases there are several tiny variations of the shortest path that are much shorter than any notably different alternative.

These properties are intentionally formulated loosely and without an exact definition here. Several concrete methods that generate sets of alternative routes found in literature [1], [4]–[10] do, explicitly or implicitly, respect all of these requirements.

Besides enabling a user to make her own choice, alternative paths can be used as reduced search space in a more complex optimization problem. For example, consider the problem of system optimal route assignment. A stochastic optimization scheme like simulated annealing or a genetic algorithm would subsequently evaluate the objective function for different elements in the search space, giving favor to better route assignments. The search space of all origin-destination paths is usually huge and contains many paths that are either likely to be dismissed by the stochastic optimizer anyway, or undesirable by the user, and therefore do not qualify as parts of practically good solutions.[2] Assuming that the objective function's evaluation is expensive, it would be economic to reduce the search space a priori to a set of reasonable routes. But does a good reduced search space have the same requirements on paths as a good selection to be presented directly to the user? Obviously the requirement of containing all of the shortest paths holds for search spaces, as well as the rather vague requirement that paths should be short. Whether two nearly identical paths should be part of a reduced search space is not that clear, but as we assumed that the objective function is expensive to evaluate, we can argue that if the search space to be explored has to be as small as feasible, more diverse paths would bear a greater potential for optimization.

For now the application as reduced search space is only intended as motivating example and will not be discussed further within this publication.

[2]We assume that a solution to the system optimal route assignment problem is practically good, if the routes are fair enough, that human car drivers are likely to comply.

III. Local optimality

Now we introduce a quantitative measure for paths, with the property that sets of paths with highest values of this measure tend to fulfill the requirements stated above. From Abraham et al. [1] we adopt the notion of local optimality: for a given path, let the "interior" of a path be the sub-path that is gained by removing the first and last edge.[3] A path P is T-locally optimal, if every sub-path whose interior is shorter than T is a shortest path. We define $\ell_{\mathrm{LO}}(P)$ of a path P from s to t (origin and destination node) as the largest number x, such that P is x-locally optimal. Then we define the local optimality ratio (LOR) $q_{\mathrm{LO}}(P) := \ell_{\mathrm{LO}}(P)/\ell(P)$ where $\ell(P)$ denotes the length of P. A shortest path has $q_{\mathrm{LO}}(P) = \infty$ and all other paths have a local optimality ratio in $[0, 1)$, where higher values correspond to more locally-optimal paths. A path with high LOR consists of long shortest sub-paths that are telescoped with long consecutive overlaps.

We claim that paths with high LOR are likely to be relevant to a user, because one can show that

- Shortest paths always have the highest LOR.
- A significant overlap with a high-LOR path leads to a small LOR.

Whether all paths with high LOR are decently short, depends on the prioritization of "short" and "divers". In our experiments we never found stretch factors beyond 2 for paths that were computed, but it is easy to construct malicious examples where an arbitrarily long path is the second-best in terms of LOR. To eliminate this problem one could either introduce a tight bound the total stretch of paths or change the normalization to $q_{\mathrm{LO}}(P) = \ell_{\mathrm{LO}}(P)/\ell(P)^s$ with some $s > 1$ giving favor to shorter paths. Though not evaluated, the latter approach could turn out to be a good tool to balance path length versus path diversity.

IV. Proposed algorithm

We will now roughly sketch the developed algorithm. Given a weighted digraph G, a source s and a destination t, we grow one shortest path tree (SPT) from s and a reversed one from t. The connected components of the intersection of both trees are the plateaus. To each plateau corresponds one single via path. The plateau length yields a lower bound on the LOR [1]: $q_{\mathrm{LO}} \geq d(u, v)/\ell(P)$. Upper bounds for all plateaus can be computed in linear time by traversing each shortest path tree once.

When searching for k-highest LOR paths, we must employ additional shortest path searches to refine the bounds of all candidate paths. We do not need to compute exact values of

[3]The interior of a path consisting of less than tree arcs has length zero.

ℓ_{LO}; we grow just enough SPTs to separate the k best paths from the rest, i.e., until the $(k-1)^{\text{th}}$ highest lower bound is higher than the k^{th} highest upper bound. Deducing lower and upper bounds on $\ell_{\mathrm{LO}}(P)$ from a list of forward SPTs rooted at nodes in P is simple. Let u (v) be the first (last) node of the plateau. Let R be the list of nodes between s and u that are roots of known SPTs, including u as last element, sorted by their occurrence in P, and let's assume that P is not a shortest path, because otherwise $\ell_{\mathrm{LO}}(P) = \infty$ anyway. Let $\hat{P}_x$ be the longest sub-path of P starting at x that is a shortest path. Then $\ell_{\mathrm{LO}}(P) \leq \min_{x \in R} \ell(\hat{P}_x)$.

$$\text{Let } g_P(x,y) := \begin{cases} \min(\ell(\hat{P}_x), \ell(\hat{P}_y)) & \text{if } (x,y) \in P \\ \ell(\hat{P}_x \cap \hat{P}_y) & \text{else} \end{cases}$$

where $(x,y) \in P$ means that (x,y) is one arc of P. Then $\ell_{\mathrm{LO}}(P) \geq \min_{(x,y) \in R} g_P(x,y)$, where $(x,y) \in R$ means the set of pairs (x,y) that lie consecutively in R. Therefore, each additional SPT has the potential to refine the bounds on $\ell_{\mathrm{LO}}(P)$ for one or more candidate paths. One can use properties of the SPTs known at a time to predict which nodes are irrelevant as SPT roots to further refine bounds on LOR, and we use heuristics to predict which nodes are most promising to grow a SPT from, in order to separate the k best paths from the rest. To restrict grown SPTs in size, we use a tight bound of 1.5 on the path's stretch factor during evaluation. Details about heuristics will be given in a follow-up publication. Though our algorithm still requires $O(N)$ SPTs in the worst case, our evaluation on OpenStreetMap data of the metropolitan area of Berlin shows, that we typically require less than $20k$ partial SPTs, scanning approximately $k \cdot \left|\{v \mid d_{sv} + d_{vt} < 1.5 d_{st}\}\right|$ nodes, where d denotes the shortest path distance.

Abraham et al. [1] argued that the computation of local optimality length requires quadratically many shortest path queries. Their objective is to test only whether $\ell_{\mathrm{LO}}(P)$ is above some threshold value and they favored a cheap 2-approximation that may dismiss paths with $T \leq \ell_{\mathrm{LO}}(P) < 2T$ false negatively.

A. Evaluation

For evaluation we used OpenStreetMap data of the area of Berlin. On the largest strongly connected component of the graph we computed arc weights as the geometrical length of an arc divided by the speed limit. That is, we computed an approximation of the free flow travel time but without taking into account delays caused by traffic lights. For $k \in \{2,3,4,5,6,8,10,12,16,20,24\}$ we chose pairs (s,t) uniform at random, neglecting pairs of nodes whose distance is less than 5 minutes, because intuitively if origin and destination are too close together the number of reasonable alternatives becomes small. For 100 (s,t) pairs, and for each value of k we computed the k-highest LOR single via paths and recorded the number of SPTs needed and the number of nodes scanned during all SPT growths. Afterwards, we computed exact values of LOR for each path that was returned.

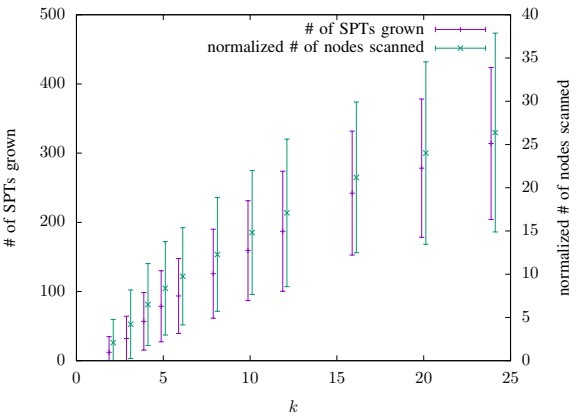


Fig. 1. Average number of SPTs grown and average normalized number of nodes scanned.

In our implementation (C++11, GCC 5.3.0), each query for k alternative paths took less than 50ms on a single core of a Intel(R) Core(TM) i5-4210U CPU, which can be considered fast enough for online queries in route planning software. We varied the hard limit on paths' stretch and found that even for values greater than 1.5, paths with stretch > 1.5 where rarely returned. Even without any stretch bound, paths with stretch > 2 where never returned.

In Fig. 1 we show that the average number of SPTs that where grown to compute the k-highest LOR paths. The figure also shows the average ratio of number of nodes scanned in all consecutive SPT searches and the number of nodes whose via-path-length is less than 1.5 times the source-destination-distance. The latter shows that indeed the average number of nodes needed to scan to find k-highest LOR paths is only about a factor k as large as the number of nodes scanned to find the shortest path. Together with the large number of SPTs grown, typically more than $10k$, this means that the SPTs grown are on average much smaller than the initial two SPTs. In Fig. 2 we show the average stretch of alternative paths depending on their LOR rank. The stretch saturates quickly for ranks greater than 2 at a value of about 1.2 but with a huge standard deviation. In Fig. 3 we show the paths' average LOR depending on the rank.

V. Conclusion and Outlook

We have extended the concept of locally-optimal paths by introducing the quantitative measure of LOR and argued, why paths with a high LOR are good candidates for alternative routes in road networks. We proposed an algorithm that computes k best single via paths in terms of LOR in graphs with static non-negative arc weights. Despite its super-quadratic worst case run-time we have demonstrated in our evaluation that it is still feasible for alternative paths online queries in urban area sized road networks.

The proposed algorithm relies on single via paths that are constructed using bidirectional SPTs. Therefore non-single-

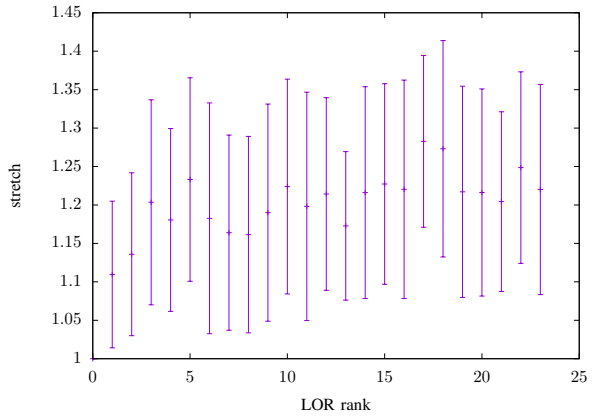


Fig. 2. Average stretch of paths found, depending on LOR rank.

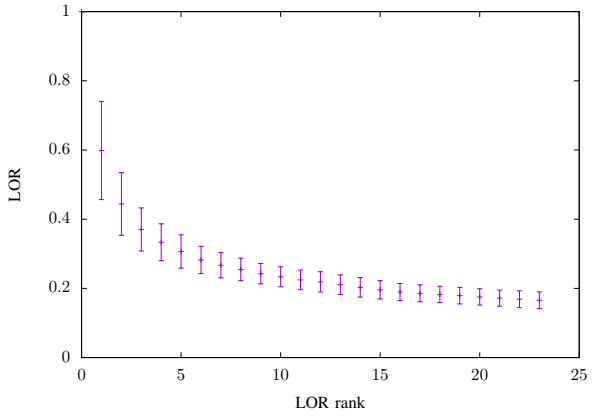


Fig. 3. Average LOR of paths found, depending on LOR rank.

via[4] paths will never be found by the algorithm. One can show that a non-single-via path's LOR cannot exceed a certain constant value, but the number of single via paths is finite, since it is bounded by the number of nodes. Since the LOR of single via paths can be arbitrarily small (typically many single via paths have $q_{\mathrm{LO}}(P) = 0$), it follows that there are non-single-via paths with a higher LOR than the n^{th} best single via path, for some finite n. One has grow additional SPTs in order to construct 2-via paths at all, and even more trees in order to classify them in terms of LOR.

But generalizing the above algorithm to an unconstrained search space is not the only motivation for stacking forward shortest path trees. In a graph where both the arc weights and the arc traversal times are time-dependent, single source shortest path search is still possible, if waiting at nodes is allowed. But since the arrival time at the destination node differs across paths, the bidirectional Dijkstra approach is not feasible. Obviously one possible approach is to avoid the need for a single destination backward SPT and use only telescoped forward shortest path searches just as in the static case of non-single-via paths.

Another topic that we are very interested in is the empirical evaluation of users' preferences for alternative paths, because as far as we know, the question of which properties qualify a path as a relevant alternative for a typical driver is only answered in terms of reasoning and speculation, just as we did it here. Obviously this is far out of scope for computer scientists, but project partners from the Department of Psychology are willing to design and perform experiments to answer these questions.

[4]A non-single-via path shall be a path that cannot be represented as concatenation of two shortest paths.

References

[1] I. Abraham, D. Delling, A. V. Goldberg, and R. F. Werneck, "Alternative Routes in Road Networks," *J. Exp. Algorithmics*, vol. 18, pp. 1.3:1.1–1.3:1.17, Apr. 2013.

[2] E. Q. V. Martins, "On a multicriteria shortest path problem," *European Journal of Operational Research*, vol. 16, no. 2, pp. 236–245, May 1984.

[3] M. Müller-Hannemann and K. Weihe, "Pareto Shortest Paths is Often Feasible in Practice," in *Algorithm Engineering*, ser. Lecture Notes in Computer Science, G. S. Brodal, D. Frigioni, and A. Marchetti-Spaccamela, Eds. Springer Berlin Heidelberg, Aug. 2001, no. 2141, pp. 185–197.

[4] P. E. Johnson, D. S. Joy, D. B. Clarke, and J. M. Jacobi, "Highway 3. 1: An Enhanced Highway Routing Model: Program Description, Methodology, and Revised User's Manual," Oak Ridge National Lab., TN (United States), Tech. Rep. ORNL/TM-12124, Mar. 1993. [Online]. Available: http://www.osti.gov/scitech/biblio/6549364

[5] V. Akgün, E. Erkut, and R. Batta, "On finding dissimilar paths," *European Journal of Operational Research*, vol. 121, no. 2, pp. 232–246, Mar. 2000.

[6] Y. Chen, M. G. H. Bell, and K. Bogenberger, "Reliable Pretrip Multipath Planning and Dynamic Adaptation for a Centralized Road Navigation System," *Trans. Intell. Transport. Sys.*, vol. 8, no. 1, pp. 14–20, Mar. 2007.

[7] K. Lombard and R. L. Church, "The gateway shortest path problem: generating alternative routes for a corridor location problem," *Geographical Systems*, vol. 1, pp. 25–45, 1993.

[8] R. Bader, J. Dees, R. Geisberger, and P. Sanders, "Alternative Route Graphs in Road Networks," Rome, Apr. 2011.

[9] M. Kobitzsch, M. Radermacher, and D. Schieferdecker, "Evolution and Evaluation of the Penalty Method for Alternative Graphs." in *ATMOS*, ser. OASICS, D. Frigioni and S. Stiller, Eds., vol. 33. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013, pp. 94–107.

[10] Cambridge VehicleInformation Tech. Ltd., "Choice Routing," 2005. [Online]. Available: http://www.camvit.com/camvit-technical-english/Camvit-Choice-Routing-Explanation-english.pdf