Simulation of Privacy-Enhancing Technologies in Vehicular Ad-Hoc Networks

—

Simulation von Privatsphärenschutz in Fahrzeug-Ad-hoc-Netzen

Der Technischen Fakultät

der Friedrich-Alexander-Universität

Erlangen-Nürnberg

zur

Erlangung des Doktorgrades

DOKTOR-INGENIEUR

vorgelegt von

David Eckhoff

aus Dresden

Als Dissertation genehmigt
von der Technischen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg

Tag der mündlichen Prüfung:	**01. März 2016**
Vorsitzender des Promotionsorgans:	**Prof. Dr. Peter Greil**
Gutachter:	**Prof. Dr.-Ing. habil. Reinhard German**
	Prof. Dr. rer. nat. Björn Scheuermann

Abstract

Equipping vehicles with communication technology is a promising approach to increase both safety and the efficiency of tomorrow's road traffic. However, without proper privacy protection, such a communication system can be exploited to compromise drivers' location privacy or to install fully automated overbearing traffic surveillance. In order to deploy effective Privacy-Enhancing Technologies (PETs), it is not only important to understand the concrete privacy risks that go along with vehicular networks, but also to be able to measure the level of privacy provided by the system.

This thesis contributes to privacy research by providing a risk analysis, a taxonomy for privacy in vehicular networks, and a review of the state of the art in privacy research. We further address shortcomings and potentials of simulation techniques and make recommendations to improve the quality and meaningfulness of privacy simulation. Based on our findings, we develop an open-source privacy simulation framework that allows evaluation of the level of location privacy enjoyed by drivers. Combined with detailed models for American and European communication standards, we provide a powerful tool not only for the analysis of packet-based privacy protection mechanisms, but also to identify performance issues of the envisioned communication protocols.

Using our simulator, we develop and evaluate different PETs that address open research topics: We introduce SlotSwap, a time-slotted pseudonym exchange scheme which protects against privacy violations by the system provider. Time-slotted pseudonyms also protect from Sybil attacks and complicate tracking by simultaneously changing identifiers. Our certificate revocation system SmartRevoc also makes use of this technology and offers an efficient and backward privacy-preserving revocation method. We show that parked vehicles can support the timely distribution of revocation lists and also considerably improve traffic safety. Lastly, we present a robust fingerprinting attack exploiting IEEE 802.11 scramblers that illustrates that one non-privacy-aware component can compromise privacy throughout the entire system. Based on our results we draw conclusions for the design of PETs in future transportation systems.

Kurzfassung

Einer der derzeit vielversprechendsten Ansätze, um zukünftig Verkehrssicherheit und -effizienz zu erhöhen, sind miteinander kommunizierende Fahrzeuge. Ohne effektiven Privatsphärenschutz kann ein solches Kommunikationssystem allerdings missbraucht werden, z.B. um Fahrer auszuspähen oder ein vollautomatisches Verkehrsüberwachungssystem aufzubauen. Doch lässt sich ein solcher Schutz nicht entwickeln, ohne vorher die konkreten Risiken in Fahrzeugnetzen zu verstehen und den Grad der Privatsphäre eines Fahrers messbar zu machen.

In dieser Arbeit präsentieren wir daher zunächst eine tiefgehende Risikoanalyse, sowie eine Taxonomie für Privatsphäre in Fahrzeugnetzen und einen Überblick über den aktuellen Stand der Privatsphärenforschung. Wir bewerten verbreitete Simulationsmethodiken hinsichtlich Qualität und Aussagekraft und zeigen mit der Entwicklung eines Simulationswerkzeuges, wie sich diese steigern lassen. Das entwickelte Framework erlaubt die detaillierte Bewertung von Privatsphäre in Fahrzeugnetzen. Zusätzlich ermöglicht eine genaue Abbildung des amerikanischen und europäischen Kommunikationsstandards nicht nur die Untersuchung von Schutzmechanismen auf Paketebene, sondern auch das Aufdecken von Performanzproblemen.

Unser Simulator erlaubt uns darüber hinaus eigene Beiträge zum Schutz der Privatsphäre in Fahrzeugnetzen zu entwickeln und zu bewerten: Gegen Privatsphärenverletzungen durch den Betreiber schützt unser pseudonym-austauschendes SlotSwap-System. Die Benutzung zeitbasierter Pseudonyme schützt außerdem vor Sybilangriffen und der damit einhergehende synchrone Addresswechsel erschwert das Verfolgen von Fahrzeugen. Zeitbasierte Pseudonyme sind auch die Basis unseres Zertifikatsperrsystems SmartRevoc, das sich fehlverhaltende Fahrzeuge effizient sperren kann, ohne deren Vergangenheit offenzulegen. Des Weiteren zeigen wir, dass parkende Fahrzeuge entscheidend zur Verkehrssicherheit beitragen und Zertifikatsperrlisten schnell im Netzwerk verteilen können. Zum Abschluss präsentieren wir einen robusten Fingerprinting-Angriff auf IEEE 802.11-Geräte, dessen Tragweite wir mit Hilfe unseres Simulators quantifizieren können. Die Gesamtheit der Ergebnisse dieser Arbeit ermöglicht eine differenzierte Einschätzung und eine gezieltere Entwicklung von Mechanismen zum Privatsphärenschutz in Fahrzeugnetzen.

Contents

Chapter 1

Introduction

In 2014 the number of traffic accidents on German roads exceeded 2.4 million, resulting in 392 000 injured persons and a total of 3377 fatalities [227]. In Europe, a total of approximately 25 700 persons were killed in traffic-related accidents and more than 200 000 "came home with life-changing, serious injuries" [73]. It has long been a goal of many (both national and pan-European) projects (e.g., Vision Zero) to reduce this number and eventually achieve a traffic system with no serious injuries or fatalities [233].

An important step towards this goal is the improvement of the vehicles themselves: Increasing the rigidity of the coachwork and adding or improving airbag and seat belt systems all contributes to safer traffic. However, the potentials of these passive safety systems seem to be practically exhausted [251]. Focusing on accident prevention, there exist numerous active safety systems such as anti-lock braking, electronic stabilization control, head-up displays, and in-car sensors for collision warning or avoidance, just to name a few. In order to work properly, these systems, that is, the Electronic Control Units (ECUs), are connected over various in-car network bus systems. A modern vehicle comes with more than 120 ECUs with approximately 3 km of cables connecting them [215]. It is therefore safe to say that today's vehicles are complex, mobile computer networks.

Since almost all components of a vehicle are somehow connected, the next logical step is to also interconnect the vehicles with each other and the infrastructure (e.g., traffic lights) to form a so-called Intelligent Transportation System (ITS). Equipping vehicles with an ECU which is connected to the in-car network and a radio transceiver allows for wireless information exchange and enables many applications to increase traffic safety. For example, a vehicle can then periodically emit a beacon message containing information such as its current speed, position, or heading. This allows a vehicle receiving this information to develop a situational awareness of its surroundings including all vehicles in its vicinity, compute their paths and

thereby prevent collisions. In addition, vehicles can inform others of, e.g., hazardous road conditions, emergency braking, or tail ends of traffic jams so the driver of the receiving vehicle can be warned and an accident avoided.

Road safety is without a doubt the most important argument for the introduction of inter-vehicle communication. However, there also exists a wide range of non-safety applications, which are strongly promising to increase traffic efficiency such as Green Light Optimal Speed Advisory (GLOSA) [24, 25, 59] or traffic information systems [58, 194, 256], in which vehicles inform others or Roadside Units (RSUs) of their current knowledge about the traffic situation. Furthermore, wireless vehicle networks can be used for comfort applications (e.g., wireless payment systems [123], video streaming [4], or even multiplayer games [236]) to provide entertainment, or as a provider for contextual information (e.g., a parking space finder [175] or location-based services [6]).

In Europe and North America, wireless ad-hoc communication between vehicles (and infrastructure), also referred to as Car-to-X communication (or Inter-Vehicle Communication (IVC)), is based on IEEE 802.11p, an amendment to the well-established IEEE 802.11 Wireless LAN [121]. This decentralized communication takes place in the 5.9 GHz band and allows vehicles to exchange information with a line-of-sight communication range of about 500 m to 1000 m, with low latency and without the need for infrastructure. This allows the realization of the aforementioned safety applications, but trying to cover larger distances via multi-hop communication can introduce high latencies and limited throughput, e.g., when used in traffic information systems [202]. Other projects have therefore examined the applicability of cellular communication [164, 223] instead of WLAN, where throughput and connectivity is less of a problem than latency [253]. While both have advantages and disadvantages, the current trend is to assume heterogeneous networks, that is, vehicles equipped with both a cellular (e.g., UMTS or LTE) and an IEEE 802.11p transceiver. The former would then be used for centralized services, while the latter supports low latency safety applications or locally bound information exchange.

The periodic and unencrypted broadcast transmission of safety beacons [84, 196] via IEEE 802.11p raises privacy concerns [63]. These messages can be received by anyone in the vicinity, be it other vehicles, infrastructure nodes, or by arbitrary persons with compatible hardware. The fact that off-the-shelf consumer hardware can be used to receive messages from vehicles intensifies this problem and, with high enough coverage, allows for the detailed tracking of vehicles through the road network, as illustrated in Figure 1.1: Even if only a few beacons are received, they can reveal the path a vehicle took if the receiver is able to relate the beacons to the same vehicle. This might not only have a negative impact on the market acceptance of ITS devices (and thereby the penetration rate), but will have serious consequences for

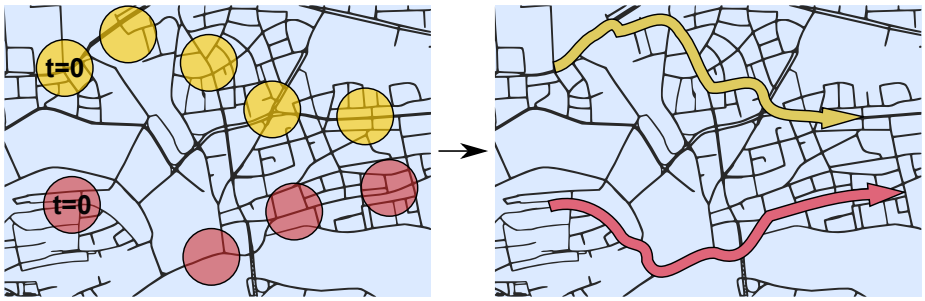


Figure 1.1 – Overhearing beacons emitted by vehicles in different locations (starting at t=0, from left to right) can reveal their track and compromise drivers' location privacy.

drivers' location privacy, especially if these systems become mandatory as currently discussed [189]. This would lead not only to a complete compromise of drivers' privacy but also make it possible for law enforcement to install a fully automated traffic surveillance system that uses periodic messages sent by vehicles to identify even the smallest traffic violations [63].

Raising awareness of these issues is not a trivial task as privacy still seems to be a somewhat nebulous concept [246]. Even though privacy challenges in vehicular networks have been identified early [49, 100], all-encompassing privacy protection mechanisms have not found their way into the standardization of either the American or the European systems [63]. One reason for this is that location privacy cannot be measured easily, that is, it is difficult to put a number on the level of location privacy of a driver. The state of the art is to do so by means of simulation, as although analytical approaches can in fact give valuable insights when it comes to lower and upper bounds, they often require substantial simplifications to still be manageable (e.g., when it comes to modeling both network and road traffic). Although Field Operational Tests (FOTs) are required to validate analytical or simulation models and are helpful to find new aspects that were simply not thought of before, they are too cost and time-intensive to explore the parameter spaces needed to evaluate Privacy-Enhancing Technologies (PETs) in vehicular networks.

To be able to make an informed choice for a PET to be deployed in future ITS's, one must be able to compare them and reproduce results presented in the literature. Unfortunately, the vast variety of metrics, tools, scenarios, and methodology in vehicular privacy research does not allow this [246]. Current revelations on data privacy violations worldwide show that privacy protection must be a strong component in ITS's from the very beginning, as retrofitting it is almost impossible [63]. This thesis tries to improve on this situation by building a privacy simulation framework with a focus on the detailed network communication models required for the evaluation of PETs. As we will show, these models are not only necessary for location privacy

research, but can give valuable insights in understanding and identifying current shortcomings in both the American and European ITS proposals.

We show the usefulness of our framework by presenting and evaluating new PETs where we believe the current state of the art needs to be extended. This ranges from pseudonym changing strategies in vehicular networks, to revocation of, e.g., stolen or misbehaving vehicles, and even to physical layer fingerprinting attacks. Our proposals aim at raising the awareness of privacy issues in vehicular networks, showing solutions for their evaluation, addressing current shortcomings, and making privacy protection both more effective and more efficient. We believe that this is required to make ITS's, and especially the vision of communicating vehicles, a reality that positively contributes to our everyday life while respecting the freedom and privacy of all people in it.

Chapter 2

Fundamentals

This chapter serves both as a motivation for this thesis as well as an introduction of the fundamentals upon which our work is based. First, we discuss the basic principles of vehicular communication (Section 2.1). We show why location privacy is at risk when this technology becomes a reality, why privacy protection has to be an inherent part of any future ITS, and what challenges have to be overcome in order to achieve this (Section 2.2).

In Section 2.3 we give an overview on how discrete-event simulation can be used as a tool to evaluate the performance of vehicular networking applications. This includes both the simulation of network traffic as well as road traffic, as these are the most important components for the simulation of vehicular networks.

We then discuss the current state of the art in the simulation of vehicular PETs based on an extensive literature review (Section 2.4). We identify the shortcomings and drawbacks of current vehicular network privacy research and explain how our work can help improve this situation.

Parts of this chapter are based on our articles published in *IEEE Security & Privacy* [63], *IEEE Vehicular Technology Magazine* [53], *GI Praxis der Informationsverarbeitung und Kommunikation* [216], *Elsevier Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment* [64], and on papers published at conferences and workshops [21, 55, 61, 65, 218, 246].

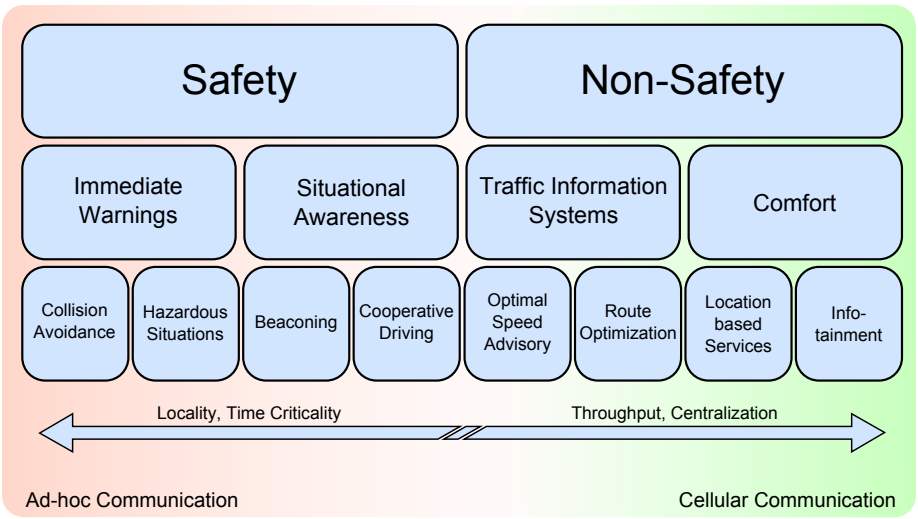


Figure 2.1 – A taxonomy of vehicular network applications based on [212]. Their conflicting requirements emphasize the need for a heterogeneous solution, such as the use of both ad-hoc and cellular communication.

2.1 Vehicular Networks and Standardization

Intelligent Transportation Systems have been a concept for at least 75 years, when visions presented at exhibitions and in magazines [14] included remote controlled highways and even the exchange of information between vehicles to improve both traffic efficiency and safety. Of course, these ideas were well ahead of their time and the technology to make these concepts a reality was yet to be invented, but the need for so-called smart highways and information exchange between all traffic participants was already understood.

One of the first wireless vehicular network systems dates back as far as 1983, when Volkswagen introduced the Wolfsburger Welle [261], a Green Light Optimal Speed Advisory (GLOSA) system that allowed drivers to adjust their velocity in order to avoid stopping at red lights. The system used infrared communication to transmit information from the traffic light to the vehicle. Infrared communication was also used in the 1989 system LISB, a field trial where infrastructure nodes informed vehicles of current traffic and gave them route guidance using infrared beacons. Successors include Euro-Scout [211] and Ali-Scout [234], a 1992 Siemens navigation system that used wireless communication to transmit the best possible route to participating vehicles. However, the technical difficulties and low market acceptance led to the cancellation of all these projects.

From the early 2000s on, the number of field trials and projects in the context of vehicular networks grew drastically [215]. The main reason was the availability of

more reliable communication based on cellular networks such as UMTS or LTE, but also the possibility to exchange information in an infrastructure-less fashion using IEEE 802.11p. Other field operational tests (such as AKTIV [222], sim^{TD} [58, 228], PATH [208], etc.) showed the applicability of vehicular networks and opened the door for various applications. To give an example, the ideas of the Wolfsburger Welle returned under the name Travolution [26], now based on IEEE 802.11p ad-hoc 5.9 GHz WLAN communication.

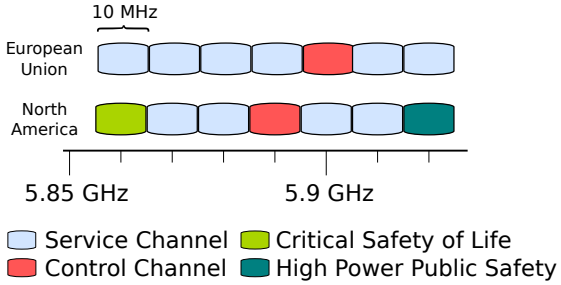


Figure 2.2 – Channels reserved for wireless ITS communication in North America [121] and Europe [82].

Figure 2.1 shows a taxonomy of these applications alongside an indicator of which communication paradigm fits their requirements best. Generally it can be said that the locality, the time constraints, and the required throughput determine whether an application is best achieved using cellular or ad-hoc communication. With decreasing time criticality and an increasing information range, the use of cellular communication can introduce various benefits, e.g., the possibility to directly obtain information from a traffic information center to optimize routes. Vehicles then do not rely on other equipped vehicles in their vicinity but can utilize their own internet connection to access different comfort applications. Safety (such as emergency braking warnings) and awareness applications (a vehicle becoming aware of other vehicles in its vicinity via periodic beacon messages broadcast) are both local and time-critical and are therefore best realized using IEEE 802.11p ad-hoc communication, also referred to as IVC or Dedicated Short-Range Communication (DSRC), as standardized in Europe and North America [78, 122, 189]. The resulting network is referred to as a Vehicular Ad-Hoc Network (VANET) and, as it is this periodic ad-hoc communication that causes privacy issues, it will be the main focus of this thesis.

The U.S. FCC and the European ECC allocated space in the 5.85 GHz spectrum for DSRC, that is, short-range to medium-range wireless vehicular communication. In this band, IEEE and ETSI reserved different 10 MHz wide channels to be used for various ITS applications. One channel is the designated Control Channel (CCH) and four channels serve as Service Channels (SCHs) as illustrated in Figure 2.2. The

control channel is used to broadcast status beacons (speed, position, etc.), emergency messages, and to advertise applications (e.g., traffic updates, traffic light information, etc.) offered on the service channels.

These channels are the basis for ad-hoc-based ITS communication systems in both North America and Europe. In North America, the system is referred to as IEEE WAVE while in Europe it is called ETSI ITS-G5. They both can be understood as a family of multiple standards defining the operation from physical to application layer, and also include cross-layer aspects such as security or management. Although the lower layers of both systems are similar (which is the goal of the harmonization task group [85]), there exist various differences from the Medium Access Control Layer (MAC) and up.

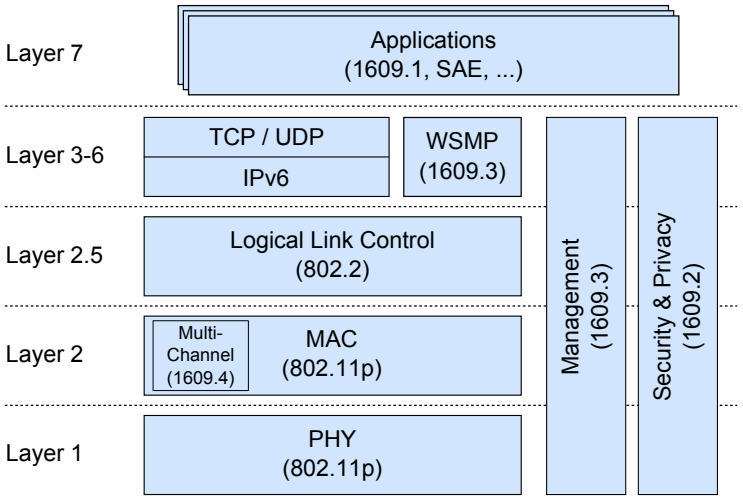


Figure 2.3 – The IEEE WAVE family of standards, consisting of multiple standards from different working groups and institutions.

2.1.1 IEEE WAVE

Figure 2.3 shows the IEEE WAVE (Wireless Access in Vehicular Environments) protocol stack, mainly consisting of the IEEE 1609 family of standards, accompanied and extended by different IEEE, SAE, and IETF documents. Channel access is controlled by IEEE 802.11p and IEEE 1609.4, where the SCH and CCH coordination is handled. This stack supports applications that are either based on IPv6 or the Wave Short Message Protocol (WSMP). Management, security, and privacy features are cross-layer mechanisms.

For the simulation-based evaluation of PETs in vehicular networks in this thesis we identify the most relevant standards in IEEE WAVE:

- **IEEE 802.11p** - *Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments* [121]: This standard will be the basis for models presented in Chapter 3, as it directly impacts the transmission of messages and therefore the possibility to overhear communication and violate users' privacy.

- **IEEE 1609.4** - *Multi-Channel Operations* [120]: This is an extension to the WAVE MAC that defines quality of service strategies and multi-channel operations such as alternating access. By building a detailed simulation model, we identified weaknesses in the protocols described in this document (details and results can be found in Section 3.2).

- **IEEE 1609.2** - *Security Services for Applications and Management Messages* [122]: This standard defines security and privacy measures in WAVE and, as we will show in Section 2.2, has several shortcomings. In Chapter 4 we present PETs that could extend this standard.

- **SAE J2945.1-2.2** - *DSRC Message Communication Minimum Performance Requirements: Basic Safety Message for Vehicle Safety Applications* [196]: This document defines the periodic transmission of safety beacons. As these un-

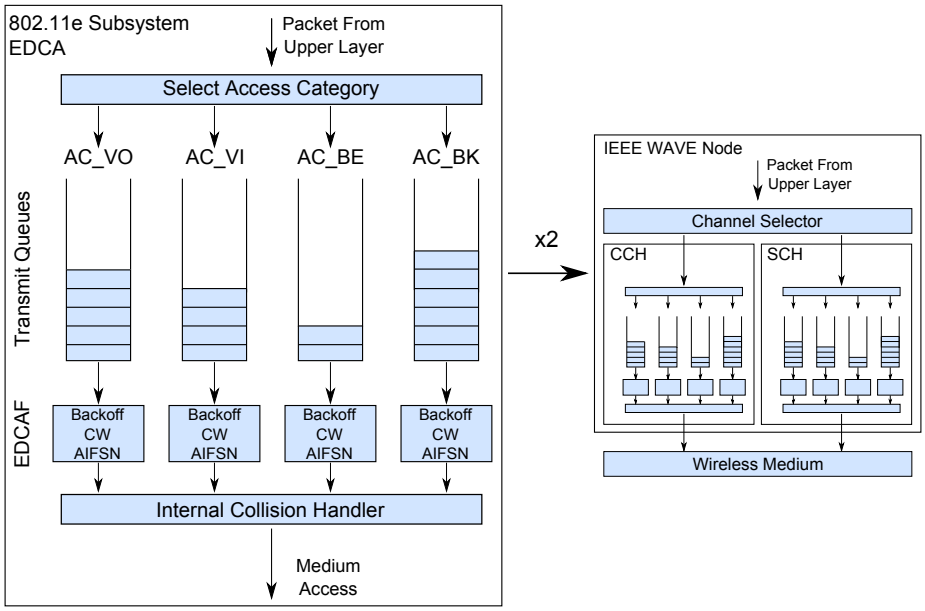


Figure 2.4 – IEEE 802.11e subsystems for IEEE WAVE-enabled vehicles. Two parallel systems are used for CCH and SCH packets, respectively. Packets compete internally and externally.

encrypted beacons contain a vehicle's position and potentially identifying information it has a significant impact on the location privacy of drivers.

Enhanced Distributed Channel Access (EDCA)

Quality of service is realized using two IEEE 802.11e subsystems (more precisely EDCA systems), one for the CCH and another one for the SCHs. This is illustrated in Figure 2.4. Packets coming from an upper layer are mapped to a certain access category, that is, either access category voice (AC_VO), video (AC_VI), best effort (AC_BE), or background (AC_BK). Each access category is represented by a transmit queue and an Enhanced Distributed Channel Access Function (EDCAF) which maintains a backoff counter and has access category-specific values for contention window size and Arbitration Interframe Spacing (AIFS), which equals Arbitration Interframe Spacing Number (AIFSN) · SlotTime + Short Interframe Spacing (SIFS). This ensures that packets from higher access categories are prioritized over packets from lower ones, not only because their contention windows are smaller but primarily because of their shorter AIFS.

Parameter	AC_BK	AC_BE	AC_VI	AC_VO
CW_{min}	aCW_{min}	aCW_{min}	$\frac{aCW_{min}+1}{2}-1$	$\frac{aCW_{min}+1}{4}-1$
CW_{max}	aCW_{max}	aCW_{max}	aCW_{min}	$\frac{aCW_{min}+1}{2}-1$
AIFSN	9	6	3	2

Table 2.1 – Standard parameters for the different ACs [121].

Parameter	Slot length	SIFS	aCW_{min}	aCW_{max}	Bandwidth
Value	13 µs	32 µs	15	1023	3 Mbit/s to 27 Mbit/s

Table 2.2 – Standard settings for IEEE WAVE [120, 121].

Following [121] an EDCAF will go into backoff mode when:
(a) a frame is requested to be transmitted, the backoff counter for that queue was 0 and the channel was busy,
(b) a packet from this queue was transmitted successfully,
(c) the transmission of a packet failed (i.e., no ACK was received), or
(d) a packet from a higher access category was ready to be sent at the same time.

The number of slots a transmit queue will back off is determined by drawing a random number from the interval [0;CW]. The value of CW is left unchanged in case (a), and, in cases (c) and (d), doubled or set to CW_{max}, depending on which value is smaller; in case (b) it is set back to CW_{min}. The backoff counter is only reduced

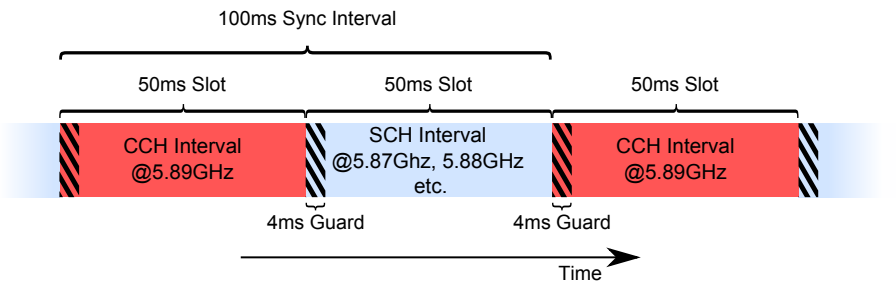


Figure 2.5 – Channel access in IEEE 1609.4 to support multi-channel transmission using a single radio. In alternating mode, a host can tune to one of the services channels in the second 50 ms of the sync interval. In continuous mode, the host will remain tuned to the CCH.

at slot boundaries, i.e., after one AIFS when the channel turned idle and after each consecutive slot. An overview of the access category-specific parameters is shown in Table 2.1; other parameters for IEEE 1609.4 and IEEE 802.11p are listed in Table 2.2.

Alternating Access

An important feature of the upcoming IEEE WAVE standard is the multi-channel operation specified in IEEE 1609.4 [120]: To allow single radio nodes to transmit on multiple channels, an alternating access scheme is proposed. Figure 2.5 shows the working principle. At the beginning of a 100 ms sync interval a host is always tuned in to the CCH to send and receive periodic safety messages. This channel is also used to advertise services offered on one of the SCHs which a host can then tune into during the second 50 ms of the sync interval. This means that the host will be unable to receive and send safety messages in this time and is limited to sending and receiving messages belonging to applications using this service channel. Between CCH and SCH there is a 4 ms guard interval in which the channel is treated as busy. This is done to accommodate timing inaccuracies and to give the chip time to change its transmission frequency. We will show in Section 3.2.3 that alternating access introduces synchronization effects and thereby packet loss. These problems contributed to the fact that, at the time of writing, it is unclear whether this access scheme will be in the final system.

Periodic Safety Messages

Each vehicle is envisioned to send periodic safety messages (or beacons), informing other vehicles in its vicinity about its current state, including speed, heading, position, and so on. These messages are called Basic Safety Messages (BSMs) and are standardized in SAE J2945.1-2.2 [196]. A valid message has to include, but is not limited to, the GPS coordinates, the elevation, the speed, the heading, an indication

whether the vehicle is braking, and the dimensions of the vehicle in terms of width and length. These messages are sent on the CCH with a default frequency of 10 Hz. Vehicles receiving messages from other vehicles can use them as input for safety applications, e.g., collision avoidance or headway warnings.

Although the SAE Message Directory SAE J2735 [195] defines other messages (e.g., traffic light timing messages, vehicle and roadside alerts, etc.), they are not of relevance for this thesis and will not be explained in detail.

2.1.2 ETSI ITS-G5

A simplified view of the architecture of ETSI ITS-G5 is shown in Figure 2.6. Like IEEE WAVE, ETSI ITS-G5 uses an IEEE 802.2 Logical Link Control (LLC) on top of an IEEE 802.11p MAC and Physical Layer (PHY). The network and transport layers are similar, but instead of having applications directly on top of them, ETSI ITS-G5 uses a facility layer for the most important message types such as the Cooperative Awareness Message (CAM), Decentralized Environmental Notification Message (DENM), Signal Phase and Timing Message (SPAT), and Road Topology Message (TOPO). In addition to management, security, and privacy features, ITS-G5 incorporates another cross-layer mechanism called Decentralized Congestion Control (DCC). This is one of the main differences between the European and American systems as DCC heavily affects channel access and message generation, and thereby also network topology.

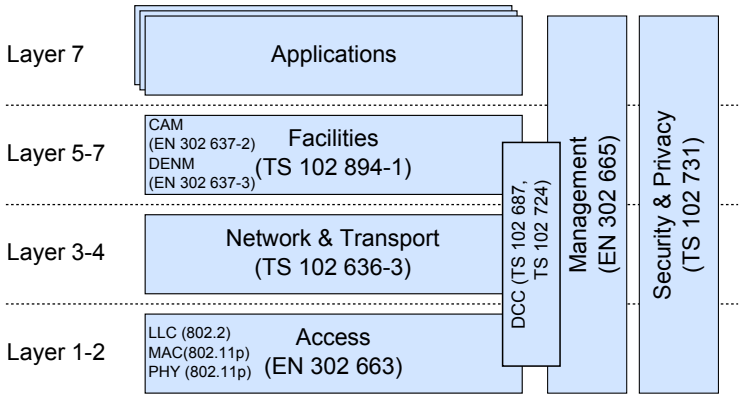


Figure 2.6 – A simplified view of the ETSI ITS-G5 family of standards.

For this thesis, the following documents of the ETSI ITS-G5 family are of particular relevance:

- **ETSI 102731-v1.1.1** - *Intelligent Transport Systems (ITS); Security; Security Services and Architecture* [75]: This standard specifies mechanisms for secure and private communication in vehicular networks and can be seen as the basis for all following security and privacy-related ETSI standards.

- **ETSI 102893-v1.1.1** - *Intelligent Transport Systems (ITS); Security; Threat, Vulnerability and Risk Analysis (TVRA)* [76]: This document lists goals and challenges for security and privacy mechanisms.
- **ETSI 102940-v1.1.1** - *Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management* [80]: This standard focuses on identifying the required security-related entities and their relationships. It also lists security parameters for certificate management and other Public Key Infrastructure (PKI) processes.
- **ETSI 102941-v1.1.1** - *Intelligent Transport Systems (ITS); Security; Trust and Privacy Management* [81]: In this standard the Public Key Infrastructure (PKI) of ITS-G5 is described, including certificate enrollment and pseudonymous communication. This can be seen as the counterpart to IEEE 1609.2.
- **ETSI 302637-2-v1.3.0** - *Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service* [84]: This standard, like SAE J2945.1-2.2 in WAVE, defines the periodic broadcasting of vehicle status information.
- **ETSI 102687-v1.1.1** - *Intelligent Transport Systems (ITS); Decentralized Congestion Control Mechanisms for Intelligent Transport Systems operating in the 5 GHz range; Access layer part* [79]: This document describes the DCC (Decentralized Congestion Control) mechanism which controls channel access of all vehicles. Shortcomings and potentials of this system are discussed in Section 3.3.

Channel Access

Channel access, just like in IEEE WAVE, is based on two EDCA subsystems on top of IEEE 802.11p with a slightly different parametrization of the EDCAFs (listed in Table 2.3). This brings the ACs closer together and decreases starvation effects as maximum contention window sizes are smaller compared to IEEE WAVE.

ITS-G5 does not make use of the alternating access scheme, that is, the radio does not periodically switch between CCH and an SCH. Not listening to the CCH at all would severely affect safety applications as no periodic message would be sent to or received from other vehicles. Vehicles with one ITS-G5 transceiver are therefore expected to continuously listen on the CCH, while vehicles with additional transceivers are free to use any SCHs at the same time [215].

Decentralized Congestion Control

One of the most significant differences between IEEE WAVE and ETSI ITS-G5 is a mechanism called Decentralized Congestion Control (DCC) that aims at keeping

Parameter	AC_BK	AC_BE	AC_VI	AC_VO
CW_{min}	CW_{min}	$\frac{CW_{min}+1}{2}-1$	$\frac{CW_{min}+1}{4}-1$	$\frac{CW_{min}+1}{4}-1$
CW_{max}	CW_{max}	CW_{min}	$\frac{CW_{min}+1}{2}-1$	$\frac{CW_{min}+1}{2}-1$
AIFSN	9	6	3	2

Table 2.3 – ETSI ITS-G5 standard parameters for Access Categories (ACs) [82]. Differences to IEEE WAVE are marked in red.

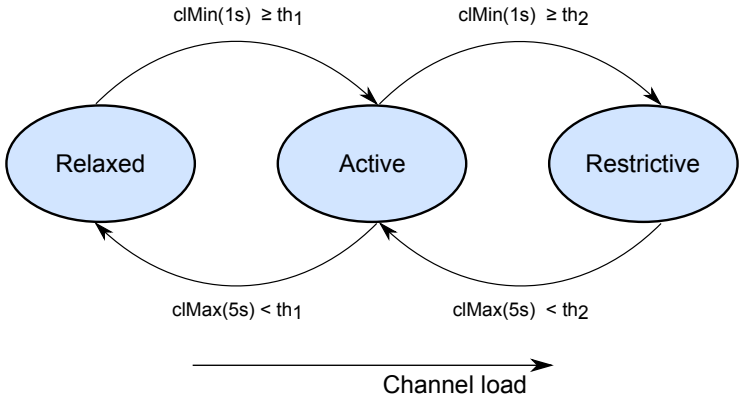


Figure 2.7 – The DCC state machine for the CCH. The thresholds th_1 and th_2 are configurable; the standard suggests default values of $th_1 = 15\,\%$ and $th_2 = 40\,\%$.

the channel load at a level where reliable operation of ITS-G5 safety applications can be achieved [79]. High channel load and congestion can lead to packet loss and higher latencies, e.g., a vehicle may not get updates from a close-by vehicle for a certain amount of time. DCC tries to control the usage of the wireless medium by adjusting MAC and PHY parameters based on the currently observed channel load. The channel load can be understood as a channel busy fraction [215], that is, the time the channel was busy over a fixed time period. This can be measured by frequently probing the channel (e.g., for 10 µs) and observing the average power level; the channel busy fraction is then simply the percentage of probes where the observed power level was higher than the default Clear Channel Assessment (CCA) threshold.

The core of the DCC mechanism is a state machine that switches states and adjusts MAC and PHY parameters based on the minimum (clMin) or maximum (clMax) observed channel load in a given time interval. Figure 2.7 shows the state machine for the CCH. It consists of three states, namely *Relaxed*, *Active*, and *Restrictive*. For SCHs this state machine can be extended by dividing the active state into several sub-states to be able to control channel access in a more fine-grained way. DCC supports five different ways of controlling channel access:

	Relaxed	Active				Restrictive
		AC_VI	AC_VO	AC_BE	AC_BK	
TPC: Power [dBm]	33	keep	25	20	15	-10
TRC: Interval [s]	0.04	keep	keep	keep	keep	1
TDC: Rate [Mbit/s]	3	keep	keep	keep	keep	12
DSC: Sens. [dBm]	-95	keep	keep	keep	keep	-65

Table 2.4 – ETSI ITS-G5 settings for the DCC state machine for the CCH. 'keep' means the value is not changed when the state is entered. Recommend values differ for the SCHs.

- Transmit Power Control (TPC): Controlling the transmit power affects the transmission and interference range of a node. This is done so the wireless channel can handle more nodes in the same area.

- Transmit Rate Control (TRC): This affects the frequency with which a vehicle may access the channel. It can also be understood as the packet interval and has a high impact on the channel usage as it directly influences when a node can transmit a packet.

- Transmit Data Rate Control (TDC): Based on the channel load the data rate of a node is altered. A lower data rate offers more robustness and reliability, but lower throughput.

- DCC Sensitivity Control (DSC): This changes the CCA threshold without changing the sensitivity of the receiver. This directly affects whether a node sees the channel as busy or idle.

- Transmit Access Control (TAC): Restricting the number of packets a node can send supports "the operational requirement of fair channel access" [79].

Table 2.4 gives an overview of the MAC and PHY settings affected by DCC. When in state *Active*, parameters can be changed individually for each EDCAF, while they are the same for each AC in the *Relaxed* and *Restrictive* states. Most parameters are not changed when DCC enters the *Active* state, marked by a `keep` in the table. This introduces a level of hysteresis and is done to avoid changing the parameters in an over-sensitive manner. The listed settings are default values for the CCH, however, they differ for the SCHs where each SCH has its own state machine and settings.

This mechanism has a significant influence on the channel access of each node in an ETSI ITS-G5 network and therefore needs to be modeled to be able to investigate PETs based on ETSI ITS-G5. By doing this, we were able to uncover shortcomings and side effects of DCC, which we will show in Section 3.3.

Message Types

Cooperative Awareness Message (CAM) [84]: These messages are the counterpart of BSMs in IEEE WAVE. They are also sent in a periodic fashion and include important parameters describing the current state of the sending vehicle. While they are also sent with a maximum frequency of 10 Hz, the facility layer makes use of additional rules to determine whether a CAM should be generated. These include the difference in heading (4°), position (5 m), or speed (1 m/s) since the last transmitted CAM. However, a minimum frequency of 1 Hz must be ensured.

While at the time of writing the standard did not explicitly mention in which AC CAMs should be sent, it is generally assumed that CAMs will be sent in the Access Category AC_VO. [61, 132, 133].

CAMs can include detailed information about the sending vehicle, most importantly, they can include a path history of the vehicle to allow for easier tracking and thereby more reliable collision avoidance for receiving vehicles. Apart from that, they include the current state of exterior lights such as brake lights or turn indicators and other, partly optional, information such as occupancy and vehicle dimensions. The potential impact on privacy of these fields will be discussed in the next section (Section 2.2).

Decentralized Environmental Notification Message (DENM) [77]: This message type is used to inform other vehicles about immediate dangers, such as emergency braking or hazardous road conditions. They are triggered by certain events and are targeted at a defined geographic area, e.g., a location reference point, circle, square, or ellipsis. To control the directed dissemination of these messages, DENMs are forwarded using geographic routing. Once they reach their area of relevance they are 'kept alive' by rebroadcasting them until their expiry time, which is included in the message. Depending on the priority of the event, a different Access Category is used. Like CAMs, they also include potentially identifying information and may be harmful to the location privacy of drivers.

Signal Phase and Timing Message (SPAT) & **Road Topology Message (TOPO)** (both [83]) are two further message types in ETSI ITS-G5. The first is used to transmit information about traffic lights to enable GLOSA applications, the second is used to exchange information about the road topology to, among other reasons, "overcome the issue of different map formats being used." [83]

2.2 Privacy in Intelligent Transportation Systems

Recent revelations on the dimension and reach of governmental programs to globally (and often illegally) collect, store, and analyze private data exceeded even pessimistic estimates. Many of these privacy violations are conducted by eavesdropping on national and global communication networks or by obligating companies to collaborate with intelligence services and grant access to potentially sensitive user information. The comprehensiveness of these interferences suggests "that if there's data, there is also abuse" [63]. Trends (such as the increasing adoption of cloud services) indicate that future technology will rely even more on high degrees of interconnectivity, further increasing the amount of personal data potentially exposed to both private and public institutions. Vehicular networks are no exception to this, and, due to their large volume of potential sensitive information, deserve special attention when it comes to privacy protection.

Field Operational Tests (FOTs) all over the world, along with the advancing IEEE and ETSI standardization processes, show that vehicular networks have long ceased to be only a vision and will become a reality within the next years. Effective privacy protection, however, is still not an integral part of these standards and is also often neglected in FOTs. In this section, which is partly based on our *IEEE Security & Privacy* article "Driving for Big Data? Privacy Concerns in Vehicular Networking" [63] and articles published at conferences [21, 246], we give a definition of what privacy means, discuss and outline research directions that approach these problems, review important related work, and identify open challenges, which (if solved) we believe will substantially help protect drivers' privacy. With current global privacy violations in mind, we analyze the current versions of IEEE and ETSI families of standards, and pessimistically discuss possible privacy issues in vehicular networks.

2.2.1 What is Privacy?

To protect drivers' privacy, it first has to be understood what the term *privacy* actually means. In his 1967 book *Privacy and Freedom*, Alan Westin defined privacy as "the ability of an individual to control the terms under which personal information is acquired and used." [252]. Twenty years later, the EU privacy directive [74] defines "personal data" as "any information relating to an identified or identifiable natural person [...]; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity."

Reducing the vagueness of these definitions, Nissenbaum defines privacy in terms of contextual integrity, which "ties adequate protection for privacy to norms of specific contexts, demanding that information gathering and dissemination be appropriate

to that context" [173]. This means that norms determine which information is appropriate to reveal in a given context and also how that information may be distributed to other parties. Information is therefore always associated with a context that governs how it may be used and a privacy violation can then be seen as the re-purposing of personal data collected for another specific purpose.

One method that helps grasp the concept of privacy is dividing it into several sub-domains. For example, Finn et al. [92] present a systematization that divides privacy into seven domains; namely, these are privacy of 1) the person, 2) behavior and action, 3) communication, 4) thoughts and feelings, 5) data and image, 6) location and space, and 7) association. The lines between these domains are blurred, however. For example, detailed location information about an individual can also allow to draw conclusions about a person's behavior and actions and give insights on the people with which a person is associated.

Additionally there is general agreement that, regardless of the domain, privacy can be divided into five different properties that can be enforced by technical means, that is, by deploying Privacy-Enhancing Technologies (PETs) [42, 181]: **Anonymity** describes the property that a target individual cannot be distinguished from a set of other subjects. In the context of vehicular networks, this usually describes a set of vehicles among which a target vehicle cannot be singled out. **Unlinkability** describes the inability to link two or more subjects, actions, or locations (e.g., two different transmissions from the same vehicle). **Undetectability** describes an adversary's inability to discern whether an item of interest, e.g., a certain message, exists or not. **Plausible deniability** (also: repudiation) describes the ability of a subject to deny having performed an action, e.g., having driven to a certain location. Lastly, **Confidentiality** refers to an adversary's inability to access the content of data.

Neither of the two classifications present truly non-overlapping privacy domains or properties. We therefore also introduce a taxonomy that classifies privacy according to the kind of data PETs are trying to protect. We can distinguish between **published data**, **observable information**, **re-purposed data** and **leaked data**. Published data describes data that was willingly made (possibly permanently) available. In the context of vehicular networks, this could be a published database of seemingly anonymized location information of different vehicles. Observable information requires some kind of presence of the adversary trying to violate user privacy. For example, an adversary setting up an access point could overhear unencrypted messages sent by vehicles to track them throughout the network. Re-purposed data includes all privacy violations by, e.g., system or service providers, who were willingly given confidential information by users but use it for a different purpose than initially stated (such as a location-based service that sells user location information to third parties). Leaked data describes data that was not intended to be accessible by an adversary but was leaked by means of illegal measures or was unintentionally

made public. This is especially critical as PETs are often not deployed to protect this data.

For a better understanding of the scope of this thesis, we classify our work using all three classification approaches (that is, privacy domain, privacy property, and data): When we talk about privacy in this thesis, we usually refer to the location privacy category, that is, the protection of a person's past, current, and future whereabouts, as this seems to be the most directly affected privacy domain in the context of private transport and vehicular networks. Naturally, the main properties investigated in this thesis are anonymity and unlinkability as these are the most obvious properties associated with location privacy. Lastly, we investigate PETs mainly in the context of observable information, that is, the messages sent by vehicles. Furthermore, we evaluate how published data by the system provider may affect location privacy of an individual.

Insufficient or even absent measures to protect drivers' location privacy could have serious implications for drivers participating in vehicular networks [49, 63, 119]. Aside from the disclosure and exploitation of personal and private information, this can include the possibility to install overly restrictive law enforcement systems and thereby affect the quality of life for all drivers. This is especially worrisome when people do not have a choice whether or not to participate, as these systems will likely become mandatory in the future, as is explicitly listed as a possible outcome of a current US DOT study [189].

2.2.2 Location Privacy and its Importance

Economically, there is no doubt that there is a large demand for personal data and many seem to accept and tolerate the industry's growing interest to collect personal information in order to generate profit. Many online services that are seemingly free of charge require the user to disclose personal information in order to be used. A potential user can then make a choice whether the benefits they receive from this service outweigh the value of information they give up. Therefore private data, or more generally spoken, privacy can be seen as a kind of currency [39, 149]. Privacy has a value attached to it and we believe that each person should be able to decide individually what that value is.

When it comes to location privacy, many people associate only a low value to their location information, as current studies indicate: A majority of the test subjects would sell one month of location data to be used commercially for as little as US$ 35 [39, 41]. Furthermore, concern about being tracked by a third party does not seem to be too strong as tracking is already done by mobile phone operators. This suggests that from a service provider's point of view, preservation of location privacy might not be a critical feature for the design of an ITS as it may not even

have a significant impact on the financial success of the system. In fact, there is even a business case for *not* preserving users' privacy as personal data can be used for targeted advertisement or sold to other companies. For example, mobile phone operators recently revealed plans to sell customer location information in many European countries [13].

Even if companies have strict data protection policies and do not willingly misuse personal information, the extent to which they can protect users' data could still be limited by governmental regulations, e.g., disallowing the deployment of privacy protection mechanisms or forcing them to cooperate. Of course, this governmental access to private data (often performed by non-governmental sub-contractors) is not limited to participating companies but can also be obtained by exploitation or compromise of the used communication networks. Collecting and analyzing this data is even easier when the provider or operator of the ITS is the government itself.

In many cases, e.g., when using a cellular phone or a location-based service, a person could simply choose not to use a service and thereby avoid revealing personal information and preserve their privacy. There is, however, a difference when it comes to vehicular networks as one of the benefits of these networks is traffic safety – something that most users will most likely value higher than location privacy. If location privacy is not protected in the context of an ITS, drivers are forced to give up their location information for the sake of personal safety. Even if a driver chose privacy over his safety (and the safety of others), upcoming plans of making IEEE WAVE transceivers legally mandated would leave them no choice.

A violation of location privacy can have severe implications; from obtrusive advertisements over disclosure of information that causes embarrassment or humiliation [203] to oppressive regimes that (in a worst-case scenario) could use it to persecute political or social minorities by using location information to link individuals to each other.

As stated above, the violation of one privacy domain can also affect other domains. This is particularly the case when location privacy of an individual is compromised: For example, knowing a person drove to the hospital could indicate a medical condition and make the person appear as a less desirable candidate for potential employers. In order to avoid this, a system has to provide anonymity, the precondition for location privacy. Anonymity is defined by Pfitzmann and Hansen as the "state of being not identifiable within a set of subjects [...]" [181]. Only when an individual cannot be identified or recognized can they preserve their location privacy.

There have been numerous publications on methods and algorithms to preserve different aspects of location privacy in the context of vehicular networks. As the standardization progresses, it will be essential to know which approaches will be realized and to what extent location privacy can be protected in ETSI ITS-G5 and IEEE WAVE. In the following we examine the current progress and its implications

on privacy for drivers. We also discuss how, in a worst-case scenario, interested parties can exploit such a system.

2.2.3 State of the Art

The need for privacy protection in vehicular networks has certainly been understood early on. Position papers have been published as far back as in 2004 [119] and 2005 [49], however only little has translated into the standards of ETSI and IEEE. In general we find that, although far from comprehensive, the ETSI family of standards covers privacy aspects in more detail than IEEE does in their IEEE 1609.2 standard [122]. While ETSI 102731-v1.1.1 notes that "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence" [75], specific recommendations for effective protection measures are either missing or insufficiently precise.

Taking a closer look into the ETSI and IEEE standards, we evaluate presented privacy measures and the associated implications: In both systems, all vehicles will periodically transmit unencrypted broadcast messages – CAMs or BSMs – including information on the vehicle's current state (ETSI 302637-2-v1.3.0 [84] and SAE J2945.1-2.2 [196], respectively). This information contains (but is not limited to) the current direction of a vehicle, its position, speed, and acceleration. The frequency of these messages is either a fixed 10 Hz or varies from 1 Hz to 10 Hz depending on the current traffic situation [84]. These messages can be (undetectably) overheard by anyone close enough to the sender (in a vicinity of about 500 m to 1000 m [214]) using freely available and fairly inexpensive hardware [21] and therefore put drivers' location privacy at risk.

Building a Vehicular Public Key Infrastructure

To prevent unauthorized users from sending messages and joining the network, the standards IEEE 1609.2-2013 [122] and ETSI 102941-v1.1.1 [81] describe the deployment of a PKI as shown in Figure 2.8. Vehicles have one pre-installed base identity which must never be used to sign Car-to-X messages, but is only used to generate or request pseudonyms from a (possibly governmental) Certificate Authority (CA). These pseudonyms are also certificates and only valid when they are (directly or through a chain) signed by the CA. Each vehicle maintains a pool of pseudonyms and uses them as its visible address, that is, to sign and send messages over the wireless channel. A receiving vehicle will only accept a message if it has been signed with a valid pseudonym. While it would be beneficial for the anonymity of a driver to use a different pseudonym for each message, it would very likely compromise safety applications of other vehicles, as this can lead to problems linking two messages

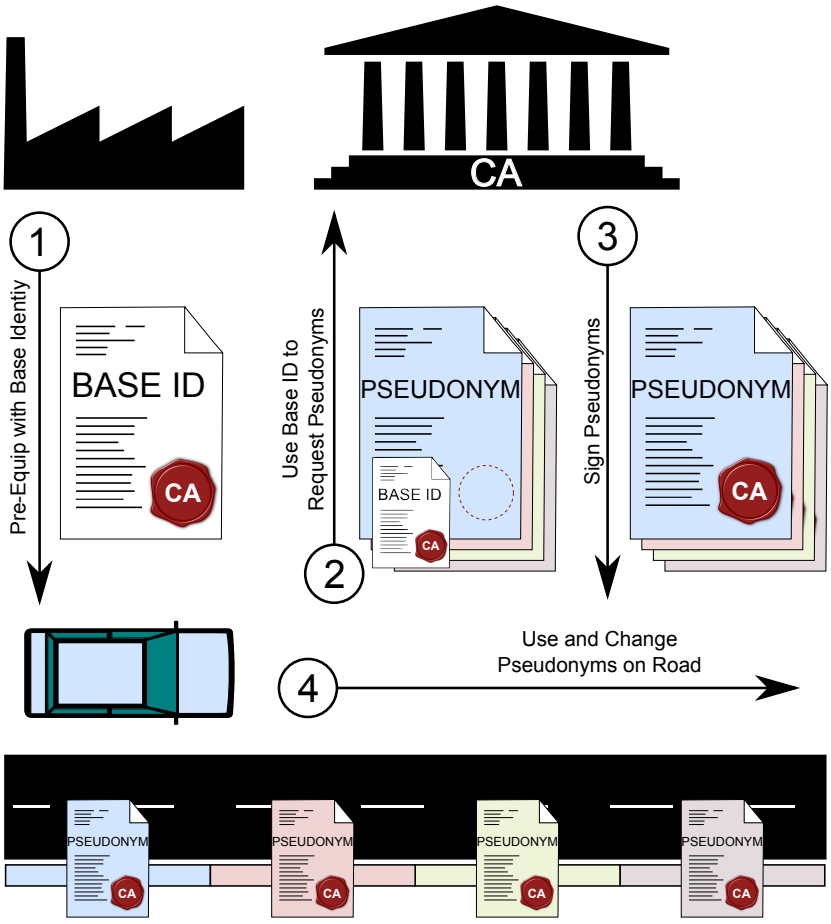


Figure 2.8 – Simplified view of the Public Key Infrastructure (PKI) in IEEE WAVE or ETSI ITS-G5.

to the same vehicle. Therefore, pseudonyms are only changed according to certain pseudonym changing strategies.

A formal explanation of the PKI in vehicular networks follows (for signs and symbols refer to Table 2.5):

Let C_{CA} be the certificate of the CA and CA^- and CA^+ the corresponding private and public keys. Each vehicle i is preloaded with a base identity certificate C_{B_i}, B_i^+ and B_i^- along with C_{CA} and CA^+. C_{B_i} is signed by the CA and allows the vehicle to request the signing of n pseudonymous certificates $P_{i,j}$ by the C_{CA}. For each pseudonymous certificate, a vehicle also maintains the corresponding keys $P_{i,j}^-$ and $P_{i,j}^+$. All certificates in a vehicular PKI contain the public key, the expiration time, a scope, and a signature along with the id of the signer.

When a vehicle sends an unencrypted message m (e.g., a periodic safety message) it will send the message, a signature of the message using the private key of the

Notation	Description
C_K	Certificate of entity or identity K
K^-, K^+	Private and public key for C_K
$h(x)$	Cryptographic hash function
B_i	Base identity of vehicle i
$P_{i,j}$	Pseudonym j of vehicle i
$E(K^-, m) = c$	Encrypting m with private key K^- to ciphertext c
$D(K^+, c) = m$	Decrypting ciphertext c with public key K^- to m
$s(K^-, m) = E(K^-, h(m))$	Signing message m using private key K^-

Table 2.5 – List of relevant notation and operations in a vehicular PKI.

currently used pseudonym, and the pseudonym certificate. The message will then be $m \wedge s(P^-_{i,j}, m) \wedge C_{P_{i,j}}$. A receiving vehicle has to check whether $C_{P_{i,j}}$ is a valid pseudonym certificate by checking the signature in the certificate. This can be done by comparing $D(CA^+, s(CA^-, C_{P_{i,j}})) = h(C_{P_{i,j}})$ and verifying if all other necessary fields are valid (e.g., the expiration time). Next, it needs to be checked if the signature for the message is correct. Because $P^+_{i,j}$ is included in $C_{P_{i,j}}$, this can be achieved by checking whether $D(P^+_{i,j}, s(P^-_{i,j}, m)) = h(m)$. This principle ensures message integrity and (pseudonymous) authenticity. It should be noted that in the case of intermediate CAs the sending vehicle must attach the certificate chain so that the receiving vehicle is able to verify the message with its pre-installed CA certificate.

Confidentiality can be achieved by, instead of sending m in plain text, encrypting it using the public key of the receiver vehicle u so that the message becomes: $E(P^+_{u,j}, m) \wedge s(P^-_{i,j}, m) \wedge C_{P_{i,j}}$.

Should a vehicle be sold, broken, or its On-Board Unit (OBU) be compromised, it is necessary to invalidate its pseudonym pool to prevent the transmission of faulty or malign messages [107]. This process is called certificate revocation and, while currently not foreseen in the ETSI family of standards, can be achieved by distributing a list of revoked certificates, a so-called Certificate Revocation List (CRL). In addition to checking signatures, a receiving vehicle must then ensure that a message was not sent using a revoked certificate.

Vehicle Tracking

A common approach to complicate linking messages with different pseudonyms to each other, and hence to prevent the tracking of vehicles, is to change the pseudonym, i.e., the source address and certificate for all sent messages, from time to time.

Determining how and when to actually change the pseudonym is a difficult challenge. There exist various proposals and ideas to change pseudonyms in vehicular

networks, as described in a recent survey [180]. However, there is no common agreement on which strategy is the most effective and which interferes the least with safety applications. A straightforward method would be to change the pseudonyms in a fixed or random time interval [66, 174]. The problem with these approaches is that a vehicle might be in a suboptimal position to change the pseudonym, e.g., a safety critical situation or not surrounded by enough vehicles to actually confuse an eavesdropping adversary. Context-based pseudonym switching is believed to be a promising approach to fix this problem: Pseudonyms are only changed when it is believed to cause confusion for an eavesdropping attacker, e.g., when vehicles with similar speed and direction are close-by [67, 100]. When these vehicles change their pseudonyms simultaneously, an attacker might not be able to link the new pseudonyms to the old ones. Another efficient countermeasure is the concept of random silent times, that is, not transmitting messages for a random amount of time after the pseudonym was changed [118]. However, this would render the vehicle invisible to the OBUs of surrounding cars and hence interfere with safety applications [148]. Other approaches include mix-zones, i.e., geographical areas where vehicles change their pseudonyms, possibly by the help of cryptographic schemes to confuse attackers [28, 98]. It has to be noted that in a simulation environment, it was shown that pseudonym changes (even with low transmit frequencies of to 1 Hz to 2 Hz) can be tracked without correlation of additional data [21, 60, 255] if a theoretical attacker was able to overhear all messages and the position information in the messages is accurate.

Even though many pseudonym strategies are to be found in the literature, the IEEE and ETSI family of standards do not recommend a specific one, nor do they list possible strategies to choose from. The documents only mention the need to "use a pseudonym that cannot be linked to [...] the user's true identity" (ETSI 102893-v1.1.1 [76]) and suggest to change it frequently "[...] to avoid simple correlation between the pseudonym and the vehicle" (ETSI 102940-v1.1.1 [80]).

Tracking becomes more difficult for an attacker when they are unable to overhear all messages but, for example, only monitors certain areas of a city. When a vehicle leaves a monitored area and changes its pseudonym before it enters another one there is a good chance to avoid re-identification by an attacker [28]. However, data included in periodic safety messages (such as vehicle width and height) could be used to correlate messages and therefore increase the chance of re-identifying a vehicle. While this is acknowledged in the IEEE and ETSI family of standards, the message formats for BSMs and CAMs do not reflect this concern. Furthermore, other messages such as DENMs, i.e., messages to inform other vehicles of hazards such as accidents or black ice, include fields that allow the identification of a vehicle and "[...] may be problematic with respect to privacy protection" (ETSI 102893-v1.1.1 [76]). The suggested use of sequence numbers in the latter standard is especially critical as it

allows for the association of sequences with specific sources. In general it can be said that the more information a vehicle discloses, the easier it becomes to link two pseudonyms and therefore to track it. It is an open challenge to identify how often and which additional data can be included in messages to avoid this problem and how accurate this data has to be to still allow proper operation of safety applications without making vehicles more or less unique. Even if some message fields are marked *optional* in the standards, the decision whether to include them will not be made by the driver but by the OBU.

The actual privacy protection measures related to location privacy found in the ETSI standards are very imprecise, or even almost absent in the case of IEEE WAVE. In an ETSI threat, vulnerability, risk, and assessment analysis, ETSI 102893-v1.1.1 [76] states that tracking can be prevented by either the use of pseudonyms or by sending encrypted messages. We deem these statements a dangerous simplification, as they suggest that the sole use of a pseudonym provides sufficient location privacy to users. This does not hold when pseudonym changes can be tracked, or (under the optimistic assumption that they can not) when the path traveled by a car using a single pseudonym can be related to a home or work addresses revealing the true identity of a driver [104, 143]. Furthermore, the use of encryption does not prevent location tracking as stated in the standard. All participating communication partners (e.g., provider operated RSUs) are still able to decrypt these messages and can therefore track the sending vehicle. Apart from that, a large portion of privacy-critical messages are periodically sent unencrypted, annulling any benefit from simultaneously sent encrypted messages.

Aside from online tracking, there is another threat to the problem of disclosing the location of vehicles: When a vehicle willingly or unwillingly sends faulty messages, all of its pseudonyms need to be invalidated. These pseudonyms are put on a CRL and distributed in the vehicular network so that messages signed with revoked pseudonyms can then be ignored by other vehicles. However, publishing a list of all or many pseudonyms belonging to one vehicle can retrospectively reveal location information of the driver. A possible solution for this is the deployment of privacy-preserving revocation schemes [56, 107] that only disclose current and future pseudonyms of a vehicle, such as the solution presented in Section 4.2. Although the responsible IEEE 1609.2-2013 standard acknowledges the need for privacy protection, this issue is not addressed.

Another privacy-related issue is the logging and storing of the large amounts of data collected by on-board units, road side units, traffic information centers, etc. Policies on which data is stored (and for how long) are subject to data protection laws and agreements in the countries of operation and cannot be sufficiently covered by the standards. While we acknowledge the necessity for strict regulations in this matter, specific recommendations are outside the scope of this work.

Fingerprinting tracking

Even under the assumption that pseudonyms prevent an adversary from tracking vehicles through the network, it has been shown that there are other attack vectors that do not rely on information in the actual messages but exploit characteristics of the transceiver chip. Several attack vectors to track users' mobility have been identified and countermeasures have already been discussed [40]. While these attacks mostly target traditional IEEE 802.11 networks, many of them can be directly applied to IEEE 802.11p networks, however, their effectiveness in highly dynamic environments such as vehicular networks has yet to be evaluated.

The physical waveform transmitted by an IEEE 802.11p radio can potentially be used to re-identify a user, e.g., when there are characteristic distortions in the signal. These distortions can be caused by small imperfections and variations in the analog part of the radio [27, 69] or the wireless channel itself [177]. This can be exploited by, e.g., using a sophisticated signal analyzer in a shielded test chamber as demonstrated by Klein et. al [130]. It has not been shown whether this method can be applied to re-identify vehicles in a highly mobile vehicular network characterized by a rather unstable channel with strong fast fading effects.

Another possible attack vector is the analysis of the signal in the transient phase, that is, the short time immediately after the transceiver chip switched to transmit mode. Powering up transmit components such as amplifiers causes the signal to have a characteristic shape that was shown to be exploitable to allow the identification of devices with an accuracy of up to 98 % [243]. Even though this approach is very reliable in static scenarios, its practical exploitation in vehicular environments has not yet been proven.

Higher layer fingerprinting includes the analysis of timestamps in TCP packets [136] or the exploitation of vendor-specific features like protocol and traffic characteristics [97]. While the first attack is possibly limited by GPS-enabled time synchronization in vehicular networks, the latter could be used to identify the model or the vendor of the hardware a vehicle is using.

In Section 4.3 we present the scrambler attack, a more robust fingerprinting technique that exploits a weakness of IEEE 802.11 transceiver chips. It can be used to identify unique users in vehicular networks and therefore needs to be addressed before the roll-out of these chips.

2.2.4 Automated Traffic Surveillance

The use of changing pseudonyms is the most important privacy feature of both the European and North American systems, but even if pseudonyms cannot be linked to each other, the problem remains that each pseudonym can still be resolved to a base

identity by the certificate authority that signed it, meaning that complete location privacy cannot be ensured.

There are different approaches to circumvent this problem: For example, Schaub et al. proposed the use of blind signatures to obtain so-called *vtokens* from the CA. The CA requests that the requesting vehicle reveals a large percentage of the vtokens, but not all. The remaining can then be used to request pseudonyms that cannot be linked to a base identity. Another approach is pseudonym swapping [67] (Section 4.1): Vehicles exchange their pseudonyms and thereby eliminate the base identity-to-pseudonym mapping at the CA.

It is unlikely that any PETs that provide privacy from the authorities are going to be installed in a future ITS as they would not be "[...] supporting law enforcement access under appropriate circumstances" (IEEE 1609.2-2013 [122]). We therefore believe that it is of utmost importance that pseudonym resolving is controlled lawfully, for example, by the means of knowledge splitting, making it impossible to resolve a pseudonym without the collaboration of multiple institutions. Regulations must include clear statements on when pseudonyms are allowed to be resolved to base identities. This kind of knowledge separation is also recommended in ETSI 102941-v1.1.1 [81]: The standard states that the authority handling the signing of pseudonyms should not have knowledge of base identifiers and that these are only known to a separate authority. Of course, these systems would not offer additional privacy protection if both authorities are run by the same institution and access to pseudonym/base identity pairs cannot be controlled externally. Also, it needs to be ensured that no third party is able to (secretly) access all data arbitrarily.

The possibility to resolve pseudonyms to base identities ensures accountability and allows the identification of vehicles that (deliberately or unintentionally) send false messages, the recovery of stolen vehicles, and the detection of hit-and-run offenses. On the downside, it could also change traffic supervision as we know it.

A vehicle that periodically sends out safety messages (such as BSMs or CAMs) including its current position and speed will also transmit them when the driver is speeding. These unencrypted messages can potentially be received by provider-operated RSUs in the vicinity. The fact that all messages are signed using a pseudonym and a pseudonym can be resolved to a base identity by an operator leads to non-repudiation, that is, the sending vehicle cannot deny having sent the message. The receiving RSU can therefore act as a WLAN-based speed camera and tickets could be issued solely on the basis of receiving a periodic safety message.

The formats of safety messages in both ETSI 302637-2-v1.3.0 [84] and IEEE WAVE (SAE J2945.1-2.2 [196]) do not only allow detection of speed limit violations but virtually all traffic offenses. Transmitted information includes the state of all exterior lights (e.g., indicator or brake light), the path history in case of CAMs and even the steering wheel angle (see Section 2.1). This allows recipients

to automatically detect when a vehicle is turning or changing lanes without the indication through a turn signal; it further makes it possible to detect whether a vehicle is running a red light, violating the right of way, or crossing a stop line without having stopped. While this can certainly be beneficial for traffic safety, e.g., by warning drivers of errant vehicles, it can also be used for automated ticketing.

Tendencies in both academia and industry show that future Car-to-X-enabled vehicles will be equipped with both an IEEE 802.11p and a cellular radio. The latter will be used for centralized services and can connect to, e.g., a traffic information server via IP. In scenarios where there is no RSU nearby, other vehicles could act as *witnesses* if they receive a message that indicates a traffic violation by reporting the incident to some provider-operated server using the cellular link. Even if vehicles are not equipped with a cellular radio, they could store the message in question and forward it to an RSU once they are within transmission range. Based on the certainty of the report (potentially derived from the number of witnesses) the errant vehicle could be fined.

Lastly, advanced driver assistance systems, including their numerous sensors such as fatigue warning systems or dashboard cameras, are already discussed to be used beyond their main purpose, for example, as a countermeasure against vehicle-related crime [134]. In a worst-case scenario, they could be exploited to support the (visual) identification of drivers and further contribute to an automated traffic surveillance system.

From today's point of view these scenarios seem far-fetched. Nevertheless, we want to point out that ITS's, in the way they are currently envisioned and standardized, give the operator (or the government) the ability to deploy these or similar privacy-compromising methods in the future. Certainly, these 'features' will most likely not be part of ITS's from the very beginning, but once OBUs are widely deployed or even legally mandated, this potential source of income for privacy violation will become far more interesting for the respective profiteers.

2.2.5 Open Challenges

There are many steps required to build a privacy-preserving ITS, some of which we will address in this thesis.

One of the most important tasks is to be able to quantify the level of privacy enjoyed by a driver in a vehicular network. Only with the possibility to put a number on different PETs are we able to compare them and recommend one over the other. The most common approach to evaluate PETs is by means of simulation. Using peer-reviewed, publicly available, established, and validated models for network communication and mobility increases the reproducibility of simulations. In this thesis we present such models, including a framework tailored for privacy evaluation

(see Chapter 3). Next, the metrics used to assess the performance of PETs need to be meaningful and easy to understand, especially when presented to a wider audience or decision makers. Unfortunately, current metrics [45] do not reflect this need and are difficult to interpret for people not working in the privacy domain.

Another open challenge is the evaluation of the impact of privacy measures on other areas such as traffic safety and comfort applications. In particular, the privacy/safety trade-off needs to be investigated more closely to comprehend the exact requirements of safety applications and to draw a reasonable line at the amount and accuracy of information included in periodic safety messages. This requires accurate mobility and driver models and an understanding of the requirements of deployed collision avoidance algorithms in terms of, e.g., latency or packet loss. While there have been various studies regarding the latter [1–3], the consideration of privacy protection mechanisms and their impact has only recently moved into the focus of researchers [148].

Tackling these challenges will not be enough without a stronger emphasis on privacy in ongoing standardization efforts, recommending practices for the technical protection of users' location information and measures to prevent institutions from easily accessing private data. This thesis tries to work towards a better understanding of the necessity of privacy protection in vehicular networks not only by showing methods to evaluate PETs but also by identifying problems and presenting solutions in the field of pseudonym resolution, privacy-preserving revocation systems, and physical layer fingerprinting. Along with the findings from many privacy researchers this knowledge needs to find its way into the design of upcoming systems. In particular, field operational tests all over the world should understand privacy as an integral part to serve as an example for future implementations, because retrofitting privacy is bound to fail.

2.3 Simulating Vehicular Networks

In this section, which is based on our book chapter titled “Simulative Performance Evaluation of Vehicular Networks” [64] in the book “Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment” published by Elsevier, we outline the principles and challenges of vehicular network simulation.

In general, performance assessment of vehicular network applications, protocols, and PETs can be approached using three different types of methodologies: analytical evaluation, Field Operational Tests (FOTs), and simulation. Their applicability depends on the type of application, as each of them have their own advantages and limitations, requiring researchers to carefully choose which method suits their needs best.

In the mathematical (or numerical) analysis of vehicular networks, system components are represented using analytical models, often based on probability distributions. These analytical models usually make use of simplifications to keep the complexity of the problem at a manageable level. For example, traffic is often modeled on a one-dimensional street using exponentially distributed gaps between vehicles. These simplifications can introduce inaccuracies leading to imprecise and misleading results, however, analytical studies can give valuable insights into the overall behavior, lower and upper bounds, and can generally help understand the designed system.

Testing the envisioned system in the field is probably the most straightforward approach and has many advantages. The obvious downsides of field testing are potentially excessive requirements in terms of cost, time, and other resources. Even in cases where these resources are available to an extent (e.g., in large field tests such as sim^{TD} [228]), the parameter space that can be explored is still limited. Drawing conclusions on the scalability of the envisioned system or detailed insights on the causes of observed behavior is often not possible. However, FOTs are invaluable for the validation of existing analytical or simulation models, and can also be used to develop new models, for example, based on empirical data collected in these field experiments [218]. In addition, real-life testing can help discover problems and system properties that have not been considered before and are therefore not accounted for in analytical or simulation models. Every system should therefore be thoroughly tested in FOTs before deployment, even though the current level of communication and collaboration between (typically industry-run) FOTs and the academic research community has room to be improved [57].

The third method for the assessment of vehicular network technology is simulation. In the last decade simulation has become the primary tool for the performance evaluation of vehicular network applications, technology, and protocols as it is a powerful tool to explore wide parameter spaces and investigate large-scale networks

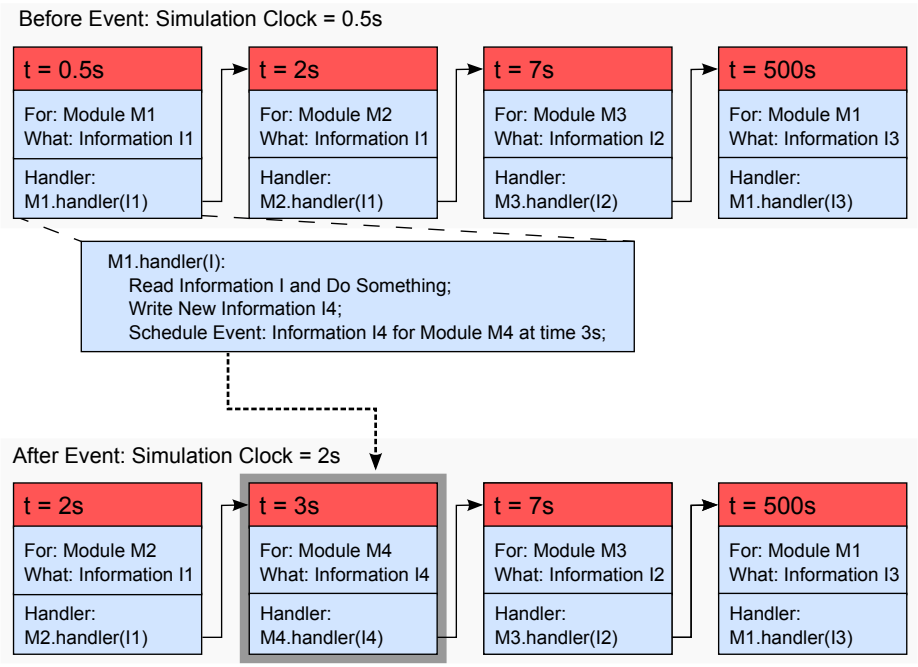


Figure 2.9 – Example event queue in a Discrete Event Simulator (DES). The event queue consists of four events before the first event in the queue is handled. In this event, a new event is generated and inserted in the event queue. The simulation clock advances to the next event once the handler has returned.

at low cost. However, just as for the analytical approach, its outcome fully depends on the detail and the correctness of the used models. Complex system components, such as multi-path radio propagation or road traffic, have to be simplified to keep the simulation run time at a reasonable level. It is a non-trivial task to determine with how much detail a given component or property has to be modeled: too abstract and it may produce unrealistic results, too complex and it becomes computationally infeasible or requires too much (possibly unavailable) data. Incorrect models may lead to false results and even invalidate the results of many simulation studies based on them in retrospect. Ideally, all used models should be peer-reviewed, validated, and open-source to minimize the chance of this happening. There exist various publicly available open-source simulation frameworks that make the setup and conduction of simulations easy and fast. Examples include Veins [219] (in which all models presented in this thesis are implemented), iTETRIS [192], and VSimRTI [205]. These frameworks make use of different DES's such as OMNeT++ [244], ns-2 and ns-3 [96], and JiST/SWANS [10]. In the following we will explain the principles of discrete event simulation and how OMNeT++, coupled with a traffic simulator, can be used to simulate vehicular networks.

2.3.1 Discrete Event Simulation in OMNeT++

In the context of vehicular networks, discrete event simulation of communication has become the most established simulation method. Contrary to continuous simulation (e.g., system dynamics simulation), where the investigated system is changed using continuous, differential equations, the state and attributes of entities in discrete event simulation are only changed at discrete points in time, so called events. Recently, hybrid simulation, that is the coupling of continuous and discrete event simulation has received much attention from the research community [48]. Vehicular network simulation is usually carried out using a special type of DES, that is, agent-based simulation. In agent-based simulation, each entity is an agent with attributes and can interact with other agents via methods. Agents are not usually omniscient and can therefore only rely on their own observations.

The core of any DES is an event queue, that is, an ordered list of all events currently known to the system. An example is given in Figure 2.9. Each event in the queue is assigned a time, an entity (or module) which the event is for, and information that is delivered to the entity when the event is processed. For example, an event could be the reception of a message from another node, with the information being the actual content of the sent packet. In the event handler, state variables are updated and future events are determined and inserted into the queue. Once an event has been processed, the simulation will dequeue the next event, advancing the simulation clock instantly to the time of that event. Therefore the simulation clock is not bound to the real clock, as is done in real-time simulation, but can advance faster or slower depending on events in the queue. The simulation ends once the event queue is empty or a predefined simulation time limit has been reached.

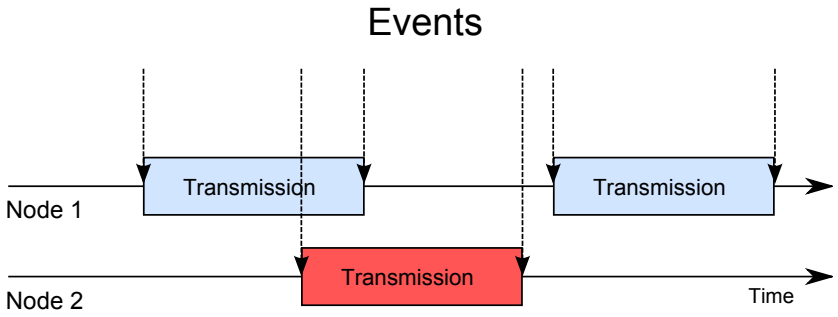


Figure 2.10 – Principle of mapping continuous processes on discrete events.

Modeling continuous processes in a DES is done by inserting events at the points in time when the state of the system actually changes. Figure 2.10 visualizes this methodology using an example from wireless network communication. Sending a packet is a continuous process for the duration of the transmission, however, the state of the system (‘there is no packet’ vs. ‘there is a packet’) changes only when the

node starts to transmit and when the transmission is over. In this example, events are generated and inserted into the queue when either node 1 or node 2 start or stop transmitting. In the context of wireless communication, deciding which node will then receive this event could be done defining a maximum range based on the transmission power, however, determining whether the packet can actually be successfully decoded is typically done in the event handler of the respective node.

Simulating the continuous mobility of nodes can be done by setting a fixed interval after which the position of a node is updated. To avoid nodes jumping from one location to another in the simulation, node positions can be interpolated (or extrapolated) between two location updates based on their speed, heading, and previous position.

Modeling in OMNeT++

OMNeT++ [244] is a widespread and well-established DES for network simulation. All models and network simulations in this thesis are implemented for OMNeT++, however they could be ported to other simulation environments such as ns-2 and ns-3 [96].

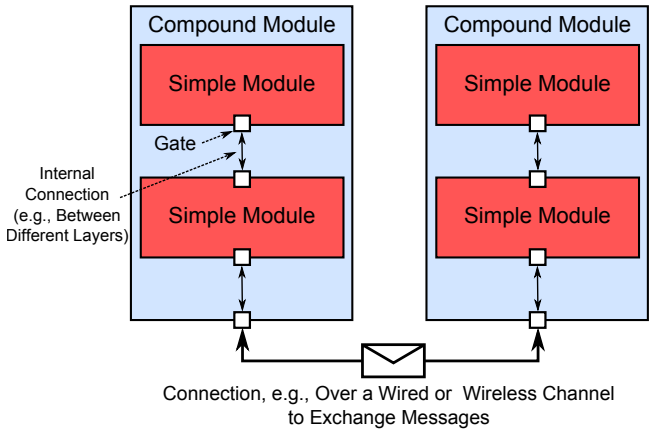


Figure 2.11 – Modeling in OMNeT++: Compound modules (e.g., network nodes) consist of multiple simple modules (e.g., different OSI layers). Gates are used to connect modules and allow the exchange of information between them.

The lowest level entity in OMNeT++ is a `simple module`. Each simple module can be assigned C++ code that is run when an event is delivered to this module. Simple Modules can be connected to other modules by the use of `gates` over which they can exchange `messages`. Receiving a message from a node then triggers an event at the receiver side and allows the node to handle this message and react accordingly. The modeling paradigm usually used in OMNeT++ is to model different OSI layers via simple modules and combine them in a `compound module`. An

example is shown in Figure 2.11. The compound module can then be considered the network node, e.g., a vehicle, and the different simple modules would then represent PHY, MAC, etc. When the application layer wants to send a message (or a packet), it will send it down to the network layer (or MAC if there is no network layer). The lower layers will add necessary information to the message by adding their headers, carrying information such as network layer addresses, routing information, and so on. Once the packet leaves the PHY, which is connected to another node (in the case of wired transmissions) or connected to a radio module, OMNeT++ will deliver this message to all receiving nodes, where the packet is then processed at each layer and handed up from PHY to the application layer.

This modeling paradigm is tailored for network simulation and therefore differs from other simulators, e.g., where behavior is modeled using the Unified Modeling Language (UML).

OMNeT++ itself only provides the DES core and does not come with models for common network protocols, such as TCP, UDP, or IEEE 802.11. However, there are various frameworks for OMNeT++, such as INET,[1] MiXiM [137], and INET-MANET[2] that contain models for different technologies. While INET used to focus on higher level layers to provide models for routing and other internet technologies, MiXiM focused more on realistic lower layer simulation, e.g., including a detailed signal interference model. Recently these three frameworks have been merged for maintenance and collaboration purposes.

When work on this thesis began, the community did not have accurate models for the upcoming technology standards IEEE WAVE or ETSI ITS-G5. Also, the channel and path loss models used were oversimplified and their accuracy and realism did not sufficiently reflect real environments [214]. An important contribution of this work is the development of models for both IEEE WAVE [65] and ETSI ITS-G5 [61], their validation, and based on them, the identification of shortcomings that may create problems in future deployments. We also extended channel models to support shadowing caused by buildings [218]. Almost all of our models were published in the free and open-source framework Veins [219].

2.3.2 Modeling Wireless Communication

In this section we will discuss how the wireless channel and the different OSI layers are modeled in order to increase the accuracy and meaningfulness of vehicular network simulation.

[1] `https://inet.omnetpp.org/`
[2] `https://github.com/aarizaq/inetmanet-2.0/`

Channel and PHY Modeling

Wireless channels play an important role in the performance of most envisioned ITS applications, but are error-prone, chaotic, and usually hard to predict [144]. As computational feasibility is an important requirement in vehicular network simulation, a channel model can only be an approximation of the real world. Determining the adequate level of abstraction is a challenging task: too abstract and the simulation results could become incorrect or misleading, too detailed and the simulation can become too complex and poorly performing or require too much (and possibly non-available) information about the scenario (e.g., building locations, different kinds of materials, etc.).

One of the simplest methods to model wireless channels is the use of a unit disk model (see Equation 2.1) where the packet success probability p_{succ} is a Boolean function of the distance d between sender and receiver: If the receiver is within a predefined maximum transmission range R of the sender the packet can be decoded, otherwise it will be lost.

$$p_{\text{succ}} = \begin{cases} 1 & \text{if } d \leq R, \\ 0 & \text{if } d > R. \end{cases} \tag{2.1}$$

When the performance of the examined application or protocol highly depends on the reception of single packets this abstract model can produce inaccurate results; however, it can still be appropriate for macroscopic simulations.

For microscopic simulations the state of the art is to assign the packet success probability p_{succ} based on the received power P_r which depends on transmit power P_t, the antenna gains of both the sender and receiver antenna (G_t and G_r, respectively), and the sum of all attenuation components L affecting the signal. Attenuation components can reflect signal power loss caused by path loss, slow fading, fast fading, or probabilistic attenuation effects.

$$P_r[\text{dBm}] = P_t[\text{dBm}] + G_t[\text{dB}] + G_r[\text{dB}] - \sum L[\text{dB}] \tag{2.2}$$

In this context, power levels P are usually given in dBm (decibel milliwatt), that is, the power ratio referenced to one milliwatt. Attenuation levels are given in dB to describe their effect on the signal as the ratio of input to output intensity. An attenuation of < 0 dB would therefore amplify the signal.

A widely used path loss model to capture the effect of decreasing signal strength over distance is the free-space path loss model (or, more precisely, an empirical adaptation thereof) which only depends on the distance d, the wave length in meters λ, and a path loss exponent α (usually set to 2, but can be changed according to the environment [186]).

$$L_{\text{emp-freespace}}[\text{dB}] = 10\log_{10}\left(\frac{16\pi^2 d^\alpha}{\lambda^\alpha}\right) \tag{2.3}$$

This model has been shown to often overestimate and underestimate the measured power level in the context of vehicular communication [218, 220]. The reason for that is a second strong component, namely the reflection of the signal from the road surface. This effect is called two-ray interference [200] and leads to constructive and destructive self-interference effects. The receive power $L_{\text{tri}}[\text{dB}]$ can be computed using the phase difference φ (which depends on the wavelength, the sender/receiver distance, and antenna heights) and the reflection coefficient $\Gamma_\perp$ (which depends on the incidence angle and on the reflection characteristics of the surface):

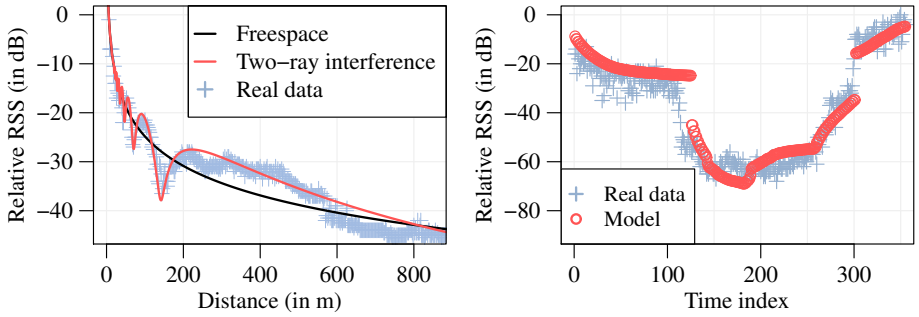
$$L_{\text{two-ray-int}}[\text{dB}] = 10\log_{10}\left(4\pi\frac{d}{\lambda}\left|1+\Gamma_\perp e^{i\varphi}\right|^{-1}\right)^2 \tag{2.4}$$

An in-depth description of the two-ray interference model is given in [220]. This model is not to be confused with the *simplified* two-ray ground model which is often used in vehicular network research [214]. The simplified two-ray ground model assumes that (under a distance threshold) radio path loss can be computed using Equation 2.3, and that for distances higher than a cross-over distance a simplified two-ray ground model assuming perfect polarization and reflection can be used. It has been shown that this model, which was developed for cellular communication, is not suitable for the simulation of vehicular networks, because when used with typical vehicular parameters, the cross-over distance would be at $\approx 885\,\text{m}$, resulting in a model that would generally be the free-space path loss model for almost all transmissions [214].

Figure 2.12a illustrates the difference between the free-space model and the two-ray interference model. It also compares them against real life measurements we made with IEEE 802.11p prototype hardware [214, 218].

Radio propagation is also influenced by obstacles such as buildings, pedestrians, trees, or other vehicles that attenuate the signal. This effect is called radio shadowing and occurs when the line of sight between sender and receiver is blocked, resulting in lower received power. If transmission attempts can be assumed independent and uncorrelated in time (on the order of seconds) and space (on the order of tens of meters), the effect of obstacle shadowing can be modeled purely stochastically, for example using the log-normal shadowing model as shown in Equation 2.5 [34]. It uses a normally distributed random variable X with standard deviation σ to determine the attenuation of the signal.

$$L_{lognorm}[dB] = 10\log_{10}(X_\sigma) \tag{2.5}$$



(a) Comparison of the free-space model, the two-ray interference model, and measurements from a field test based on [214]

(b) Comparison of real data and the shadowing model in Equation 2.6 based on [218]

Figure 2.12 – Path loss (two-ray interference, Equation 2.4) and the radio shadowing model (Equation 2.6) used in this thesis.

When transmissions are made within a short period of time, or the geometric conditions do not change between transmissions, however, purely stochastic models cannot be used. In these cases, models have to be deployed that take into account the exact positions of obstacles such as buildings [101, 163, 218] and vehicles [23, 221] to determine the level of attenuation caused by shadowing.

To account for radio shadowing caused by buildings, we use the computationally inexpensive model we developed and presented in [218]: Figure 2.12b shows that the signal attenuation caused by buildings can be modeled based on the number n of intersections with the line of sight between receiver and sender. The first component accounts for attenuation caused by the number of walls ($2n$) the signal has to penetrate, the second component approximates the attenuation caused by the interior of the obstacles using the total length l of all intersections. The attenuation is computed using two scaling parameters β and γ: β calibrates the amount of attenuation caused by each wall and depends on the material (e.g., brick, concrete), γ is given in dB/m and serves as a rough approximation for the attenuation caused by the interior of a building.

$$L_{build}[dB] = \beta[\text{dB}] \cdot 2n + \gamma[\text{dB/m}] \cdot \sum_{i=1}^{n} l_i[\text{m}] \tag{2.6}$$

Small scale fast fading is usually modeled using probabilistic models that take the receive power determined by deterministic propagation models as input. Popular examples include Rayleigh Fading [110] which models fading based on two uncorrelated Gaussian random variables, Rician Fading [190] that takes the existence of a strong line-of-sight component into account, and Nakagami-m [171] which models

multipath fading based on m paths and has been shown to be a suitable choice for the simulation of vehicular networks [231].

After the receive power P_i of a packet i has been determined, the simulation has to decide whether the receiving vehicle can decode the packet. In wireless network simulation this decision is usually made based on the Signal-to-Interference-plus-Noise Ratio (SINR) of the packet, that is the power level P_i divided by the sum of the power levels of all other packets on the channel plus the background noise N:

$$\mathrm{SINR}(i) = \frac{P_i}{N + \sum_{i \neq j} P_j} \tag{2.7}$$

Based on SINR it is possible to derive the bit error probability, i.e., the probability of failing to decode one bit. In the case of Quadrature Phase-shift Keying (QPSK) modulation and under the assumption of an Additive White Gaussian Noise (AWGN) channel it can be computed using Equation 2.8. The bit error probability can also be derived from empirical data [99].

$$\mathrm{BER} = \frac{1}{2}\,\mathrm{erfc}(\sqrt{\mathrm{SINR}[\mathrm{dB}]}) \tag{2.8}$$

The probability whether a packet of length l bits can be decoded successfully is therefore:

$$p_{\mathrm{succ}} = (1 - \mathrm{BER})^l \tag{2.9}$$

Then, the decision of whether or not a packet can be decoded is made by drawing a random number $v \in [0, 1)$ and comparing it against p_{succ} to determine the final decision: The packet is handed to the MAC if $v < p_{\mathrm{succ}}$.

Medium Access Control Layer and Upper Layers

A straightforward approach would be to combine the simple unit disk model with an idealized PHY and MAC that disregard collisions to yield a packet success rate of 100 % and always assume the channel idle to transmit without latency. This oversimplification can still be a valid approach if the actual performance in terms of latency or throughput does not significantly affect the investigated application or protocol. An example would be the investigation of inter-rendezvous times in a large-scale Delay Tolerant Network (DTN) with tens of thousands of vehicles, where a detailed channel, PHY, and MAC model would possibly not significantly change the outcome and would lead to unreasonably long simulation run times

However, in vehicular networks the majority of applications (everything safety-related [2], and all applications with firm, hard, or soft deadlines [172, 183]) have specific requirements or dependencies in terms of latency, throughput, or packet

success rates, making the use of an idealized MAC an unsuitable choice. Not only do transmissions from other vehicles affect all these properties, but also can Quality of Service (QoS) mechanisms within the MAC impact when a message will actually be sent. The MAC model used in packet-level simulation should therefore almost always be a model of the MAC used in the examined network, that is, a model of either the MAC used in IEEE WAVE (Section 2.1.1), ETSI ITS-G5 (Section 2.1.2), or UMTS/LTE/LTE-A in the case of cellular communication.

It can be observed that when accurate models are not available the research community tends to parametrize models of similar or related standards (e.g., IEEE 802.11b instead of IEEE 802.11p) to approximate their behavior to the desired one [65, 125]. This can lead to inaccurate and misleading results as we will show in Section 3.2. In particular, when simulating and studying distinct properties such as multi-channel applications [131] and interference [30], an exact representation of the envisioned architecture is required. But even if established and validated models are available, they often come with a vast number of parameters, many of them severely influencing the actual performance of the examined system. These settings should therefore be chosen according to current trends and recommendations and always be included in simulation studies. For example, although IEEE 802.11p allows data rates from 3 Mbit to 27 Mbit, FOTs commonly follow a recommendation to primarily use a rate of 6 Mbit. Also, fixed transmission powers that lead to too small or too large communication ranges should be avoided.

Reproducibility can still not be ensured, as a recent study showed: An investigation on the comparability among IEEE 802.15.4 and IEEE 802.11 models used in different network simulators revealed that, although all these models claimed to follow the standard, the simulation outcomes were alarmingly different for each of them [87]. Ideally, all employed models should therefore be open-source, well documented, widely-used, and possibly peer-reviewed. In addition, authors should always give necessary details of used models, their parametrization, and the scenario that was evaluated, to improve reproducibility.

Although most decentralized vehicular ad-hoc network applications and protocols use a three-layer stack (PHY/MAC/APP), both IEEE and ETSI standards support IP-based communication over the IEEE 802.11p link. Furthermore, RSUs are envisioned to be connected to each other and to a central server over an additional network interface such as Ethernet or a cellular transceiver. Other vehicular network applications require connections to centralized services (e.g., traffic information systems or location-based services) that are only reachable via IP. For the simulation of many vehicular network applications it may not be necessary to fully simulate the network layer, i.e., it can be sufficient to abstract from it using delay characteristics of the communication link [222]. However, especially in the case of cellular links,

empirical values for latency and throughput can be hard to obtain and are also likely to not be constant.

When evaluating the performance of a specific vehicular network application it can be crucial to account for other applications also running on the same IEEE 802.11p-enabled vehicle. For example, the periodic safety messages in both IEEE and ETSI already generate a non-negligible network load that will affect all other applications using the same communication channel [61]. Also, internal contention with packets from other applications and cross-layer mechanisms such as congestion control in ETSI ITS-G5 can introduce further latencies. Lastly, privacy and security mechanisms (e.g., changing pseudonyms or overhead from attaching certificates) need to be considered when developing protocols and applications for IEEE 802.11p-based vehicular networks.

2.3.3 Road Traffic Simulation

In the beginning of VANET research, many believed that VANETs are just an application for Mobile Ad-Hoc Networks (MANETs), a field that has already been studied for years with many established protocols for management, routing, energy awareness, and so on. However, it turned out that many of the already developed MANET protocols do not work properly in the context of VANETs [213]. Not only were the basic conditions slightly different (seemingly infinite energy, different use-cases) but also do the distinct mobility patterns of vehicles require careful consideration when designing protocols, applications, and also PETs. Compared to general MANET movement, vehicular mobility is bound to a road network, typically in a mixture of high- and low-density areas. High relative speeds and, thus, fast changing network topologies and possible strong partitioning require tailored protocols to work properly. Even in the context of vehicular networks, different scenarios yield entirely different mobility patterns: Driving on a freeway will subdivide possible communication partners in two groups, that is, traffic in the same direction (long connection times) and oncoming traffic (very short connection times), while in a city, a vehicle might experience a sudden drop in neighbors when entering a side street from a busy main road. It is therefore crucial to simulate vehicular network technology using realistic mobility patterns.

One way to achieve this is the use of traces, that is, position information, e.g., collected by equipping vehicles with GPS receivers and a logging device. This trace file can then be played back to simulate road traffic. The advantage of this approach is that these mobility profiles are directly taken from the real world, granting a high level of realism if the error-prone GPS readings have been corrected. However, there are non-negligible drawbacks: Creating these traces is a cost-intensive task, as a considerable amount of vehicles have to be equipped in order to represent real

traffic. In general, the maximum simulated traffic density is bound by the number of equipped vehicles that generated the trace. Duplicating vehicles from the trace can circumvent this problem but will in return decrease the level of realism of the simulated mobility.

Instead of creating their own traces, researchers can make use of publicly available ones ([12, 117, 139, 242]) that can be used to simulate urban mobility. Many of them are not generated using private transport, but public vehicles such as taxis or buses. The problem with this is that the mobility of these vehicles is atypical and may not represent regular traffic. Lastly, many public traces have a resolution in the range from 1 s to even 1 min per entry, requiring the movement in-between to be interpolated, introducing further inaccuracies and unrealistic movement.

A second, more flexible solution is the use of a traffic simulator. Traffic can be simulated on different scales, namely microscopic, mesoscopic, or macroscopic, with plenty of free and commercial simulation tools for each. In microscopic traffic simulation each vehicle is simulated individually and will be influenced by other vehicles around it. At the mesoscopic scale the focus lies more on the movement of small groups of vehicles (e.g., platoons), while in macroscopic traffic simulation, entire traffic flows are evaluated to study their effect on the traffic system. Privacy simulation in this thesis is concerned with the evaluation of the privacy level enjoyed by individual drivers. Evaluating the effects of network traffic on privacy, we employ packet-level simulation studying single nodes on the network, therefore the only suitable scale for traffic simulation is microscopic [212]. Popular examples for microscopic traffic simulators include VISSIM [158] and Simulation of Urban Mobility (SUMO) [141], the latter being the traffic simulator used for all simulations in this thesis.

In microscopic simulators, vehicles are assigned routes through a predefined road network, with acceleration and deceleration of a vehicle being determined by a car-following model that, amongst other things, takes into account current speed and the distance and speed of the leading vehicle(s). These car-following models are combined with lane-change models to determine when a vehicle will change lanes in order to prepare for a turn or to overtake another vehicle.

In the context of PETs it is important to take a closer look at the microscopic mobility models controlling a vehicle's movement. When mobility generated by a simulator is taken as an input for a tracking algorithm to measure how well a PET performs (that is, complicates tracking), it is necessary that the underlying mobility is as realistic as possible. Assume a mobility model that does not allow vehicles to overtake or to change lanes – tracking would become trivial as the order of cars on the road would not change. The confidence in simulation results evaluating different PETs rises with an increasing level of realism. The car-following models used in

this thesis are the Krauß model [142] and the Intelligent Driver Model (IDM) [239], which we will now explain more detail.

The Krauß Model

The Krauß model is a space-continuous, time-discrete, accident-free, single-lane, stochastic, microscopic car-following model [142]. Like the Gipps Model [102] it is very efficient to compute due to its variable time-step length. The model equations are given in Equation 2.10.

$$\begin{aligned} v_{\text{safe}}(t) &= v_l(t) + \frac{g(t) - g_{\text{des}}(t)}{\tau_b + \tau} \\ v_{\text{des}}(t) &= \min\{v_{\text{max}}; v_{\text{safe}}(t); v(t) + a(v)\Delta t\} \\ v(t+\Delta t) &= max\{0; v_{\text{des}} - \eta\} \\ x(t+\Delta t) &= x(t) + v\Delta t \end{aligned} \tag{2.10}$$

The vehicle determines a safe velocity v_{safe} based on a desired gap g_{des} to the vehicle in front with regard to the drivers' reaction time τ, the current gap g, the velocity of the leading vehicle v_l, and the timescale $\tau_b = \overline{v}/b$. a and b depict a vehicle's acceleration and deceleration, respectively. This size of this gap can be determined in different ways, for example, it can be set to $g_{\text{des}} = \tau v_l$. A desired velocity v_{des} is then chosen as the minimum of the maximum velocity v_{max}, the safe velocity v_{safe}, and the maximum achievable velocity in the next time-step with regard to acceleration a. The actual speed that is assigned to the vehicle is not v_{des} but $v_{\text{des}} - \eta$, $\eta > 0$ being a random factor to account for driver imperfection. The next position of the vehicle $x(t+\Delta t)$ can then simply be computed using the determined velocity v.

The problem with this model is that it strongly relies on the randomized variable η, which when set to 0 leads to degeneration of the model [142]. Furthermore, Krauß states that this random factor "cannot be justified from real car-following behavior" [142] as it changes a vehicle's speed instantaneously, leading to very unsteady speed and acceleration curves and unrealistic behavior in traffic jams. Also, the impact of the correction value η is dependent on the time-step length of the simulator, causing more fluctuations in a vehicle's velocity when a smaller time-step is chosen.

Intelligent Driver Model

Treiber et al. developed the Intelligent Driver Model (IDM), a space and time-continuous, single-lane car-following model that overcomes shortcomings of the

Gipps and Krauß models (such as unrealistic mobility in congested scenarios). The fully deterministic model can be broken down into two equations:

$$\dot{v} = a\left[1 - \left(\frac{v}{v_0}\right)^{\delta} - \left(\frac{s^*}{s}\right)^2\right]$$
$$\text{with } s^* = s_0 + s_1\sqrt{\frac{v}{v_0}} + Tv + \frac{v\Delta v}{2\sqrt{ab}} \tag{2.11}$$

The vehicle's acceleration $\dot{v}$ is a function of the desired velocity v_0, the current velocity v, the desired gap s^*, and the current gap s to the vehicle in front. The acceleration component δ is usually set to 4 [239]. One of the main differences to the Krauß model is that, in the IDM, the desired gap s^* is computed using both the velocity v and the approach rate Δv, which is the velocity difference to the vehicle in front. IDM parameters are: safe time headway T, maximum acceleration a, desired (comfortable) deceleration b, and minimum distances in a jam s_0 and s_1. Treiber et al. recommend values for these parameters [239, 240], with s_1 commonly set to zero. IDM has been shown to be able to reproduce real traffic very well, making it a good choice to be employed in microscopic traffic simulations even though it is not as computationally inexpensive as the Krauß model. The model can also be used in a time-discrete manner to obtain speed and position updates of vehicles [113] and, depending on the time-step length, will still produce realistic acceleration and deceleration profiles [212].

Gipps, Krauß, and IDM all have in common that they are collision-free. At a microscopic scale, there is a requirement for future mobility models to be able to include atypical driving behavior, as this is an important requirement for the investigation of safety applications [126]. Without these critical situations (e.g., red light violations, too small safety gaps, speeding) the benefit of safety applications can only be approximated vaguely through other metrics.

After careful consideration we chose to primarily use SUMO's slight adaptation of the Krauß model, as the IDM implementation in SUMO was discovered to show undesired side-effects. When not stated otherwise, usage of the Krauß model can be assumed.

Lane-Change Models

Lane-change models are tightly connected with car-following models to capture decisions on whether and when a vehicle changes lanes. Determining whether a vehicle will change lanes is usually achieved by evaluating a set of rules or conditions, e.g., if the vehicle could move faster in the new lane and still maintain all necessary safety gaps. Example input parameters of these models are the maximum speed and

the smallest acceptable gap (in terms of time or space) between vehicles, and also range in complexity up to the level of politeness of a driver. A lane-change model often used in today's traffic simulation is MOBIL (short for: Minimizing Overall Braking Induced by Lane change) developed by Kesting et al. [241]. Paired with, e.g., IDM it can be used to simulate microscopic multi-lane traffic. The basic model can be described by the use of two simple rules:

$$\begin{aligned} \widetilde{a}_c - a_c + p\,(\widetilde{a}_n - a_n + \widetilde{a}_o - a_o) > \Delta a_{th} \\ \widetilde{a}_n \geq -b_{\text{safe}} \end{aligned} \tag{2.12}$$

The first rule ($\widetilde{a}$ and a being the acceleration after and before a possible lane change for either the current vehicle c, the new follower n, or the old follower o) requires the situation after a possible lane change to be better than the current one by a certain global threshold Δa_{th}. The politeness factor p determines how much a vehicle incorporates other vehicles; when set to 0, the decision will be made solely by evaluating $\widetilde{a}_c$ and a_c. The second equation ensures that lanes are only changed when the deceleration of the new follower n does not exceed a given safe limit b_{safe}. The model can be extended to also account for European traffic regulations where the rightmost lane is appointed the default lane and overtaking on the right is disallowed.

SUMO uses similar models to control the lane changing behavior of vehicles. Over the years, the used models have been adjusted frequently and default models have been occasionally replaced. A detailed explanation of the recent lane-change models used in SUMO can be found in [72].

Scenario Modeling

The ability to realistically simulate the microscopic behavior of single vehicles does not automatically lead to realistic traffic patterns and flows. The underlying scenario in which vehicles move plays a major role and has been shown to have a large influence, not only on mobility, but also on many network metrics such as channel load, neighbor count, and so on [53]. There exist a number of different scenarios used in vehicular network simulation; we try to list the most common ones and sort them from synthetic to realistic:

- *Random waypoint mobility*: One of the simplest approaches to generate (random) mobility without the need for car-following models: nodes choose a random location on the map and will move there. Upon arrival, a new location is randomly chosen and the node will move there and the process is repeated. This mobility model is still widely used in MANET simulation and has even

been used in vehicular network simulation to reflect urban movement [167], however, it was soon shown to not correctly reflect traffic characteristics and is likely to produce incorrect results [11, 260].

- *Manhattan grid:* Movement is bound to vertical and horizontal lines spanning a grid. Origin and destination are still generated randomly, but vehicles will only move along the grid. While this simplified scenario is completely synthetic, it can be used to simulate the grid-like layout in many American cities, when the parameters of the grid are chosen accordingly.

- *Road topology*: Instead of a grid, vehicles move along a simplified road network based on a real map, lacking information about traffic light positions, lanes, or speed limits. The level of realism is, as for the Manhattan grid scenario, rather low [216].

- *Detailed geodata*: Map data including the number of lanes, turn restrictions, speed limits, and traffic light positions are imported to generate more realistic traffic flows and road utilization. Further, adding obstacles such as houses that influence radio propagation will affect packet success rates and network topology [218]. An example for a good data source is the OpenStreetMap Project [109], a user-maintained map database that (depending on the area) provides detailed maps which can be used as a basis for vehicular network simulation.

- *Demand models*: Even if the underlying road network is fully based on real maps, random mobility of vehicles will not necessarily reflect real traffic flows [93]. Generating realistic traffic demand is therefore an important step towards realistic traffic simulation. Attributing different properties (e.g., residential or industrial) to areas of the map, and taking into account the time of day, can generate more realistic traffic flows than randomly assigning origin-destination pairs [112].

- *City-wide mobility*: There have been several attempts to model and validate city-wide microscopic traffic [35, 242]. In the case of LuST [35], a 24-hour scenario of Luxembourg City, large parts of the map have been manually tweaked to compensate for modeling errors introduced by automatic conversion from OpenStreetMap data [16]. Validation is done by using data from regional traffic agencies and comparing it to simulated traffic. Despite the scenarios being published, they have not been able to establish themselves as a default scenario to be used in vehicular network simulation as of yet.

Current efforts include the integration of public transport such as buses and trains as well as cyclists and pedestrians. Also, vehicles that are currently not moving

need to be modeled when investigating network applications making use of parked vehicles [56, 68, 156, 161] as we will show in Section 4.2.

In all of these scenarios, it is also important to account for border effects of the simulated road network. One way to circumvent this problem is to define a Region of Interest (ROI) and only investigate vehicles in this region but to still simulate traffic in a larger area around it. Otherwise, roads on the border of the simulated road network are likely to be less frequented as they are seldom part of a shortest path through the network.

Simulation of Urban Mobility (SUMO)

SUMO is a microscopic traffic simulator developed at the DLR (German Aerospace Center) in Berlin, Germany [141]. It is a free and open-source simulation tool that is widely used in vehicular network research and is also the traffic simulator used in this thesis. Among others, it supports the Krauß and IDM car-following models along with several lane-change models.

A simulation scenario in SUMO consists of a road network and a traffic demand (or route) file. The road network can either be completely synthetic and built with tools that come with SUMO, or it can be obtained from converting map data from OpenStreetMap [109], Vissim [90], MATsim [8], and many more.

A simple road network in SUMO is an XML-file consisting of `edges` (streets) and `nodes` (junctions and connections between streets). An edge connects two nodes and may have multiple lanes (with assigned speed limits). To represent intersections, the `junction` element is used to describe the area where roads cross including right-of-way rules, and `connection` elements are used to describe which outgoing lanes can be reached from an incoming lane. Additionally, traffic lights and their programs can be modeled in detail and can be assigned to connections. Buildings or parking areas are not part of the road network as they do not affect the mobility of vehicles. Instead, they are visualized using the `poly` element and are assigned a certain type that can be referenced later, for example, when importing buildings into the network simulator. A screenshot of SUMO showing a zoomed-in region of a full featured road network is shown in Figure 2.13.

Traffic demand can be modeled in different ways as shown in Listing 2.1. It is possible to define different vehicle types using the `vtype` element, assigning each of them different values for acceleration, deceleration, length, speed, and car-following model parameters: in this example `sigma` for the driver imperfection used in the Krauß model. A `route` is simply a sorted list of connected edges. SUMO allows for both the specific assigning of routes to certain vehicles (in this case, `car1` of type `car` is assigned `route0`) or the specification of traffic flows. In the given example,

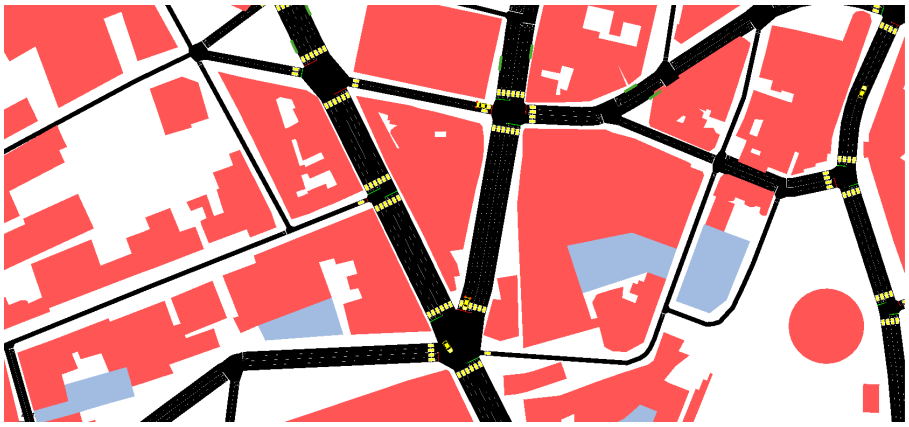


Figure 2.13 – Screenshot of the traffic simulator SUMO [141] using the LuST scenario [35] with modeled buildings (red) and parking areas (blue).

```
<vTypeDistribution id="cars">
<vType id="truck" accel="0.50" decel="3.50" sigma="0.85" ↘
     length="18.33" maxSpeed="33.03" vClass="transport" />
<vType id="car" accel="0.70" decel="3.50" sigma="0.35" ↘
     length="4.68" maxSpeed="36.07" vClass="passenger" />
</vTypeDistribution>

<route id="route0" edges="12 13 14#1 14#2 18"/>

<vehicle id="car1" type="car" route="route0" depart="0"/>
<flow id="cars" type="cars" route="route0" begin="0" ↘
     end="2500" period="1.2" departLane="best"/>
```

Listing 2.1 – Definition of vehicles and routes in SUMO.

starting from 0 s (and ending at 2500 s) every 1.2 s a car from the vehicle distribution `cars` with assigned `route0` is spawned.

In summary, SUMO is a powerful traffic simulator that meets the requirements to be used as a mobility source in vehicular network research. Some limitations remain, such as the error-prone conversion of OpenStreetMap data [16] and the lack of support for 3D maps. However, after careful review of other existing simulators, SUMO was chosen as the traffic simulator to be used throughout this thesis.

2.3.4 Veins: Coupled Mobility and Network Simulation

With the ability to use realistic mobility patterns in network simulation one of the major requirements for vehicular simulations is met. Tools like SUMO can either run at the same time and feed back information to the network simulator about the

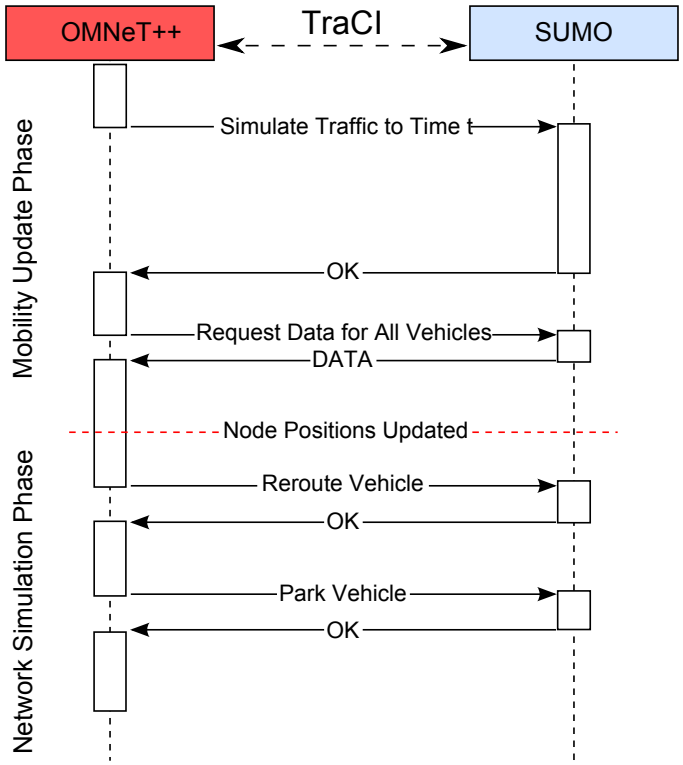


Figure 2.14 – Example information exchange in Veins between OMNET++ and SUMO using TraCI.

nodes' position and speed, or they can be used to generate traces which are then played back in the simulator. The problem with both approaches lies within the very nature of vehicular networks, that is, the goal to influence traffic by means of information exchange. Many traffic safety and traffic efficiency applications that inform the driver of certain traffic or road conditions and can influence vehicles' mobility. A simulation environment where the mobility cannot be changed during run-time has no method to evaluate the effectiveness of these applications. It is therefore necessary to either integrate or bidirectionally couple network and traffic simulators, so that the mobility of the nodes will change the network topology and vice versa. Coupling the simulators instead of integrating them has several advantages: Firstly, using two different simulators that are each maintained and developed by the experts in the respective field is likely to be less error-prone than having both simulators developed by one group. Secondly, coupling two simulators over a common Application Programming Interface (API) makes it easier to individually develop and upgrade each simulator independently as long as the API remains stable.

There are different frameworks that couple network and mobility simulators. Veins [219], developed at the Chair for Computer Networks and Communication Systems at the University of Erlangen, Germany, is the one with the longest track record and one of the most used in the vehicular networks research community. It bidirectionally couples OMNeT++ with SUMO over the Traffic Control Interface (TraCI) to enable node movement in OMNeT++ according to their position in SUMO and also to control vehicles from OMNeT++ in SUMO.

The second focus of Veins is to be an easy-to-use framework for the simulation of vehicular networks by providing example scenarios for different applications including the use of RSUs, traffic rerouting, and so on. Having started as an extension for the INET and MiXiM framework [137], it is now a stand-alone OMNeT++ extension with a focus on peer-reviewed, tested, and validated models for IVC, such as IEEE 802.11p, IEEE 1609.4, and ETSI ITS-G5. Almost all developed models described in this thesis have been published as part of the Veins framework. This does not only help other research groups and industry simulate vehicular networks using the correct models, but also ensures that errors in the models can be identified more easily with the help of the research community.

Traffic Control Interface (TraCI)

TraCI is a binary protocol over TCP with SUMO acting as the server and OMNeT++, or, more specifically, a Veins management module, as the client. It allows the client to control and bidirectionally exchange information with the traffic simulator.

This exchanged information can include vehicle positions and speeds, road topology, or the position of parking spaces and houses. To further allow for route (re-)planning, information on routes, travel times, or even traffic light phases can be exchanged. Commands sent to SUMO include the stopping of a node, setting of a target speed, or assigning of a new route, and can also affect the graphical user interface to, e.g., change the color of a vehicle. As both Veins and SUMO are still in development, the functionality supported by TraCI is steadily increasing to cope with upcoming requirements of new trends in vehicular network simulation. The open-source nature of both tools gives researchers the possibility to extend and change TraCI implementations.

Figure 2.14 shows an example information exchange: Even though each simulator is a separate process, one has always to wait for the other to finish its computations and hand back control. The controlling process is OMNeT++, as it triggers SUMO via TraCI to simulate traffic until a certain time-step has been reached. After SUMO has finished doing so, it will call back OMNeT++, allowing the network simulator to subsequently request the new vehicle positions and speeds to update the network topology accordingly. It is also possible to subscribe to certain properties which

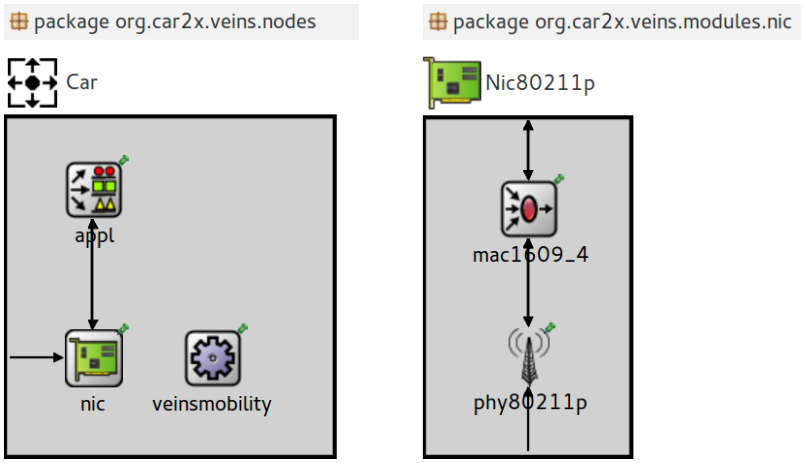
will be reported back automatically by SUMO whenever the traffic simulation has been advanced. After this, the network simulation will take place, and depending on the events, other commands can be invoked to control a vehicle's movement or to retrieve more specific parameters. As the controlling entity is the discrete event simulator, this flow is generated by using events: According to a predefined mobility update frequency, Veins will periodically schedule a specific traffic simulation event that, when handed to the controlling Veins module, will cause the simulation to send a TraCI command so that SUMO advances the traffic simulation.

Modeling in Veins

With Veins now including models for all common IVC technologies and also compound modules for different types of network nodes, setting up a simulation can be done quickly. The vehicle node is shown in Figure 2.15a: It is a simply structured three-layer node, where the application layer is directly connected to the Network Interface Controller (NIC). The `veinsmobility` module shares an interface to the TraCI controller in Veins and is responsible for changing the node's mobility parameters based on the traffic simulation. A node can have multiple applications or even intermediate layers to enable network and transport layer functionality. As an example, Veins already includes an application layer that periodically sends BSMs to exchange safety information with other vehicles.

The NIC is itself a compound module as can be seen in Figure 2.15b: It consists of an IEEE 802.11p PHY and an IEEE 1609.4 MAC. The layers cannot only exchange packets but also control information, e.g., the PHY will inform the MAC when a packet has been transmitted successfully or when the channel becomes idle or busy. EDCA QoS mechanisms and also multi-channel operations are implemented entirely in the MAC. The MAC receives a packet from one of the connected application layers and applies scheduling mechanisms accordingly. When a packet is ready to be sent, it is handed to the PHY and the channel is accessed. For packets received and handed up by the PHY, the MAC will check their destination address and give the packet to the application layer or discard it. Signal attenuation and SINR of a packet will be computed in the PHY of the receiving node. This is done by the help of models that have been forked from the MiXiM framework [137].

Cross layer mechanisms such as security and privacy aspects have to be implemented in each layer separately. Other types of nodes, such as Stationary Support Units (SSUs) or constantly parked vehicles can be modeled by simply replacing the `veinsmobility` with a static mobility. RSUs would be additionally connected to a traffic information center to allow information exchange over a non-IEEE 802.11p interface such as Ethernet or cellular communication.



(a) Compound Module: A vehicle consists of one or more application layers, a network interface and a mobility module

(b) Compound Module: network interface, consisting of MAC and PHY

Figure 2.15 – Model of an IEEE WAVE-enabled vehicle in OMNeT++.

Recently, Veins has been extended to also support heterogeneous networks, that is, vehicles that are equipped with both an IEEE 802.11p device and an LTE transceiver [108]. There exist various other extensions such as ones for the simulation of autonomous vehicles that drive in platoons [206], the consideration of electric vehicles with a focus on realistic battery models [201], or the possibility to control the mobility of vehicles in a more fine-grained fashion, e.g., by commanding them to change their state to free-wheeling or coasting [59]. Furthermore, it has been shown that Veins can also be extended by human driver behavior models [52], moving away from the assumption that supplying information to a node in the network ultimately leads to a change in mobility – the driver can just as well choose not to react to this information, or act in a suboptimal way.

Other Frameworks

Veins is a well-established and widely used simulation framework in both academia and industry. Several other coupled vehicular network simulators exist today – and more are added to cater to specific use cases. They differ in terms of implemented modules, level of coupling, and used traffic or network simulators. Table 2.6 gives an overview of coupled vehicular network simulation frameworks.

TraNS [182] is similar to Veins as it also uses TraCI to couple SUMO with a discrete event simulator, however, instead of OMNeT++, TraNS uses the network simulator ns-2. As the development of TraNS has been suspended, it is based on

Toolkit	Network simulation	Traffic simulation	Coupling
Veins	OMNeT++	SUMO	Bidirectional
TraNS	ns-2	SUMO	Bidirectional
iTETRIS	ns-3	SUMO	Bidirectional
VSimRTI	Multiple	Multiple	Bidirectional
VGSim	JiST/SWANS	Nagel-Schreckenberg	Integrated
NCTUns	(Proprietary)	(Proprietary)	Integrated
SWANS++	JiST/SWANS	STRAW *	Unidirectional
GrooveNet	(Proprietary)	Roadnav *	Unidirectional

* Roadnav and STRAW can use TIGER scenarios that include most U.S. roads and their classifications

Table 2.6 – Summary of simulation frameworks, based on [226].

older versions of SUMO and does not have accurate models for IEEE 802.11p based communication.

The iTetris program [140], which was funded by the European Commission to build a platform for the evaluation of solutions based on ETSI ITS-G5, uses TraCI to couple SUMO with ns-3, the successor of ns-2. It cannot be freely downloaded as it is only available to members of the iTetris community.

VSimRTI [205] goes one step further in the modularization of individual simulators. It is not targeted towards a specific simulation kernel, but provides a generalized and open-source framework for coupling different simulators. Adapters to VSimRTI exist for all major network and road traffic simulators.

VGSim [155] integrates both microscopic traffic simulation and network simulation in one tool using Jist/SWANS and the Nagel-Schreckenberg model to simulate freeway traffic [170]. NCTUns [248] is similar in that it is also an integrated solution for the simulation of traffic and wireless networks. SWANS++ and GrooveNet [165] unidirectionally couple network and traffic simulators and therefore make it difficult to investigate the impact vehicular networks can have on mobility.

Reviewing the existing alternatives in terms of availability, up-to-dateness, extendability, and user community, we chose to use and further extend our Veins simulation framework.

2.3.5 Performance Evaluation

In general, best practices for the evaluation of computer networks also hold for the evaluation of vehicular network applications and protocols. There exist several publications that focus on these practices and can serve as a good guideline [147,

199, 247]. Some characteristics of vehicular networks, however, require special attention.

Often caused by the traffic simulation component, simulations of vehicular networks are prone to having transient simulation phases. These are phases at the beginning of a simulation where the simulation is not in a more or less steady state. Assume a traffic simulator generating a large number of vehicles at the beginning of the simulation: It will take some simulation time until the simulated traffic will reflect real traffic conditions, as vehicles need to disperse from their initial positions before traffic levels on main roads become realistic. A vehicular network application studied in this transient phase will, e.g., experience a different level of packet loss than when studied after the traffic simulation has become more or less steady. Thus, accurately detecting (and discarding) transient simulation phase(s) is particularly important.

Another direct consequence from using traffic simulation as a source for node mobility is that the traffic has a significant influence on the network performance. From this it follows that multiple independent replications of a simulation must not only choose different random seeds for the network simulation, but also for the traffic simulation. The effect of randomness on the communication side of the simulation is often negligible compared to the effect different traffic patterns have on the simulation outcome. From this it follows that applications or protocols should be investigated in different scenarios (e.g., changing traffic density, different road topology, new routes) rather than being repeated over and over again using one specific mobility scenario.

In decentralized, distributed environments such as vehicular networks, fairness is often a fundamental requirement. The methods and metrics chosen to represent the system's performance should show whether a system performs very well for *some* vehicles at the cost of degrading the performance for others. Imagine a PET that protects the privacy of half of all vehicles very well, but completely discloses the other half. Showing only the mean values would entirely hide this bimodal distribution. There are several methods to visualize fairness, including showing error bars, variances and standard deviations, confidence intervals, simple scatter plots, histograms, or the plotting of Empirical Cumulative Distribution Functions (ECDFs). Unfortunately, these methods are often neglected as we will show in the next section, reducing the meaningfulness and interpretability of presented results.

Choosing the right metrics is a difficult task. The diversity of vehicular network applications makes it almost impossible to introduce general-purpose metrics and requires tailored metrics for some fields. Network performance metrics, e.g., packet loss or latency, are possibly not the best choice to indicate the actual performance of, say, a safety or privacy algorithm. For example, traffic safety applications could be assessed with regard to the number of prevented traffic accidents, while privacy

mechanisms could be evaluated by measuring the time a vehicle can be tracked through the network by some kind of adversary. It should be ascertained that the metrics selected describe every important aspect of the evaluated system. For example, in the case of MAC protocols, the properties of the evaluated system can often be divided into three parts: timeliness, efficiency, and robustness. It is trivial to change an existing scheme with good balance between these three to become better in one department by sacrificing performance in another. Staying in the scope of the MAC example, it is easy to reduce the number of collisions (robustness) when at the same time increasing latency (timeliness) and decreasing throughput (efficiency). Selectively presenting metrics that only cover one (or two) of these fields would then give the misleading conclusion that the new scheme outperforms the existing one.

In general, for the sake of reproducibility it is important that all relevant parameters, scenarios, and used software are properly reported. Especially in the field of privacy research, fair reporting, easy interpretability, and good comparability of results are the fundamental requirements to help PETs become an integral part of future communication systems.

2.4 Measuring Privacy Using Simulation

Just as for general vehicular network applications and protocols, discrete event simulation has become a popular tool to evaluate the performance of PETs. Often, the system is evaluated using tailored, very complex metrics in specific scenarios with various assumptions regarding mobility and the adversary to protect against, and a vast number of parameters.

In this section, which is based on our conference paper "Privacy Assessment in Vehicular Networks Using Simulation" [246],[3] we explain the common methodology in terms of metrics and adversary models. In addition, by performing a systematic literature review, we investigate the current state of the art in privacy simulation and discuss current trends, along with benefits and shortcomings of different methodologies.

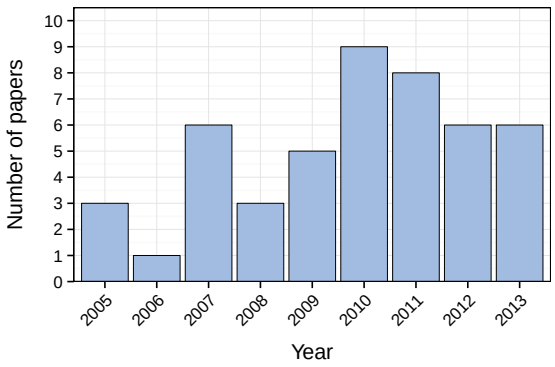


Figure 2.16 – Surveyed papers by year using discrete event simulation for privacy evaluation in vehicular networks.

As part of this, we wanted to ensure that the literature reviews followed a fixed set of steps to encompass as much of the relevant literature as possible, and at the same time reduce the influence of biases the authors may have. Kitchenham adapted a three-phase approach to literature reviews from its origin in medical research to computer science [129]. The first phase is the planning phase, where a set of research questions is specified, inclusion/exclusion criteria are defined, and the search engines and keywords to be used are determined. In the second phase, the conducting phase, the literature search is carried out, and papers are classified according to a taxonomy, including an assessment of paper quality. This phase also includes data synthesis and answering the research questions. The reporting phase is then concerned with writing up and publishing the results.

[3] This paper was written in collaboration with Isabel Wagner from the University of Hull, England. As both authors contributed equally to all parts of the work, it is difficult to single out contributions made by one particular author.

Out of 180 initial candidate papers we identified 48 papers (30 conference papers, 18 journal publications) that evaluated PETs using discrete event simulation. More information about the methodology and the literature review along with a spreadsheet containing full details for all selected papers can be found in [246]. Figure 2.16 shows how the years of publication are distributed between 2005 and 2013.

2.4.1 Adversary Models

As explained by Díaz [45], the evaluation of PETs strongly depends on the specific capabilities of the adversary against which the system under investigation is protected. Naturally, the results of simulation studies can change dramatically based on the chosen adversary. The adversary can be classified along five different dimensions:

Internal vs. external describes whether an adversary is part of the system or not. In vehicular networks, a possible internal adversary can be a participating vehicle, a road side unit, or the system provider. The main difference to external adversaries is that internal ones may have more attack possibilities because, e.g., they can read encrypted messages when they are part of a cryptographic group, they possess a signed certificate, or they are trusted by other participants. External adversaries can be seen as outside attackers who try to compromise users' privacy in a system they are not part of.

Local vs. global refers to the coverage or geographic extent of an adversary's operations. In a vehicular network, a global adversary has access to all communication in the system, i.e., can overhear all messages sent regardless of their geographic position. This only implies the ability to record these message but does not suggest that an adversary can also decrypt them in the case of confidential conversations. A local adversary, on the other hand, could be operating an access point at a single location and is therefore restricted to overhearing or sending messages within a certain range. This dimension of adversary type is not binary, as an attacker could operate distributed multiple access points to increase their coverage.

Active vs. passive describes whether the adversary is only passively observing the system or actively participating in it. Passive adversaries include the deployment of packet sniffers along the road to overhear messages sent by vehicles to, e.g., track them throughout the network. These adversaries are hard to detect as they do not have a perceivable influence on the system they are attacking. An active attacker could send specific messages to trigger responses from the system's participants. This can include forged messages (e.g., by a traffic information center) that lead to the disclosure of private information (e.g., the current destination of a vehicle). These attacks can be very effective if the attacker is an internal one who is trusted by other vehicles.

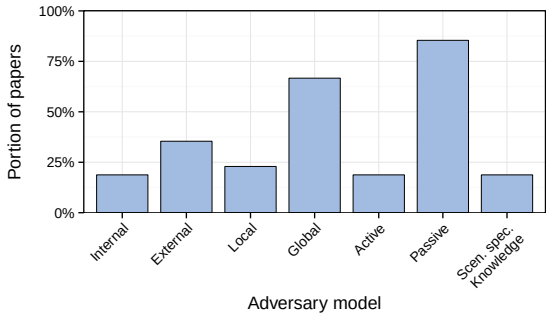


Figure 2.17 – Adversary models used in the surveyed publications.

Static vs. adaptive refers to the adversary's strategy and behavior. A static adversary will start with a certain strategy and will not change it during the attack. This does not necessarily mean that this type of attacker is easy to defend against, as the initial parametrization or strategy might already be effective. An adaptive adversary learns about the system and may adapt their strategy or change parameters accordingly, for example, by deriving turning probabilities at an intersection to increase tracking accuracy based on previous observations.

Prior knowledge describes the level of information about the attacked system the adversary has before starting an attack. We identify three different types of knowledge: no prior knowledge, domain-specific knowledge, and scenario-specific knowledge. A general-purpose attack can work with no prior knowledge about the system, e.g., a tracking algorithm must not necessarily 'know' what objects are tracked. In the context of vehicular network research, attacks are often domain-specific, that is, the adversary knows that they try to track vehicles and can therefore use specific or even tailored algorithms. This knowledge includes but is not limited to WLAN frequencies, channel allocations, used technology, and boundaries of vehicular movement. Scenario-specific knowledge describes all information that is specific to a certain situation or attack use case. For example, an adversary trying to track vehicles through an intersection can take advantage of considering certain turn restrictions of this specific intersection. This can extend to initial identities of nodes, city layouts, or statistics about traffic flow.

In Figure 2.17 we show the state of the art in terms of used adversary models in vehicular network privacy research. Please note that the bars of one dimension may not necessarily add up to 100 % as often papers did not classify their adversary in the specific dimension or evaluated multiple scenarios covering different adversary types. Our results clearly show current research focuses on global ($\approx$ 70 %), passive ($\approx$ 80 %) adversaries, that is, eavesdroppers who can listen in to all of the communication in the network.

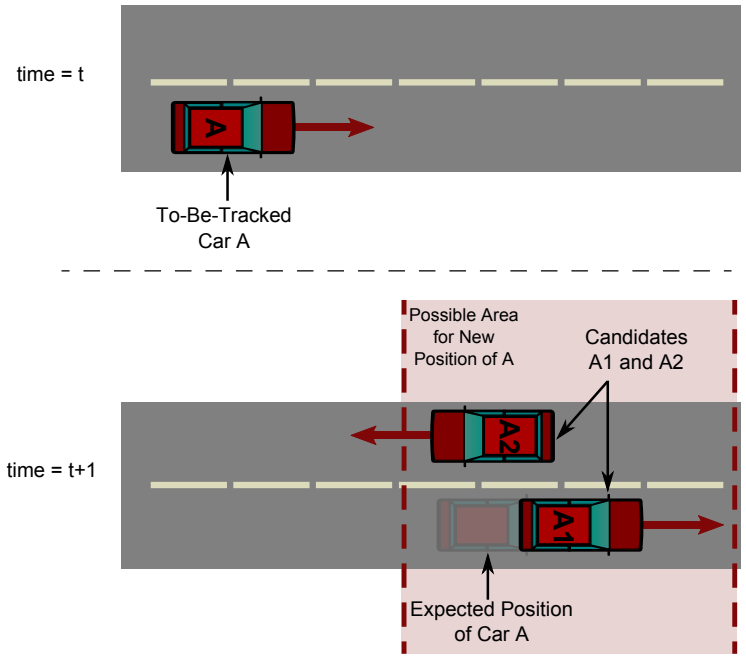


Figure 2.18 – Example situation for the illustration of privacy metrics.

Arguing that this is sufficient because a global passive adversary represents the strongest kind of attacker is invalid as there is no obvious ordering in the strength of an adversary. For example, the relationship between an external, eavesdropping adversary and an internal adversary who might have access to some of the cryptographic material (e.g., participating vehicles or providers of RSUs), or an active adversary who might be able to alter communications, is not immediately clear. Investigating or discussing the effects of other adversary types is a gap in existing research.

We observed that no paper classified their adversary models along the static vs. adaptive dimension. From this we conclude that all adversaries used were of static nature, as attacker adaptivity is certainly a feature that would have been described in the paper. Assuming that an attacker will not learn and adapt might lead to an overestimation of the effectiveness of examined PETs, as we believe real-world adversaries would in all likelihood try to optimize or change their attack parameters.

Only about one in five papers considered the effect of scenario-specific knowledge. However, in the age of big data, it is becoming increasingly likely that an adversary would have access to data that they might be able to correlate with their own observations. This may either increase their chances of success, or allow them to draw surprising new conclusions. Investigations in this direction seem to be promising areas for future research.

2.4.2 Privacy Metrics

Assigning a number to the level of location privacy enjoyed by drivers is a challenging task, and there is a large amount of different privacy metrics used in discrete event simulation that try to achieve this. There is no common agreement on which privacy metric to use, and there seems to be no single metric that satisfies all requirements of different research groups. The investigated privacy metrics not only differ in what privacy property they evaluate but also in terms of complexity and their designated target audience. In the context of vehicular networks, these applied metrics can be grouped into five main categories: anonymity set size, entropy, adversary's success rate, statistics on pseudonym changes, and maximum tracking time. We use the example illustrated in Figure 2.18 to explain their functionality.

Anonymity Set Size (ASS)

The Anonymity Set Size $|AS|$ describes among how many other vehicles it is not possible to distinguish a target vehicle. The advantages of this metric lie in its simplicity and ease of calculation. Consider the example in Figure 2.18 and that at time t the adversary knows that vehicle A is the target vehicle. In this moment the anonymity set size is 1 meaning the driver's location is fully disclosed. Now further assume that between time t and $t+1$ vehicle A changes its pseudonym to $A1$ and another vehicle with pseudonym $A2$ appears. Even though it is very unlikely that A has turned around and is now using pseudonym $A2$, it cannot be completely ruled out, leading to an anonymity set size of 2 because AS then consists of both $A1$ and $A2$.

This assumption that all vehicles in the anonymity set are equally likely to be the target is probably the biggest problem of this metric as discussed in [207]. In addition, the metric measures only anonymity, and disregards the other privacy properties. Also, there is no way to represent an adversary's prior knowledge. Further, this metric heavily depends on the total number of vehicles in the simulated scenario, making it difficult to compare results from different simulation studies. In general, the concept of the ASS is similar to the well-established k-anonymity metric [230] which describes that a specific database record is indistinguishable from k other records. The literature has shown that records can be de-anonymized even if k-anonymity is fulfilled [210], and also that k-anonymity for location privacy, i.e., the anonymity set size, is insufficient for similar reasons.

We observe that about one third of the papers in our survey use the anonymity set size as a metric to evaluate privacy (see Figure 2.19). Given the criticism that has been directed at this metric for more than a decade now (and the number of viable alternatives in the literature) it is somewhat surprising that it is still in such

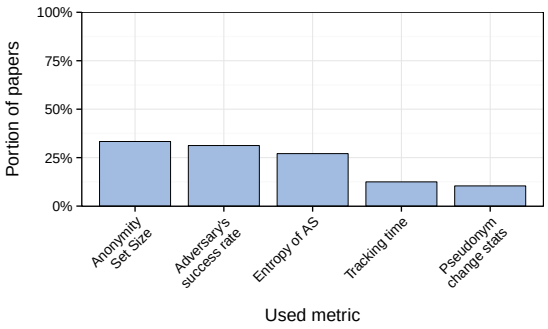


Figure 2.19 – Privacy metrics used by the surveyed papers.

widespread use. As the first group of bars in Figure 2.20 shows, there is no visible decline in the use of this metric over the years.

Entropy

To account for the fact that not all vehicles in the anonymity set are equally likely to be the target, authors have turned to entropy as a privacy metric. It purely focuses on the anonymity dimension of privacy, disregarding other privacy properties. Entropy is a concept from information theory expressing the uncertainty in a random variable. The entropy of an anonymity set can therefore represent the adversary's beliefs about the likelihood of individual vehicles and reaches its maximum when all members of the anonymity set are equally likely to be the target. Formally, the entropy $\mathscr{H}(X)$ is expressed as:

$$\mathscr{H}(X) = -\sum_{i=1}^{|AS|} p_i \cdot \log_2(p_i) \tag{2.13}$$

where $|AS|$ denotes the size of the anonymity set and p_i commonly refers to the adversary's estimation of the probability of i being the target vehicle X. Note that the entropy can also be used for other types of uncertainty, e.g., p_i could be the adversary's uncertainty in assigning trips to individuals. Looking back at the example in Figure 2.18, the entropy can now reflect the unlikeliness of A having turned around. For a frame of reference we first assume that the adversary is completely unsure whether $A2$ or $A1$ is the new pseudonym of A. In this case the entropy would yield $0.5 \cdot \log_2(0.5) - 0.5 \cdot \log_2(0.5) = 1$. Now assume a smarter adversary that assigns a turn-around probability of 1 % and therefore a likelihood of 99 % of A being $A1$. The entropy would then yield $0.01 \cdot \log_2(0.01) - 0.99 \cdot \log_2(0.99) \approx 0.08$. It can be seen that in this particular example the entropy can reflect the actual privacy enjoyed by the driver of A considerably better than the anonymity set size.

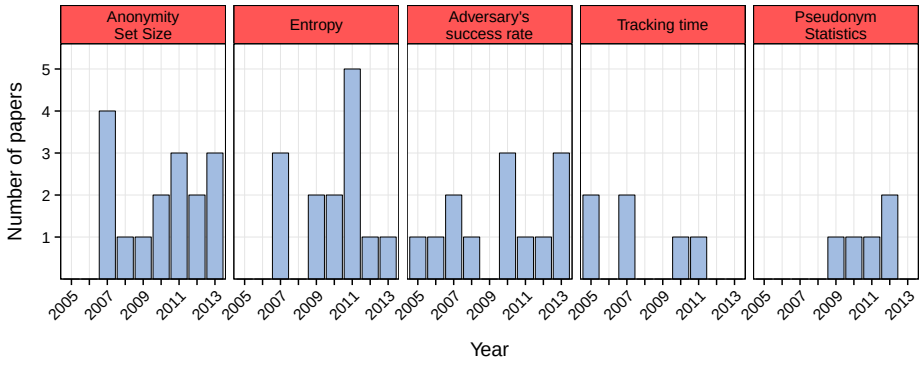


Figure 2.20 – Privacy metrics by type and year.

Similar to the anonymity set size, entropy also depends on the absolute number of vehicles in the anonymity set and therefore in the scenario. One possible solution to this was proposed by Díaz et al. [46], the so-called degree of anonymity. It normalizes the entropy using the maximum possible entropy value $\mathcal{H}_{\max} = \log_2(|AS|)$, resulting in a range of $[0, 1]$ for the degree of anonymity $d = \frac{\mathcal{H}(X)}{\mathcal{H}_{\max}}$.

Another drawback of entropy was discussed by [115]. They argue that entropy can give misleadingly high values, because even if an adversary cannot distinguish between two vehicles, if those two vehicles are in fact standing next to each other then the adversary has still successfully inferred the target's location.

As can be seen in Figures 2.19 and 2.20, entropy is a popular metric to be used in privacy research. The vast majority of papers we classified as using entropy computed the entropy using the adversary's belief whether a specific vehicle is the target as the input probability. In most papers, the distribution of this probability depended on the strength of the adversary or the employed tracking algorithm, however in almost 25 % of papers the probability distribution was assumed to be uniform. In these cases entropy does not carry any additional information compared to the anonymity set size.

Adversary's Success Rate

The adversary's success rate is an unspecific general-purpose metric that expresses to what extent an adversary is able to achieve their goal. It is often used as an easy-to-understand number that can be used to measure any of the privacy properties, depending on the specific way it is defined. This leads to a large variety of different metrics that make comparisons between studies difficult (if not impossible), as authors usually define their own adversary with specific goals in unique scenarios.

Even in the small example shown in Figure 2.18 we can define different adversary goals: If the goal was to track vehicle A with 100 % certainty, then the adversary

would not be successful in this case as they cannot completely rule out $A2$ as a possible successor of A. If the goal was to track vehicle A only with high certainty then the attack could still be counted as successful.

Of the 15 papers (see Figure 2.19) in our survey that utilized some form of an adversary's success rate, we identified eight different variants, a combination of ratios, probabilities and chances relating to successful attacks, adversary guesses, vehicle identification, tracking, pseudonym mapping, anonymity set sizes, and so on.

While the adversary's success rate may be very scenario-specific and imprecise it can still serve as a good method to illustrate and indicate how privacy-preserving a system is if the adversary and their goals are described clearly. Especially when communicating to a broader audience it can be beneficial to use an easy-to-grasp metric such as the adversary's success rate.

Maximum Tracking Time

The maximum tracking time also measures how successful an adversary is at attacking the system, but is much more specific as it defines exactly what the adversary's goal is, namely, tracking vehicles for as long as possible. As such, the metric focuses on the unlinkability property, that is, being unable to link two messages sent with different pseudonyms.

The metric assumes that the adversary will eventually be (completely) confused by pseudonym changes and measures the time until this happens. It does not consider the possibility that an attacker might be able to relink vehicles at a later time or become less confused.

The definition for the maximum tracking time varies. Some authors define it as the time an adversary can track the target vehicle with 100 % certainty, meaning the time the anonymity set size for this target vehicle remains 1. Others have defined it as the time until an adversary is successfully tricked into believing that another vehicle is the target vehicle. The latter definition is also sometimes called the *Maximum (or Mean) Time to Confusion* [116]. Depending on the definition, the outcome of Figure 2.18 would be different. The first definition would result in stopping the time at $t + 1$, the second definition would not stop the tracking time when $A1$ is in fact vehicle A. It can be seen that the latter definition requires global knowledge about the simulation as it compares the adversary's guesses with the actual identities.

According to the fourth group of bars in Figure 2.20, the maximum tracking time seems to have fallen out of use recently; the latest paper using it dating from 2011. Even though this metric has some disadvantages, its strength lies within its public outreach capabilities as it is easy to understand and market. Just like the

adversary's success rate it can be the right choice depending on the publication venue and audience.

Statistics on Pseudonym Changes

As the use of pseudonyms is the main PET in vehicular networks, it is a straightforward choice to use tailored metrics based on different statistics regarding pseudonym changing. This can include the total number of changes, the frequency with which they are changed, the number of successful and failed changes, and so on. Six papers in our survey used metrics from this category.

These metrics can only be seen as an indirect method of evaluating privacy as they do not actually measure to what extent the drivers' privacy is protected but only how well a particular mechanism is working. The strict focus on pseudonyms limits these metrics to only investigating unlinkability as a privacy property.

Their advantage is that they work independently from the adversary model and therefore allow comparison between different pseudonym changing strategies. This advantage can also be seen as a downside as the informative value of the metrics is limited and their applicability restricted to one certain type of privacy mechanism.

Other Metrics

The large amount of metrics used in the surveyed papers did not allow for a complete classification as many were not used frequently enough to merit their own category. Many of them are rather complicated or hard to understand, which might also be the reason why they have not been adopted by other researchers. Some are combinations of existing metrics, such as the entropy or anonymity set size combined with user-specific parameters or information accumulated by an adversary [159]. Others are specific to a particular privacy protection mechanism and as such have limited applicability; examples include the silent time to be observed after a pseudonym change or the number of destination locations revealed.

Several metrics used in analytic evaluations have not yet found their way into the context of simulation studies. A good survey on the topic is [209]. As an example, [95] argue that entropy-based measures do not suffice to measure unlinkability and therefore introduce the *expected distance unlinkability* measure that accounts for the 'inner structure', i.e., the similarity between the adversary's choices, and the robustness of these choices. Another example is the expectation of the adversary's distance error introduced by [115] capturing how well an adversary is capable of estimating a user's position. Metrics from other application areas also show promising potential to be adopted into simulation studies. Examples include mutual information [33] and differential privacy [54]. Mutual information is an information theory concept that could be used to measure the commonalities between a user's

real location trace and the adversary's estimation of it. Differential privacy was originally developed for use in statistical databases, but could be used to allow a vehicle to gauge whether the data it transmits will violate the user's privacy level.

2.4.3 Presentation and Reproducibility

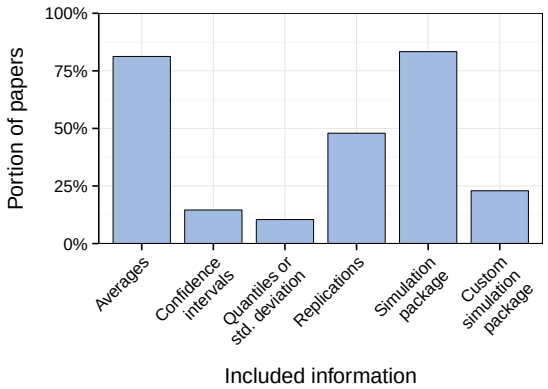


Figure 2.21 – Information reported for the simulations carried out in the surveyed papers.

To further understand why privacy concepts and protection mechanisms have not yet become a fundamental component of FOTs and ITS standards, we looked into the way results are presented in papers evaluating privacy in vehicular networks.

We defined criteria to indicate the quality and reproducibility of simulation studies. We classify papers regarding whether the authors reported only on average values or also introduce confidence intervals, quantiles, or standard deviation as measures for statistical error. The importance of reporting confidence intervals and measures of statistical error has been discussed in [178], which states that simulation studies lacking these features risk their credibility. This can then lead to non-consideration of these papers in other research groups or projects. Furthermore we noted whether the number of replications and the used simulation package was reported (and if the package consisted of publicly available models), an important factor for the reproducibility of reported results.

We observed (Figure 2.21) that only about one quarter of papers report confidence intervals or a measure of statistical error (quantiles or standard deviation). This means that the validity of simulation results is questionable for three quarters of the surveyed papers [178]. 80 % of the papers report only averages, which is problematic as it obfuscates information about the result distribution. This is especially true in the context of privacy where fairness is an important factor. Only 60 % of these papers claim to have conducted independent replications, leaving 40 % for which it remains unclear how these values were actually derived. In total, one out of five

papers present results from a single simulation run, effectively drawing conclusions based on a data set of size 1.

About 80 % of the papers report on the used simulation package, but many fail to indicate whether the used models are publicly available or how to obtain them. We observe that, even though some include an URL to a model library, the given URL is no longer working, emphasizing the need for archiving and maintaining used simulation models. About a quarter of the papers used custom, non-public simulation packages, and thereby make it impossible for other researchers to repeat and reproduce the presented simulation results. Apart from one, all papers report on simulation parameters in at least some degree of detail.

In general, we conclude that none of the surveyed simulation studies are in fact easily repeatable for non-involved authors, strongly hindering their integration in standards or FOTs. An important step towards reproducibility is a common simulation framework for privacy evaluation. This should not only include peer-reviewed models for the underlying technology such as ETSI ITS-G5 and IEEE WAVE, but also implementations of different adversary models and metrics alongside a set of scenarios, tracking algorithms, and a list of sound, predefined parameters, released under a permissive open-source license such as the GPL. Articles could then simply refer to this framework, name the adversary and scenario, and only list parameters that were different from the default configuration. Ideally, the used parametrization, model configuration, or even raw simulation results should be made available online to allow other researchers to build on, evaluate, and improve published PETs. This could be done by the help of digital online repositories at universities for easy long-term sharing of simulation setups. Furthermore, we believe that the use of meaningful and comprehensible privacy metrics is just as important to make the nebulous concept of privacy easier to grasp, paving the way for significant improvements in privacy protection mechanisms.

Chapter 3

Building a Privacy Simulation Framework

This chapter describes the building of a privacy simulation framework on top of the Veins simulator.

In Section 3.1 we explain the basic components of the simulation framework that are required to evaluate PETs in vehicular networks. We identify the need for detailed models for IEEE WAVE and ETSI ITS-G5 communication models which we present and evaluate in Section 3.2 and Section 3.3, respectively. These models allowed us to identify shortcomings of the envisioned systems and therefore contributed to improving them before standardization is finalized.

Parts of this chapter are based on our articles published at the *Vehicular Technology Conference* (*VTC2012-Spring* and *VTC2014-Fall*) [60, 65], the *Wireless On-demand Network Systems and Services Conference* [61] as well as other workshops and conference publications [62, 218].

3.1 A Privacy Simulation Framework

The lack of an open privacy simulator to evaluate PETs that try to complicate tracking by an adversary makes it almost impossible to compare different privacy protection mechanisms or to reproduce results from other research groups. A privacy simulation framework should consist of a comprehensive, yet manageable set of models that can be used to assess the level of location privacy enjoyed by drivers in an ITS. For usability and maintenance reasons it was the natural choice to develop these models on top of our already well-established Veins simulation framework.

The privacy simulator consists of three different building blocks: a set of scenarios (city scenarios, freeways, traffic circle, intersection, etc.), a set of implemented metrics and, most importantly, the tracking module. We will show that, when these are paired with detailed models for radio communication (including IEEE WAVE and ETSI ITS-G5 PHY and MAC, both developed as part of this thesis), it is possible to analyze the effectiveness of pseudonym changing strategies. The simulator can also be used to assess the negative impact of disclosing additional information, e.g., including overly accurate vehicle dimensions or other identifying information in periodic messages.

3.1.1 Vehicle Tracking

The level of privacy provided to drivers in a vehicular network can only be given with regard to a certain adversary. In the context of pseudonyms, and assuming the data sent by the vehicles is the only observable information, the adversary will try to break the pseudonymity and link different pseudonyms to be able to track a vehicle through the network. The vehicle tracking component of the simulator can therefore be seen as the adversary and the evaluation of different PETs is how well they are protecting the drivers' privacy against this tracking component. If the tracking algorithm used is not sophisticated enough, results given by the simulation will be misleading as a real adversary could have used a more advanced tracking algorithm. The tracking algorithm should also not use data that is not (or cannot be) available to an adversary as it would then underestimate the provided level of privacy.

There exists a large number of different tracking systems [18], many of them designed for specific purposes. In the field of vehicle tracking, it is common to use a tracking system design as depicted in Figure 3.1. The starting point for each tracking system is always a set of observations $O = \{o_1, \ldots, o_n\}$ made by an adversary. An observation can be obtained in various ways and it can consist of an arbitrary amount of information. For example, observations made by an adversary who set up a camera system on an intersection would consist of timestamps, positions, colors,

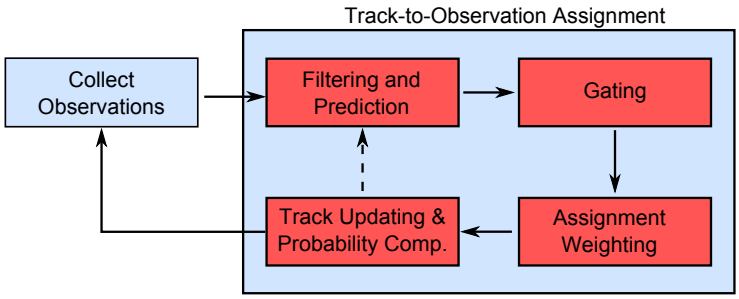


Figure 3.1 – Structure of a vehicle tracking system.

and object dimensions, while observations obtained using a radio receiver would include information contained in the received message and other information the adversary can correlate. In the context of vehicular networks this applies to all information contained in the periodic BSMs (or CAMs, respectively) sent by all vehicles (see Section 2.1).

An adversary relying on information received over the radio channel has to be modeled carefully. If the adversary is believed to be global with access to all information, every sent message by a vehicle can be overheard and therefore handed to the tracking system. This approach may be valid when investigating a worst-case passive attacker, but will not produce meaningful results when simulating an actual, realistic attacker who has set up different radio receivers to collect information emitted by vehicles. This adversary will also experience packet loss, path loss, fast fading, or radio shadowing and may therefore not be able to completely receive all messages in the areas covered by their radio receivers. Especially in dense traffic, vehicles will influence each other in terms of radio transmissions introducing interference, packet collisions, or latency fluctuations. This emphasizes the necessity to deploy realistic channel and radio models, but more importantly, they are required to properly simulate pseudonym changing strategies that incorporate information received by other vehicles to determine whether or when the pseudonym will be changed.

The goal of the adversary is to create a track for each vehicle. A track T_i is a finite sequence of observations, e.g., sent messages, that the adversary believes belong to the same vehicle. The problem of tracking can now be defined as finding the correct observation that belongs to an existing track. This is illustrated in figure Figure 3.2: Assume an adversary has already successfully tracked three vehicles using observations made at time $t = 1$ and $t = 2$. At time $t = 3$ the set of observations O includes three received broadcast messages. All tracks and observations are used as the input for the tracking algorithm to assign an observation $o_i \in O$ to a track T_j, or if not possible, to start new a track or end an existing track.

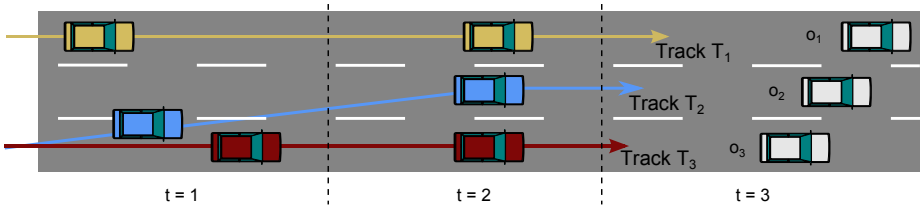


Figure 3.2 – Tracking can be seen as the problem of assigning a new observation o_i to a track T_j.

Filtering and Prediction

The first step in a tracking system is to filter the collected observations. Observations made with the help of sensors are usually subject to noise and are therefore inaccurate to some degree. Depending on the type of sensor and noise, there exist different filter mechanisms to adjust the readings and thereby increase their accuracy. In the context of position data, this is usually done by help of a Kalman filter [127]. The basic idea behind the Kalman filter is to predict the state $\bar{X}_t \in \mathbb{R}^n$ using the previous state X_{t-1} and the known input U_t (a bar over a variable (e.g., $\bar{x}$) indicates that this value is a prediction). The state vector can include position, speed, acceleration, and so on. For the sake of simplicity, let us assume the state vector will only include (x, y) position and velocity. The input vector U_t can then be seen as actions by the driver that directly affect the motion of the vehicle such as accelerating (and decelerating), or steering. With the assumption that these readings are also subject to Gaussian noise q, this leads to a system of linear equations:

$$\bar{X}_t = A \cdot X_{t-1} + B \cdot U_t + q \tag{3.1}$$

The coefficient matrices A and B determine how the previous state and the dynamics of the vehicle will affect the next state. As these can get large with increasing dimensions of the state vector, it is more feasible to, e.g., predict the x and y position separately, using v_x and v_y (and acceleration a_x and a_y respectively), leading to the easily manageable equations of motion:

$$p_t = p_{t-1} + v_{t-1} \cdot \Delta t + \frac{1}{2} \cdot \Delta t^2 a_t \tag{3.2}$$

$$v_t = v_{t-1} + \Delta t \cdot a_t \tag{3.3}$$

The second part of the Kalman filter is the sensor prediction. Based on the result of Equation 3.1, the vehicle can predict what kind of information Z_t it should receive from the GPS sensor, assuming Gaussian noise r.

$$\bar{Z}_t = H \cdot \bar{X}_t + r \tag{3.4}$$

H is referred to as the measurement matrix that transforms the state into the measurement space. Both q and r are zero-mean, white, Gaussian noise with covariance Q and R respectively. Note that in general Q and R may also change over time, but for the sake of simplicity will remain constant here.

The values in these matrices can be derived either from taking measurements and comparing them against real values or, e.g., by taking the accuracy of the sensor (possibly described in the sensor's manual) into account. Based on Equations 3.1 and 3.4 the Kalman equations allow us to **predict** the mean x_t and covariance P_t of state X_t.

$$\bar{x}_t = A \cdot x_{t-1} + B \cdot U_t \tag{3.5}$$

$$\bar{P}_t = A \cdot P_{t-1} \cdot A^\intercal + Q \tag{3.6}$$

These predictions can then be used to **correct** the noisy measurements Z_t to obtain a corrected mean x_t:

$$K_t = \bar{P}_t \cdot H_t^\intercal \cdot (H_t \cdot \bar{P}_t \cdot H_t^\intercal + R)^{-1} \tag{3.7}$$

$$x_t = \bar{x}_t + K_t(Z_t - H \cdot \bar{x}_t) \tag{3.8}$$

$$P_t = (I - K_t \cdot H) \cdot \bar{P}_t \tag{3.9}$$

The factor K is referred to as the Kalman gain (or Kalman matrix) and determines how much the correction term $Z_t - H \cdot \bar{x}_t$ will affect the estimate. Should the predicted GPS measurement match the actual measurement, then this term becomes zero and the corrected mean x_t will be the predicted state $\bar{x}_t$. The noisier the measurement Z (reflected by the covariance matrix R) the smaller will this correction be. After the mean has been corrected using the Kalman gain, the covariance P of state X is also updated. Both values are then fed back into Equations 3.5 and 3.6 for the prediction of X_{t+1}.

In a vehicular network, the position data received by vehicles does not necessarily require filtering. Ideally, the transmitting vehicle itself should already transmit filtered position information as they have direct access to all sensors, that is, all components of the input vector U_t. The need for accurate position information is particularly relevant considering that transmitted position information is an important input for the safety applications of receiving vehicles. Also, the OBUs of the vehicles are expected to have limited processing power, making it possibly infeasible to run a Kalman filter for each neighboring vehicle. If vehicles transmit unfiltered

position information, the adversary can apply filters using the included information in BSM and CAM broadcasts.

Once the observations have been filtered, the adversary predicts (or extrapolates) the next expected observation of each track. Position prediction can be done using Equation 3.1. However, observations may include much more than only position information and every piece of information can be used by the adversary to track a vehicle. For example, if there is an observation with a pseudonym that is already known and part of a track, it is obvious that this observation belongs to the respective track. All these variables can become part of the state X_t and can be predicted by the adversary to track vehicles more effectively. At the end of the prediction phase, there exists exactly one estimated successor state $\bar{e}_i$ for each track T_i.

Gating

Gating is the process of eliminating all unlikely successors for each track to increase the performance of the tracking system by decreasing the overall number of required comparisons between observations and predictions. In addition, in a multi-hypothesis tracking system, it reduces the number of possible hypotheses and thereby also the required memory. It is a per-track operation that identifies all $o_i \in O$ that cannot be used (or have a likeliness below a certain threshold) to continue a track T_j. We refer to the set of all remaining possible successor observations for track T_j as $\dot{O}_j$.

Gating can be done in multiple dimensions, the most obvious one being the geographic one. Assume again the situation illustrated in Figure 3.2: When finding the possible successor for each track, some observations may be neglected because it might have been physically impossible for the vehicle associated with a track to reach the given position.

In the following we discuss all gating mechanisms that are implemented in our privacy simulation framework.

A straightforward approach to identifying possibly unreachable observations is presented in [198]. For each track, only observations lying in an annulus are considered, that is, the overlapping area formed by two concentric circles around the last position of the track. The radii of the two circles are determined by a minimum and maximum velocity ($v_{\min}$ and $v_{\max}$) and are given by $r_1 = v_{\min} \cdot \Delta t$ and $r_2 = v_{\max} \cdot \Delta t$. Determining the minimum and maximum velocities is not trivial and can be done in various ways: For example, $v_{\max}$ can be the legal speed limit plus some margin, however, this would allow vehicles to evade tracking by speeding. Setting the maximum speed to the overall maximum speed thought to be achievable by a vehicle will increase the radius of the outer circle and therefore make the gating process less efficient. The minimum speed $v_{\min}$ can either be set to zero or a reasonable minimum speed according to a certain scenario.

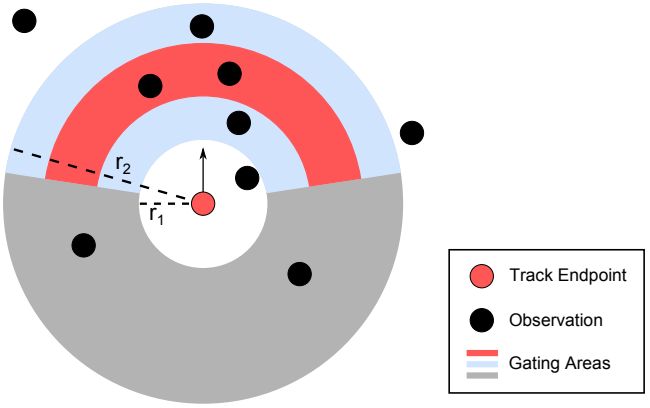


Figure 3.3 – Example illustration of different gating mechanisms: gray = only speed, blue = with angle, red = with acceleration. Only observations within the respective gating area are considered for the continuation of the track.

Due to its simplicity, this gating mechanism may keep various observations that are physically impossible to continue a certain track. This is especially problematic when the probabilities of a track-to-observation assignment are not computed independently but are set to $1/n$ (assuming n possible successor observations). Also, metrics such as the anonymity set size will give misleading results when impossible observations are not eliminated. This leads to the conclusion that this type of gating should not be used as a basis for evaluating PETs going forward, but only to compare results against already existing studies that utilized this straightforward mechanism.

A second step to further reduce the gating area is the consideration of the vehicle heading. As can be seen in Figure 3.3, the first approach (gray area) would allow for the vehicle to have turned around. Depending on the observation time interval Δt, this can be unlikely or even impossible. It is therefore useful to limit the maximum change in heading, that is, the yaw rate $\dot{\psi}_{\max}$, of a vehicle in order to eliminate more observations. The maximum yaw rate of a vehicle, given in degrees per second, depends on the current velocity of a vehicle. Typical maximum values reach from 75 °/s at very low speeds to 5 °/s on freeways. More sophisticated values can be derived using the Einspurmodell [191]. When the observation time is short, this method can result in a considerably smaller gating area (blue region), as all observations with a heading of $|\phi_o - \phi_{\text{track}}| > \dot{\psi}_{\max} \cdot \Delta t$ can be neglected for the continuation of this track. It has to be noted that the feasibility of this approach heavily relies on the underlying vehicle simulation model: For example, some simulators (including SUMO) allow vehicles to turn around almost instantly, possibly leading to a failure of tracking when the maximum yaw rate is used.

The gating area can further be reduced using a method described in [185]: instead of only considering minimum and maximum velocities, it takes the maximum

acceleration $a_{\max}$ and deceleration $b_{\max}$ into account. This leads to a computation of the radii $r_1 = v \cdot \Delta t + \frac{1}{2} \cdot b_{\max} \cdot \Delta t^2$ and $r_2 = v \cdot \Delta t + \frac{1}{2} \cdot a_{\max} \cdot \Delta t^2$. Even with very high values for $a_{\max} = 10ms^{-2}$ this method was shown to reduce the gating area by up to 26 % compared to approaches only considering velocity (cf. Figure 3.3, red region). [185] further states that the maximum radius of the outer circle can be reduced even more when the velocity of the observation v_o is taken into account. Depending on the observation time Δt they assume a worst case where a driver fully accelerates and then brakes down to the target speed v_o. Only if this worst-case maximum distance is greater than the actual distance between the observation and the track end point is the observation taken into account as a possible successor for the track.

Gating is not limited to geometric areas but can extend to all kinds of information. For example, if vehicles transmitted their (most likely rounded) dimensions, all observations (that is, received broadcast messages) containing different vehicle dimensions could be disregarded. In the case of pseudonyms, and assuming pseudonyms are unique and not exchanged between vehicles, an attacker could discard observations with pseudonyms that are already associated with other tracks, eliminating the effect of isolated, non-coordinated pseudonym changes already at the gating stage. In general, it can be said that the gating process is often dependent on the PET or the privacy vulnerability itself. Knowledge about the PET can be used to reduce the number of possible observations and some privacy vulnerabilities may even lead to a situation where an attacker can exclude all observations but one, e.g., when they are able to predict a certain state and only one observation matches their prediction.

Assignment Weighting

After all unlikely observations are discarded for a certain track, the tracking algorithm estimates the likelihood of all remaining observations to continue the track. For that, a rating mechanism is needed. The most obvious rating is to assign each observation in the gating area the same probability [198], regardless of its distance or difference compared to a predicted position $\bar{e}_i$ of the track T_i. In most cases this will lead to a false sense of privacy, maximizing metrics like entropy and reducing the maximum tracking time. This mechanism therefore corresponds to a weak adversary that cannot make use of the information included in the observations.

In the context of location privacy, assigning a rating depending on the geometric distance between the extrapolated position of the track and the observation seems an obvious choice. Authors in [198] suggest to extrapolate a track's position using the last known velocity and heading ($x_t = x_{t-1} + v_{t-1} \cdot \Delta t \cdot cos(\phi)$ and $y_t = y_{t-1} + v_{t-1} \cdot \Delta t \cdot sin(\phi)$), however, this should be extended to also using the vehicle's

acceleration (see Equation 3.2). Each track-to-observation assignment for a track T_i can then be weighted using the Euclidean distances $\|\bar{e}_i - o_j\|$, and a probability can be assigned normalizing the distance using the sum of Euclidean distances $\sum_{o \in \dot{O}_i} \|o - \bar{e}_i\|$ from the track to all observations in its gating area $\dot{O}_i$.

Reducing observations to only their position for the computation of the assignment weights is not preferable as all information contained in an observation can be used by a potential adversary. Therefore [18] and others propose the use of the Mahalanobis distance [160] to incorporate all possible dimensions of the target's state $X_t \in \mathbb{R}^n$. It is defined as $\sqrt{(\bar{e}-o)^\intercal \cdot S^{-1} \cdot (\bar{e}-o)}$ with S being the covariance matrix. Assuming the covariance matrix to be diagonal, that is, the variance σ^2 of each dimension to be uncorrelated, the Mahalanobis distance d_m between the estimated state $\bar{e}$ and an observation o regarding K dimensions of the state can be given in the form of:

$$d_m(\bar{e}, o) = \sqrt{\sum_{i=1}^{K} \frac{(\bar{e}_{[i]} - o_{[i]})^2}{\sigma_{[i]}^2}} \tag{3.10}$$

To expand on the principle of the Mahalanobis distance (or also the *squared statistical distance* or *normalized Euclidean distance*), assume the assignment weight depends on the actual positions p, the velocities v, and the heading ϕ of the estimated state $\bar{e}$ and an observation o. Then the distance becomes:

$$d_m(\bar{e}, o) = \sqrt{\frac{(p_{\bar{e}} - p_o)^2}{\sigma_p^2} + \frac{(v_{\bar{e}} - v_o)^2}{\sigma_v^2} + \frac{(\phi_{\bar{e}} - \phi_o)^2}{\sigma_\phi^2}} \tag{3.11}$$

The variances can be seen as a weighting mechanism for each of the terms, as they reflect the uncertainty of the prediction.

Track Updating & Probability Computations

After the distance for each possible track-to-observation assignment has been calculated, the tracking algorithm has to decide which observation continues which track. This solution has to be unambiguous, that is, one observation must be used to continue only one track. Also, the tracking system can determine whether a new track has started (e.g., if an observation could not be assigned to a track) or a track has ended (if no suitable observation has been found to continue the track for a given time interval). This overall solution is referred to as a *hypothesis* and reflects the adversary's current view of the system.

Finding a solution can be done in various ways, e.g., a greedy algorithm selects the observation with the lowest distance to continue a track, regardless of the overall situation. This means that if an observation o_i exists with a small enough distance

to estimation $\bar{e}_j$, the algorithm would assign o_i to T_j even if this observation was the only one to continue another track T_k. This method is error-prone and produces suboptimal results [18,185]. It is therefore desirable to use an algorithm that obtains a global optimum in terms of statistical distances, that is, assigns observations to tracks in a way that the sum of statistical distances is minimal. Depending on the scenario, a second objective can be introduced, namely maximizing the number of continued tracks. When it is unlikely that a track ends, e.g., on a freeway, it can be preferable to sacrifice one good assignment for two lower-quality assignments.

This problem can be mapped to the auction house problem [18]: First, an $n \times m$ track-to-observation assignment matrix with n tracks, m observations, and the corresponding entries a_{ij} to be values indicating the quality of the assignment of T_i to o_j is created. Finding the global optimum would then be the selection of ≤ 1 entries per row so that the sum of all selections becomes maximal. Alternatively, [185] suggests the track-to-observation assignment to be converted to a graph $G = (V, E)$ with $V = \dot{V} \cup \tilde{V}$ and $\dot{V}$ being the track endpoints and $\tilde{V}$ being all observations. Edges E are a subset of the Cartesian product $E \subset \dot{V} \times \tilde{V}$, making G a directed graph with edges only from track endpoints to observations. Each edge e is assigned a cost c_e depending on the statistical distance between the track endpoint and the observations. The following method requires these costs to be high for small distances and vice versa, for example, by assigning each edge the negative value of the statistical distance [18]. Tracks are only connected with observations in their gating area. The goal is to find a solution $E_s \subset E$ that satisfies the properties that each $v \in \dot{V}$ has an outdegree ≤ 1 and each $v \in \tilde{V}$ has an indegree ≤ 1 while maximizing the sum of all costs $\sum_{e \in E_s} c_e$.

To achieve that the algorithm maximizes the number of continued tracks and does not choose one good matching over two lower ones, [185] suggests adding the sum of all costs $\sum_{e \in E} c_e$ to each c_e so a solution with k edges will always be preferred to a matching with $k-1$ edges. The maximum matching problem can then be solved by the Edmonds algorithm [70] in $O(n \cdot m \cdot log(n))$ time as implemented in the Lemon template library [43].

In a multi-hypothesis tracking system (and also for the computation of hypothesis probabilities) it is required to compute more than only the optimal solution. Having obtained the optimal solution E_s using the above approach, the algorithm described in [169] will produce the n-best solutions (or hypotheses) to the assignment problem. It achieves this by subsequently removing possible assignments $e \in E_s$ from the solution and finding a new solution $\hat{E}_s \subset E$ with $e \notin \hat{E}_s$. For a detailed description please refer to [18, 169, 185]. To increase the computational performance of the following steps, it is recommended to discard $n - k$ solutions with a score below a certain threshold and only keep k hypotheses, resulting in a set of hypothesis $H = \{H_1, H_2, \ldots, H_k\}$.

For the evaluation of privacy protection mechanisms it is necessary to be able to assign probabilities to track-to-observation assignments and consequently to hypotheses. As discussed in Section 2.4.2, this probability is an important input for metrics such as the entropy or success rates. For that, we deploy the Joint Probabilistic Data Association (JPDA) method as described in [18] and [185].

For every assignment of track i to observation j a Gaussian likelihood value g_{ij} is computed. This is done using the statistical distance (cf. Equation 3.10), the number M and covariance matrix S_{ij} of these dimensions.

$$g_{ij} = \frac{e^{-d_m(\tilde{e}_i,j)^2/2}}{(2\pi)^{M/2} \cdot \sqrt{|S_{ij}|}} \tag{3.12}$$

Furthermore assume the set of all selected assignments to be G and P_D to be the probability of successfully detecting an observation (in the context of wireless networks, this can be related to the packet loss rate and the probability of a track actually ending). Then the unnormalized probability of a hypothesis $p'(H_k)$ can be computed using the extraneous return density β (in this case, the density of new tracks in the gating areas), the number of continued tracks m, the number of discontinued tracks e, the number of unassigned observations u, and the product of all assignments $g \in G$.

$$p'(H_k) = (1 - P_D)^e \cdot (P_D)^m \cdot \beta^u \cdot \prod_{g \in G} g \tag{3.13}$$

The normalized hypothesis is then simply computed using all (non-discarded) hypotheses:

$$p(H_k) = \frac{p'(H_k)}{\sum_{h \in H} p'(h)} \tag{3.14}$$

Then the probability of observation o_i to continue track T_j is the sum of the normalized probabilities of the hypotheses $H' \subset H$ in which this assignment is present:

$$p(T_i, o_j) = \sum_{h \in H'} p(h) \tag{3.15}$$

When the number of updated tracks (or the number of unassigned observations) in each hypothesis is the same, Equations 3.12 and 3.13 can be simplified accordingly [185].

3.1.2 Metric Implementation

The importance of easy-to-understand and meaningful metrics has been highlighted in Section 2.4.2. A privacy simulation framework for vehicular networks should

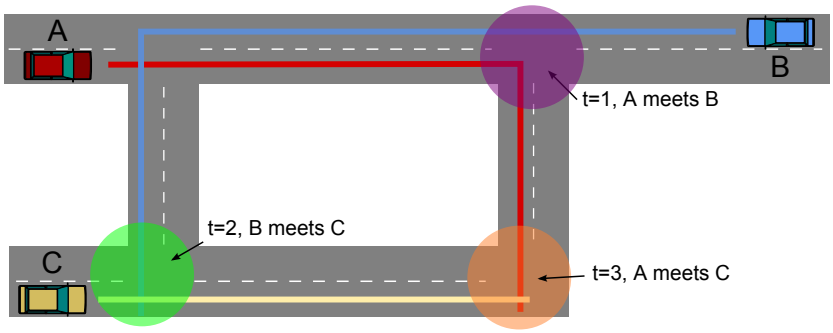


Figure 3.4 – Example scenario for the illustration of metrics.

therefore provide several basic (and extendable) metrics to enable researchers to choose the most suitable one based on their target audience or scenario.

The implementation for most metrics listed in Section 2.4.2 is more or less straightforward. The adversary's success rate and maximum tracking time can be directly measured by comparing the adversary's (or algorithm's) belief of which vehicle is which and the actual identity of the vehicle. Furthermore, statistics based on pseudonyms can be derived from the actual pseudonym changing strategy.

However, measuring the anonymity set size and especially the entropy, despite their simplicity, is not as straightforward as it seems. In the following we will illustrate the problem and how it can be addressed.

Assume a scenario with three vehicles A, B, C as illustrated in Figure 3.4 and an adversary who tries to track their drivers a, b, c. Further, the adversary receives one beacon message per vehicle at each time t_0 to t_3. The actual paths of the vehicles are indicated by colored lines. An adversary does not necessarily know that two beacons sent from A actually belong to the same vehicle. They rather see three beacons (= vehicles) at time t which are used to continue already existing tracks. They then assign possible drivers to these tracks.

In a scenario like this, the adversary may want to answer two different questions:

1. Given a received vehicle broadcast at time t, who are the possible drivers?

2. Given an individual x, which are the vehicles x could be driving at time t?

The answer to the first question is represented by a vehicle's anonymity set $A_{X,t}$ which contains all individuals who could possibly steer the vehicle at time t. The answer to the second question can be derived from the target's anonymity set $A_{x,t}$ which contains all vehicles this target can be possible driving at time t. From this it follows that there exist two different types of anonymity sets, one for vehicles and one for targets. Both sets are closely related but not identical, and in the following paragraphs we will illustrate how.

For the sake of simplicity we will refer to the beacon (or the observation) belonging to vehicle A emitted at t as At. At $t = 0$ the anonymity set for each driver $\in \{a, b, c\}$ contains only one vehicle and the anonymity set for each vehicle $\in \{A, B, C\}$ only contains one driver. Now further assume that at $t = 1$ an adversary cannot distinguish between the vehicles that sent beacons $A1$ and $B1$. This leads to the anonymity set $A_{a,1} = A_{b,1} = \{A1, B1\}$, meaning target a and b could possibly be in either of these vehicles. This also means that the drivers of vehicle A and B can be both a and b, which is represented by $A_{A,1} = A_{B,1} = \{a, b\}$. When vehicles A and B part ways again and B at a later point ($t = 2$) meets vehicle C, this does not only affect A_b and A_c but also A_a. After this $A_{c,2} = \{B2, C2\}$ and $A_{a,2} = A_{b,2} = \{A2, B2, C2\}$. At this point driver c cannot be the driver of the vehicle that emitted $A2$, therefore $A_{A,2}$ remains at $\{a, b\}$. However, it cannot be ruled out that driver a steers the vehicle that emitted $C2$ ($A_{C,2} = \{a, b, c\}$). This shows that based on the perspective, the members of the anonymity sets differ.

time	A_A	A_B	A_C
t=0	1.0a	1.0b	1.0c
t=1	0.9a,0.1b	0.1a,0.9b	1.0c
t=2	0.9a,0.1b	0.09a,0.81b,0.1c	0.01a,0.09b,0.9c
t=3	0.811a, 0.099b, 0.09c	0.09a,0.81b,0.1c	0.81c, 0.091b , 0.099a
time	A_a	A_b	A_c
t=0	1.0A	1.0B	1.0C
t=1	0.9A,0.1B	0.1A,0.9B	1.0C
t=2	0.9A,0.09B,0.01C	0.1A, 0.81B,0.09C	0.1B, 0.9C
t=3	0.811A, 0.09B, 0.099C	0.099A, 0.81B, 0.091C	0.09A, 0.1B, 0.81C

Table 3.1 – Weighted anonymity sets for vehicles and targets assuming a 90 % adversary in the scenario illustrated in Figure 3.4.

For many metrics, and most importantly the entropy, each member of the anonymity set needs to be assigned a probability. For the sake of readability we will now refer to the latest beacon At of vehicle A simply as A. As notation we will simply prefix each member with the assigned probability. Assume the same scenario from Figure 3.4 with an adversary who is able to track vehicles with 90 % certainty when they meet. This means, that when two close-by vehicles send beacons, the adversary assigns a 90 % probability to the correct track-to-observation assignment. Normally, this probability would be computed by a tracking algorithm as described in Section 3.1.1. Table 3.1 shows the weighted anonymity sets for both vehicles and targets after each time-step. After A and B ($t = 1$) met, the anonymity sets for the drivers are therefore $A_{a,1} = \{0.9A, 0.1B\}$, $A_{b,1} = \{0.9B, 0.1A\}$, and $A_{c,1} = \{1.0C\}$. Now when B encounters C at $t = 2$, the anonymity set for A_a has to be updated, accounting for the fact that the 10 % probability of driver a being in vehicle B is now shared between

B and C. This results to $A_{a,2} = \{0.9A, 0.1(0.9B, 0.1C)\} = \{0.9A, 0.09B, 0.01C\}$ and $A_{b,2} = \{0.81B, 0.09C, 0.1A\}$. For the anonymity set of target c the fact that vehicle B could also be vehicle A does not need to be considered, leading to $A_{c,2} = \{0.9C, 0.1B\}$ as B does not represent the entire path of B but only the current position. The same holds for the possible drivers of A, as target c cannot be in A. Therefore $A_{A,2} = A_{A,1} = \{0.9a, 0.1b\}$. However, the history of the track becomes relevant when answering the question who is steering vehicle C because $A_{C,2} = \{0.01a, 0.09b, 0.9c\}$. Now consider vehicle C and A meet at $t = 3$, then

$$\begin{aligned} A_{A,3} &= \{0.9A_{A,2}, 0.1A_{C,2}\} \\ &= \{0.9(0.9a, 0.1b), 0.1(0.9c, 0.09b, 0.01a)\} \\ &= \{0.81a, 0.009b, 0.09c, 0.009b, 0.001a\} \\ &= \{0.811a, 0.099b, 0.09c\} \end{aligned}$$

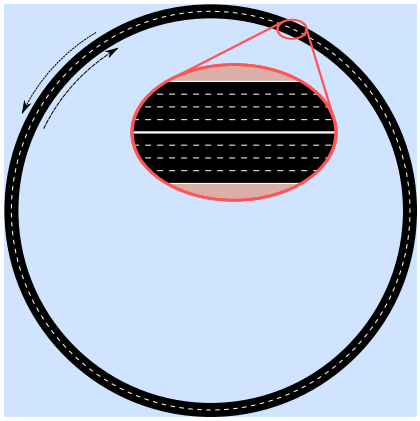
Probabilities for $A_{C,3}$ can be calculated accordingly. $A_{B,3}$ does not need to be updated as the vehicle that emitted this beacon is alone in its vicinity. It can be seen that the weighted anonymity sets for vehicles can be calculated quite easily as they are generated using the previous anonymity sets of the vehicles that cannot be distinguished. This means that when vehicle A meets vehicle B and an adversary tracks with probability P, the new sets are computed by:

$$A_{A,t} = \{P \cdot A_{A,t-1}, 1 - P \cdot A_{B,t-1}\} \tag{3.16}$$

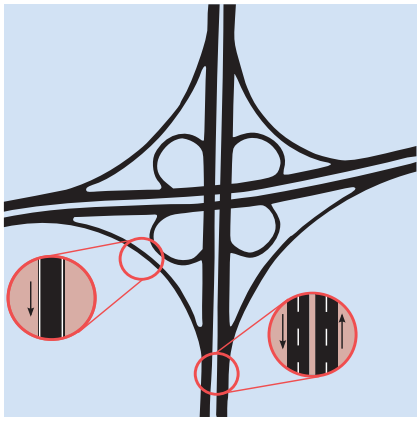
These updates can be computed quite efficiently as they only consider observations and tracks in their respective gating area (as can be seen in Table 3.1, the set of A_A and A_B does not change at $t = 2$ and $t = 3$, respectively.) This is not the case for the update of the targets' anonymity sets, as they are also affected when a vehicle in their anonymity set meets another vehicle. However, due to the fact that the sum of probabilities in one anonymity set always has to be 1 and the sum of the probabilities of one particular target (or vehicle) among all anonymity sets this target (or vehicle) is in also has to be 1, any A_x can be computed using the anonymity sets A_X and vice versa. For example, the members and probabilities of A_a can be derived by simply adding all vehicles in which a can possibly be (that is, the anonymity sets A_X that include a) with the assigned probability. This weighted anonymity set allows to easily derive values for metrics like the entropy, the maximum time to confusion and also the adversary's success rate in scenarios where the adversary has only one guess to determine a target vehicle (by picking the one with the highest probability).

3.1.3 Scenarios

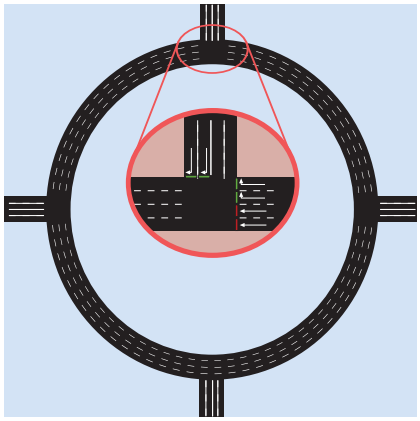
An important step to reproducibility and comparability of simulation studies is the development of common scenarios. For this, we created and include several scenarios in the simulation framework to cover most of the typical environments for vehicular networks, such as sparse and dense traffic, low- and high-speed scenarios as well as steady and dynamic network topologies. They range from primitive, synthetic intersections to scenarios based on actual crowd-sourced map data of entire cities. In this section we describe the scenarios created and used in the simulations carried out for this thesis.



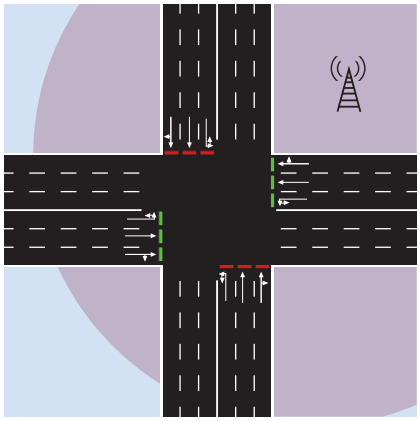
(a) Circular freeway scenario with four lanes in each direction.



(b) Motorway junction scenario with multiple on- and off-ramps.



(c) Large traffic circle scenario.



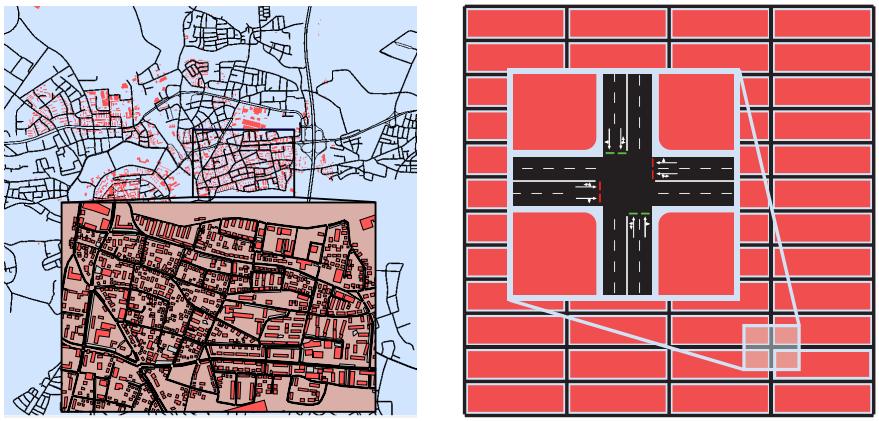
(d) Traffic light regulated intersection scenario.

Figure 3.5 – An overview of developed synthetic small-scale scenarios.

Figure 3.5 shows four completely synthetic scenarios. For dense traffic at high speeds, we created an endless freeway to simulate networks with bimodal connectiv-

ity (Figure 3.5a). Vehicles heading in the same direction experience a rather stable, longer lasting connection due to lower relative speeds, while information exchange with oncoming vehicles is challenging due to high relative speeds and the resulting short connection times. To avoid that vehicles that already passed each other meet again, a Region of Interest can be set that does not include the entire circle. This way, vehicles that leave the ROI and re-enter it will be treated as different vehicles by Veins, eliminating unwanted simulation artifacts. The motorway junction shown in Figure 3.5b introduces connections in-between these two extremes of the freeway scenario while also introducing a very high-density spot where the freeways meet. Note that although SUMO does not support three-dimensional traffic yet, the center of this scenario is not an actual intersection but two overlapping freeways where traffic on one freeway has no effect on the other.

Figure 3.5c and 3.5d are smaller-scale scenarios to simulate low to high traffic densities at lower speeds. In the traffic circle scenario, traffic moves in one direction while in the 4-way intersection each incoming street is connected to all outgoing streets, resulting in a more diverse mobility. The figure also shows an access point set up by an adversary who tries to track vehicles passing the intersection, as used in simulations in this thesis. This feature is optional and can be parametrized in terms of communication range or position.



(a) Suburban scenario based on the city of Ingolstadt, Germany. The extract shows an example ROI.

(b) Synthetic scenario based on the Manhattan grid with large built-up city blocks (red).

Figure 3.6 – Urban and suburban scenarios used throughout this thesis.

To simulate traffic in cities and suburban environments we used the scenarios shown in Figure 3.6. This SUMO road network was created based on crowd-sourced geodata from the OpenStreetMap Project [109] including buildings and parking

spaces in the city of Ingolstadt, Germany.[4] Its purpose is to represent a typical European city with a mixture of sparse and dense traffic. Similarly to the LuST scenario [35] (see Section 2.3.3), intersections were repaired manually to avoid unrealistic mobility patterns caused by conversion errors. For performance reasons, we simulated network traffic only in the illustrated ROI but generated traffic on the entire map to avoid border effects (e.g., unused roads close to the edges of the ROI). The Manhattan grid scenario (Figure 3.6b) was used to simulate a worst-case scenario in terms of obstacle radio shadowing. Built-up city blocks of 80 m · 240 m pose a challenge to many IVC applications as they make communication with vehicles on intersecting or parallel streets almost impossible when not close to or directly on the intersection [218]. This is particularly the case for safety applications when beacons of potentially colliding vehicles are not received until the vehicles are close-by or can actually be seen. Scenarios like this are ideal to demonstrate the potential benefits of relay algorithms to increase situational awareness of vehicles [68].

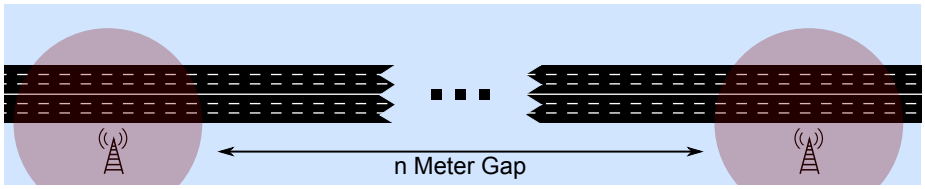


Figure 3.7 – Distributed attack scenario on a freeway with a blind spot between adversary access points.

Lastly, to demonstrate the versatility of our privacy simulation framework we created a scenario where an adversary is not able to completely cover the road network but has placed two access points with an arbitrarily large gap between them (Figure 3.7). They therefore have no knowledge of what is happening in this blind spot but try to track vehicles based on the received beacon messages in the covered areas. Identifying information such as unchanged pseudonyms, sequence numbers, or even WLAN fingerprints can help an adversary link vehicles that passed both covered areas and thereby create interpolated paths.

3.1.4 Demonstration of Capabilities

To illustrate the capabilities of our privacy simulator, we investigate one of the typical questions in vehicular network privacy research, that is, how beneficial pseudonym changing actually is. For this we investigate the fully covered intersection scenario (Figure 3.5d), and the 4-lane freeway scenario with two adversary access points with an 800 m gap between them (Figure 3.7). A full list of simulation parameters

[4] The map for Ingolstadt, Germany, including manual repairs, was provided by Tobias Gansen within the scope of the sim^{TD} project [58]

can be found in Table 3.2. Adversary model and privacy context are classified using the taxonomy presented in Sections 2.2.1 and 2.4.1.

Parameter	Value
Adversary Model	External, local, passive, static, domain-specific
Privacy Domain	Location privacy
Privacy Property	Unlinkability
Data Source	Observable information
Metrics	Adversary's success rate
Scenario	Intersection, blind spot freeway (800 m gap)
Technology	IEEE WAVE
Beacon Frequency	1 Hz
Max Pseudonym Validity t_p	0.5s, 2s, 20s, 50s, 100s, ∞
No. of Vehicles	25-300

Table 3.2 – Setup and parameters for the preliminary simulation study.

In both scenarios we use a non-cooperative pseudonym changing strategy similar to the ones employed in some field operational tests [228]. In our version, vehicles draw a random number $r \in [0, t_p]$ and change all identifying information (source addresses, sequence numbers, etc.) after r seconds. The special case $t_p = \infty$ is used as a validation setting where pseudonyms are not changed and vehicles will use the same source address throughout the simulation. They can therefore be easily re-identified by an adversary. When the adversary receives a message sent with a pseudonym that matches an already existing track endpoint, they can eliminate all other observations for this track and assign a 100 % probability to the given assignment. The adversary was successful if they are able to track a vehicle through the scenario.

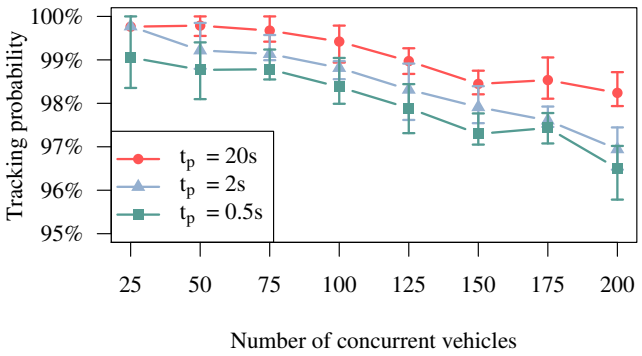


Figure 3.8 – Impact of address changing on average tracking success in the intersection scenario. Error bars show the 25 % and 75 % quantiles over all simulation runs.

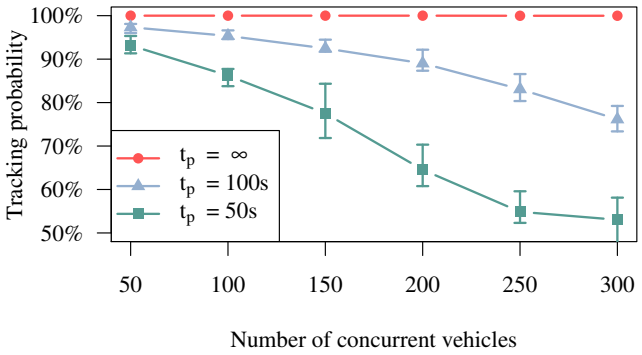


Figure 3.9 – Impact of address changing on tracking in the blind spot freeway scenario. Error bars show the 25 % and 75 % quantiles over all simulation runs.

We observe that it was almost impossible to confuse the adversary in the intersection scenario (Figure 3.8). In this already challenging scenario with an unusual low beacon frequency of 1 Hz and a very high-density of over 200 vehicles on the intersection the adversary was still able to track over 96 % of all vehicles even when new pseudonyms were used for every sent message ($t_p = 0.5$ s). We found that the main reason for a failed tracking here was packet loss. This indicates the effectiveness and correctness of the deployed tracking algorithm, and also the importance of accounting for properties of the wireless channel. In terms of privacy these results can be considered worrisome since they show that when an adversary is able to overhear messages it is nearly impossible to avoid being tracked. This is an important finding, confirming earlier results obtained with sparser and less realistic traffic [255].

The radio blind spot in the freeway scenario (Figure 3.9) had a considerable impact on the tracking probability. When vehicles did not change pseudonyms the adversary was unsurprisingly able to track every single vehicle without exception. However, with increasing vehicle density and lower pseudonym validity times ($t_p =$ 50 s) the adversary could be confused by up to 50 %. It has to be noted that traffic generated by SUMO seemed to be more dynamic than on actual freeways (frequent lane changes and overtaking, trucks not consistently using the rightmost lane). We therefore expect that a larger gap is required in a real-world scenario to confuse the adversary in equal measure.

We have illustrated that our our privacy framework provides the necessary features to investigate different aspects of privacy in the context of vehicular networks. We will further demonstrate these capabilities in Chapter 4, where we evaluate several protection mechanisms that we developed to help improve location privacy.

3.2 Simulation of IEEE WAVE

Most privacy protection mechanisms and also attack vectors to compromise drivers' privacy in vehicular networks rely on messages transmitted by vehicles. Simulating a pseudonym changing strategy that uses incoming packets as an input or a tracking algorithm that relies on overhearing beacon messages from vehicles without a detailed model of the communication stack could therefore lead to inaccurate and misleading results. In this section, which is based on our paper "On the Necessity of Accurate IEEE 802.11p Models for IVC Protocol Simulation" [65], we show that using models for communication technology other than the one that will be actually deployed later is insufficient when investigating applications, privacy aspects, or protocols.

For that, we implemented a complete model for IEEE 1609.4 and IEEE 802.11p. This is not only the basis for the research we present later in this thesis, it also allowed us to identify shortcomings in the current version of the communication standards and thereby contribute to improving IVC in general.

Research on Inter-Vehicle Communication (IVC) – starting in the early 2000s, long before first drafts of WAVE – commonly relied on network models of the different WLAN standards. When work on this thesis began, it was still common to use models for IEEE 802.11b to simulate vehicular networks due to the lack of models for IEEE 1609.4 and IEEE 802.11p.

To increase the meaningfulness of these simulations and their applicability to study applications and protocols of future ITS's, other researchers adapted these IEEE 802.11b models to operate in the 5.8 GHz band – examples include [114,154,250]. In their 2012 paper "Comparing Apples and Oranges? Trends in IVC Simulations"‘ [125] Joerer et al. show that the majority of IVC research does not use IEEE 802.11p models for IVC simulation. Even though, intuitively, the use of the different protocols will lead to different network behavior, to the best of our knowledge, there has been no qualitative and quantitative evaluation or comparison of these models in the context of IVC simulation.

As shown in Section 2.1.1, IEEE 802.11p can be seen as the lower part of the MAC, while the upper part consists of two EDCA subsystems for the multi-channel operation defined in IEEE 1609.4. This upper part has a significant influence on when and which channel is accessed and when and in which order packets are transmitted. It is therefore crucial to have detailed simulation models, in particular when simulating more than one application or multi-channel scenarios in general.

In this section, we explain in which scenarios the use of the correct network model has only a negligible effect and show where WAVE-enabled simulation produces considerably different results compared to traditional WLAN models. Furthermore, we demonstrate to what extent it is possible to change parameters of an existing

WLAN model to match those of WAVE in order to produce more realistic results. In summary, it can be said that we give an answer to the question: “When is it okay to simulate vehicular networks with a different model than WAVE?”.

We were not the first to implement detailed models for the IEEE WAVE family of standards. In fact, the need for these models was understood from the very beginning of IEEE WAVE simulation. For example, Wang and Lin presented a fully functional model of WAVE for the NCTUns simulator [248]. Gukhool and Cherkaoui developed a similar model for the ns network simulator and give detailed insight on the challenges of creating such a model. They furthermore compared packet loss ratios of IEEE 802.11a and IEEE 802.11p on different vehicle speeds [105]. Similarly, we implemented a WAVE model for the OMNeT++-based Veins simulator to be the basis for our IVC simulations.

Publications related or similar to the simulation study in this section include the work of Wang et al., who evaluated the performance of the IEEE 802.11p backoff scheme and show that under certain circumstances, such as highly dynamic vehicular communication environments, the backoff mechanism does not work optimally [249]. This is something we also encountered when analyzing the performance of the WAVE protocol stack in dense scenarios. Dhoutaut et al. investigate the impact of radio propagation models on ad-hoc network simulations [44] and find that packet loss in vehicular environments is prone to occur in bursts. Our simulations confirm their results. In [71], Eichler presents extensive studies on the performance of IEEE WAVE in vehicular networks and shows that in high load scenarios data throughput decreases while the message delay slightly increases. Among others [29,65,238,249], Chen et al. confirmed these findings and gave additional insights on the reception probabilities depending on the physical distance between nodes [32]. While their configuration of the physical layer (transmit power, sensitivity, fading) was simplified and not based on real measurements or hardware, they highlight important issues in the IEEE 1609.4 standard. Again, our models show the same behavior, giving a strong indication of the correctness of our implementation.

3.2.1 Model Implementation

We implemented the model for our Veins simulation framework [219] for the well-established OMNeT++ network simulator [244] on top of the MiXiM model library [137].

The MAC is implemented as a simple module with connections to upper and lower layers. It includes both EDCA subsystems for the CCH and SCHs, and supports alternating access (see Section 2.1.1). The events the MAC has to react to can be categorized into three different types:

1. Events triggered by upper layers, e.g., a packet handed down from the network layer or an application.
2. The PHY notifies the MAC about certain events, such as the change of the channel state and the sending or receiving of packets.
3. A MAC internal timer triggered, e.g., when a packet can be sent.

Performance was of particular importance when developing the IEEE WAVE MAC, as a computationally costly MAC model would considerably slow the already extensive simulation of hundreds to thousands of vehicles. Because of this, contrary to all other MAC models implemented in INET or MiXiM, we only use one internal timer to manage all MAC operations. This timer will be set to the next possible event, e.g., the next transmission of a packet. Whenever a message from the PHY or upper layer arrives, this timer is canceled and rescheduled accounting for possible changes.

To illustrate the operation of this timer, assume the following example: The channel is busy and the application layer generated 4 different packets, 1 for each EDCA queue. The packets are queued according to their priority, but the internal timer will not be scheduled as the channel is busy, and the MAC has to wait until the PHY notifies it of an idle medium. Once this happens, the MAC will schedule the timer according to the EDCA queue that can send the earliest depending on AIFS's and current backoff values. When the channel turns busy before the timer expired, it has to be checked whether the channel was idle long enough for any queue to reduce the backoff counter, that is, if the medium was idle longer than the respective AIFS. When the timer triggers, the packet from the front of the winning queue is handed to the PHY for transmission. If necessary, backoff values are updated and internal contention is handled. When the PHY reports the successful transmission of the packet, it can be removed from the queue, and, should the channel be idle, the timer will be scheduled again. This method is considerably more efficient than maintaining multiple timers per queue to handle AIFS's and backoffs.

At the PHY, we needed to implement functionality for channel switching and were able to correct several issues in Veins (e.g., incorrect interference calculation and missing channel sensing after transmitted packets) in the process. Due to the changed timings and bandwidth of IEEE 802.11p, we were not able to use the implemented packet error model, which was geared towards IEEE 802.11b. Therefore, we derived a packet error model from the findings of Fuxjäger et al. [99], who provided accurate frame error ratios by developing a fully functional IEEE 802.11p software radio. Analogously to Equations 2.8 and 2.9 in Section 2.3, we compute the probability of a successfully transmitted packet with packet length l at a data rate of 18 Mbit/s with 16-QAM OFDM using the following equation:

$$p_{\text{succ}} = 1 - 1.5 \operatorname{erfc}\left(0.45 \sqrt{\text{SNIR}_{\text{min}}}\right)^{l} \tag{3.17}$$

The implemented single-radio model supports alternating access and multidimensional (time, frequency, space) interference computation as provided by MiXiM [137]. In order to cover the whole WAVE protocol stack, we furthermore developed a basic application layer on top of the MAC that is able to send messages following the WSMP according to the standard. In this simulation we use the free space model for line-of-sight connections and the obstacle model proposed in [218] (see Equation 2.6, Section 2.3) when the path is interrupted by a building. Parameters for the obstacle model as well as the maximum transmission range, transmission power, and minimum receiving power were taken from real-world experiments.

Parameter	Value
Models	11b, 11b5, 11p, 11pDC
Number of applications	2
Beacon AC	AC_VO
Beacon interval	10 Hz
Transmission range	Path loss ($\approx$ 1400 m)
Scenarios	Grid, suburban (Ingolstadt), freeway
Traffic density	Low, medium, high
Metrics	Neighbor count & lifetime, channel load, received packets

Table 3.3 – Simulation parameters used for the comparison of IEEE 802.11 models.

3.2.2 Evaluation Method

Even though IEEE 802.11a is more similar to IEEE 802.11p in terms of timings and frequency, we chose to compare IEEE 802.11b as both INET and MiXiM did not have an IEEE 802.11a model at the time of writing. It was also shown that IEEE 802.11b models, both standard and adjusted for IEEE 802.11p parameters, are used considerably more often than their IEEE 802.11a counterparts, amounting to approx. 40 % in vehicular network simulation in the year 2011 [125].

We chose to compare three models: standard IEEE 802.11b referred to as *11b*, an adapted version of IEEE 802.11b operating in the 5.8 GHz band with changed radio sensitivity, transmit power, etc. referred to as *11b5*, and our WAVE implementation called *11p*. For fairness reasons, we set the transmission range of all three models to be approximately the same.

Using our Veins framework, we examined three different scenarios with different movement patterns and traffic densities ranging from low (Manhattan grid [68]) to medium (urban scenario based on the city of Ingolstadt, Germany) to high traffic

density (freeway, two-lanes). For a description of these scenarios please refer to Section 3.1.3.

Vehicles emit beacon messages with a frequency of 10 Hz including a field indicating the type of application running on the vehicle. Each vehicle was assigned one out of two possible applications, both identical and working as illustrated in Figure 3.10. While beacon messages can possibly be received by all vehicles within the transmission range, only vehicles running the same type of application will actually respond with 60 packets of 1000 B. In the 11p model, application data was sent over a Service Channel (SCH) while beacon messages were only sent on the Control Channel (CCH). To illustrate the effect of multi-channel operation, we investigated a scenario where both applications ran on distinct SCHs, meaning that packets from different applications will not compete for the channel and also not interfere. We refer to this scenario as 11pDC for *distinct channels*.

Based on the metric proposed in [124], we use the *communication density* as a metric to measure the channel load for a given vehicle. It is simply the ratio of the amount of time the currently used communication channel was *busy* when the radio was in receive mode to the total lifetime of a vehicle. A communication density of 0.6 therefore means that for a given vehicle the medium was busy for 60 % of the time. Furthermore, we investigate neighbor count and lifetimes as these are an important input for many privacy protection mechanisms [67, 100, 198]. An overview of the used simulation parameters can be found in Table 3.3

3.2.3 Implications of Alternating Access

In terms of latency, it is clear that in a low channel load scenario the used models differ only marginally, due to slightly different inter-frame spacings and slot times (DIFS is 50 µs and slot length 20 µs in IEEE 802.11b compared to 13 µs slots and shorter AIFS depending on the access category). The results completely differ when applications with different priorities are used, due to effects such as prioritization and starvation introduced by the EDCA scheduling in the IEEE WAVE MAC as we showed in an earlier publication [62].

Due to the alternating access scheme, that is, periodic channel switching, beacon frequencies (assuming uniform inter-arrival times) higher than 10 Hz can be problematic in IEEE WAVE, as a beacon for the CCH will then inevitably be generated during an SCH interval causing a worst-case delay of 54 ms until it can potentially be sent. This effect is not captured using the 11b or 11b5 models. Note that this also holds for service channel messages generated during a control channel interval, meaning that IVC simulation of delay-sensitive ($\leq$ 60 ms) applications cannot rely on IEEE 802.11 simulation models other than IEEE WAVE.

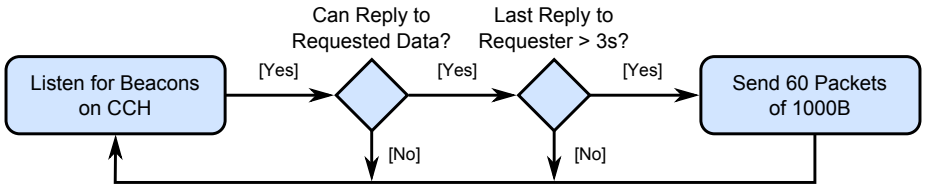


Figure 3.10 – Used application layer to generate network traffic.

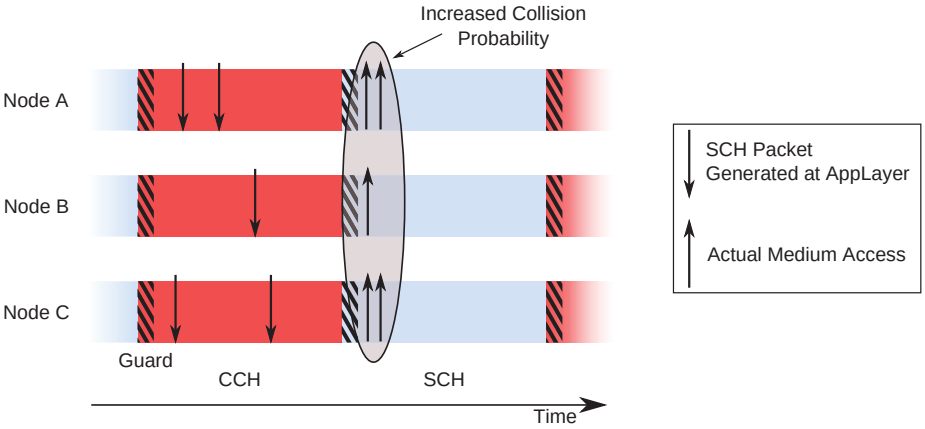
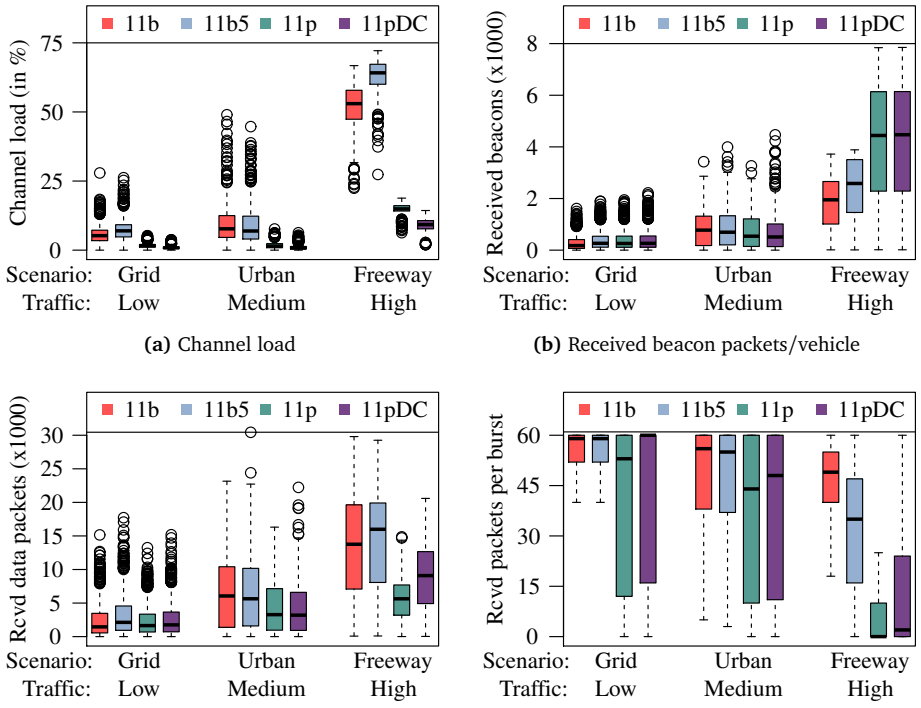


Figure 3.11 – Synchronization at the start of an SCH interval leads to possible packet collisions in IEEE WAVE.

We furthermore observed that this channel switching mechanism introduces synchronization effects as illustrated in Figure 3.11 and therefore leads to a high packet collision probability at the beginning of an interval. When an application generates an SCH packet during a CCH interval, this packet has to wait for the beginning of the next SCH period to be sent. The probability of a packet being sent at the beginning of an interval is therefore increased. This becomes particularly problematic in denser traffic when multiple vehicles have queued SCH transmissions during the CCH interval (or vice versa). Even with backoff values assigned to the higher priority queues, a packet collision is likely due to the low values for CW_{min} and CW_{max} in AC_VO or AC_VI. This also leads to an increased probability that packet losses occur in bursts, confirming the findings of Dhoutaut et al. [44]. This affects many of the metrics we investigate in our simulation study.

3.2.4 Channel Load and Packet Metrics

Figure 3.12 shows our first set of results in the form of box plots. The boxes reach from the first and third quartile, the whiskers extend to the furthest away data point which is no more than 1.5 inter-quartile range from the box. Outliers are visualized by circles. We examined the differences for the channel load (or communication density) for our different models, that is, the relative busy time of the wireless



(a) Channel load

(b) Received beacon packets/vehicle

(c) Received data packets/vehicle

(d) Received data packets per burst

Figure 3.12 – Comparison of network metrics in different simulation scenarios.

channel as observed by a vehicle. Figure 3.12a shows our findings in different traffic density scenarios. Even at lower density scenarios, but more prominently at higher traffic densities, the channel load of traditional WLAN-based models is considerably higher than for the 11p models. The higher the channel load at a node, the higher the probability a newly generated packet cannot be transmitted right away, but has to go into backoff. This increases delays and decreases throughput. Having an almost exclusive CCH for safety-related beacons such as CAMs or BSMs reduces the average channel load to avoid high latency and packet loss of important messages due to packet collisions with application packets. The benefit of the multi-channel operation becomes more apparent when the two applications in our scenario use different SCHs, as the load caused by application data packets is now distributed over two channels, meaning that vehicles running a particular application do not sense a busy channel when vehicles with another application exchange data.

As a second step, we measured the amount of beacons that were successfully received by a vehicle (Figure 3.12b). A successfully transmitted beacon message is the basis for the proper operation of many IVC applications, meaning that a notable difference here will (with high probability) affect the performance of the IVC application. This is particularly the case for safety and privacy-related applications

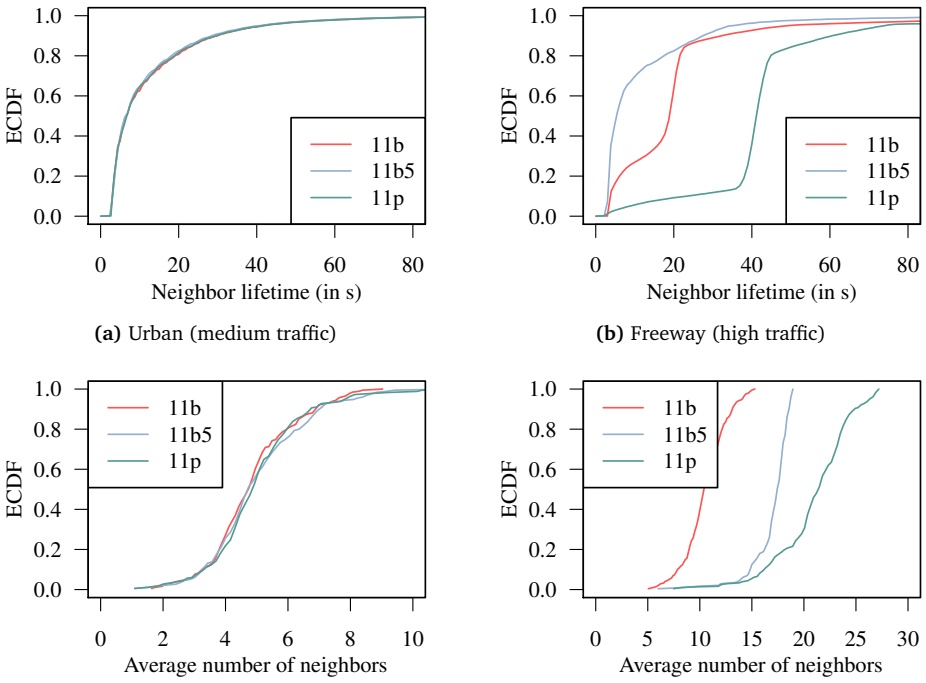
as received beacon messages are the most important input not only for collision avoidance but also for context-aware privacy mechanisms. We observe that for low and medium density scenarios, i.e., the ones with low communication density, the differences are rather small and mostly caused by the different packet error model (cf. Equation 3.17). However, when increasing the traffic density (resulting in higher channel load) the 11p model clearly outperforms the 11b-based models due to the fact that beacon packets do not compete with data packets in 11p while in 11b they can possibly interfere and collide with each other, resulting in a lower amount of received beacon messages.

Figure 3.12c shows the total amount of all received data packets per vehicle, Figure 3.12d shows the packets received per burst (60 packets of 296 B). Although the total number of packets received is similar for all models in low and medium traffic, we observe that the inter-quartile range for the 11p and 11pDC models is considerably bigger when examining the number of successfully received packets per burst. This is almost exclusively caused by the synchronization effects described above, leading to high packet loss probabilities at the beginning of an SCH interval. Due to the fact that in our simulations SCH packets are always generated in the CCH intervals, this effect becomes very noticeable. With increased traffic density, and thereby channel load, these synchronization effects lead to severe problems in the overall performance of IEEE WAVE. Furthermore, with a channel load higher than 50 % for the 11b-based models, it is obvious that the alternating access switching scheme, limiting the time to send data packets to 50 %, can no longer provide the necessary throughput needed by the implemented application. Using distinct channels helps, but cannot overcome the problems caused by packet collisions shortly after switching to an SCH.

3.2.5 Neighbor Metrics

Many ITS applications depend on the number of other vehicles in transmission range and on how long these connections last. In Figure 3.13 we plot the differences between the simulated models regarding the amount of neighbors per vehicle and the lifetime of such a neighborship in form of an ECDF. A vehicle is considered a neighbor of another vehicle after a beacon message was successfully received and will remain a neighbor until no beacon message was received for 3 s.

In sparse traffic (Figure 3.13a and 3.13c) we observe only marginal differences between the 11p and 11b-based models, caused by their very similar transmission range settings. Contention with application data did not play an important role for the 11b models in these scenarios, as the channel load was still low enough to allow undisturbed channel access for beacon messages. This changes in the high traffic density freeway scenario (Figure 3.13b and 3.13d), where it becomes clear that the



(a) Urban (medium traffic)

(b) Freeway (high traffic)

(c) Urban (medium traffic)

(d) Freeway (high traffic)

Figure 3.13 – Amount of neighbors and lifetime of a neighborship per vehicle.

distinct channel for periodic beacon messages in IEEE WAVE is advantageous for safety and cooperative awareness in general. Neighbor lifetimes and counts were considerably better using the 11p model.

Figure 3.13b shows that a large amount of neighborship relations lasted approximately 40 s. This is caused by oncoming traffic with which vehicles could communicate for about this duration when the channel was not congested, given the simulated transmission range and vehicle speeds. The distribution is heavy-tailed due to the low relative speeds of vehicles going in the same direction. Naturally, the amount of vehicles met moving in the same direction is considerably lower than the number of oncoming vehicles. In 11b and 11b5 counts and lifetimes are lower because beacon messages have to compete with application data for channel access, resulting in packet loss and a busier channel. The results also show that adjusting an 11b model to better match the properties of IEEE WAVE does not produce the desired effect.

3.2.6 Concluding Remarks

We showed several scenarios where the use of IEEE 802.11b-based models is not practicable as the differences compared to a full IEEE WAVE system are too big.

Studying applications sensitive to delays of 60 ms and lower requires a realistic model due to the alternating access scheme deployed in IEEE 1609.4. In general, the fact that application data and beacon messages do not share the same channel has many implications that need to be considered when choosing a simulation model. Examples include channel load, packet error rates, and latency.

In particular, the evaluation of safety applications and privacy properties requires realistic network models as they will affect both neighbor count and lifetimes. This is an important input for many safety-related features such as collision avoidance and intersection management, but also for context-aware pseudonym changing strategies that take surrounding vehicles into account to determine if and when a pseudonym should be changed. The used MAC and PHY models also affect what an adversary can possibly observe in the network, as lost beacon messages or high channel load will complicate the tracking of vehicles.

We also identified a major synchronization problem with the alternating access scheme that leads to packet loss. This should be addressed before roll-out as the consequences of many nodes simultaneously accessing the channel are undesirable. Our work contributed to increasing awareness of this issue, and at the time of writing, there are plans to also include other access schemes in IEEE WAVE to alleviate this problem.

3.3 Simulation of ETSI ITS-G5

Having established that a realistic model of the used communication stack is crucial for the meaningful simulation of many privacy protection mechanisms, we further wanted to investigate the actual differences between the European ETSI ITS-G5 system and its American IEEE WAVE counterpart. Although they both use an IEEE 802.11p MAC and PHY, their upper MAC shows decisive differences. Not only does ETSI ITS-G5 not make use of a channel switching scheme, it also introduces Decentralized Congestion Control (DCC) (see Section 2.1.2) which can have a strong influence on channel access, that is, when and how packets are sent. In this section, which is based on our paper "A Performance Study of Cooperative Awareness in ETSI ITS-G5 and IEEE WAVE" [61], we show that, despite their similarity, the performance of the two systems can greatly differ.

For this we implemented a full model of the lower ETSI ITS-G5 layers including a complete model of DCC for our Veins framework and compared it with our IEEE WAVE implementation from Section 3.2. We show that for some scenarios, particularly privacy and awareness problems in higher traffic density, the European system will behave considerably differently. This again emphasizes on the need to use the correct models to investigate IVC. Our results and findings contributed to the development of the upcoming ETSI ITS-G5 family of standards. In particular, we identified problematic behavior of the DCC state machine which should be addressed before the system is finalized.

At the time of writing there existed only little work on the performance of ETSI ITS-G5: Subramanian et al. compare ITS-G5 with WAVE and give valuable input on how to improve channel access in vehicular environments [229]. Our findings confirm some of their results but not all, caused by the fact that our parametrization of the physical channel is closer to real hardware when it comes to transmission ranges (and fading) or carrier sense thresholds. Also, the configuration of the DCC state machine seems to differ from the current values suggested by the standard [79]. Lastly, it is unclear whether realistic mobility models were used to simulate the performance of both MAC models.

Kloiber et al. showed that depending on the beacon frequency, the *update delay*, that is the delay between two decodable messages from the same sender, can exceed values beyond which safety applications can no longer function [132, 133]. While their study was based on the ETSI ITS-G5 MAC and the transmission of CAMs, they could not include the DCC state machine in their work. We use a derived metric of their update delay in our work in order to quantify the extent to which safety functions are influenced by packet loss.

First performance studies of other systems, such as the Japanese ITS operating at 715 MHz were presented by Sai et al. [197] in 2012, but with a particular focus on

the impact of obstacle shadowing compared to systems operating at 5.9 GHz (such as IEEE WAVE and ETSI ITS-G5). They show that the choice of frequency immensely influences the network topology and its dynamic.

Although various performance evaluations of IEEE 802.11p-based systems exist, we are not aware of a study of similar extensiveness. We put a particular focus on the comparison of the standardized mechanisms, using realistic models for mobility and the wireless channel. Our results show considerable differences in the reception probabilities of periodic beacon messages, potentially influencing both safety and privacy protection systems.

3.3.1 Evaluation Method

We described and compared both the American system and the European system in detail (Section 2.1.1 and Section 2.1.2). In summary, it can be said that the lack of alternating access in ETSI ITS-G5 alongside the deployment of the Decentralized Congestion Control algorithm accounts for the biggest distinction and can therefore be expected to cause most of the performance differences. The fact that ETSI ITS-G5 does not periodically switch channels but allows for multiple transceivers to send and receive on them simultaneously makes the system insusceptible to WAVE's packet loss problems discussed in Section 3.2.3. However, this also means that single-transmitter systems are unable to use SCH applications. Furthermore, the introduction of DCC is steered to counter channel congestion by means of TPC, TRC, TDC, DSC and TAC. In this section, we want to quantify these differences and discuss their influence on beacon-dependent safety and privacy applications. For this, we use metrics such as latency, packet delivery rates, the ratio of known neighbors, and also channel usage and conditions in general.

A complete overview of our simulation parameters can be found in Table 3.4. We have simulated every possible permutation of the listed parameters but – due to space constraints – we will only highlight the most significant simulation scenarios and findings. The hardware-relevant settings for this simulation study were chosen in close collaboration with the German car manufacturer Audi (a member of the ETSI working groups), and have also been confirmed by various real life experiments with IEEE 802.11p hardware. Settings for DCC, queue, and CAMs were taken from the corresponding ETSI standards [79, 82, 84] at the time of writing. Traffic densities were chosen to cover real traffic conditions from free-flowing to slow-moving traffic, but no gridlocks.

3.3.2 Channel Load Measurements

As a first step, we investigated how the channel load differs in IEEE WAVE and ETSI ITS-G5. Throughout all scenarios we observed that with high enough node

Parameter	Value
Scenarios	Freeway, motorway junction, traffic circle, Manhattan grid, suburban (Ingolstadt, Germany)
Path loss	Two-ray-interference [220] with 1.895 m antenna height, obstacle shadowing [218]
Traffic Density	Low, medium, high
Car-following model	IDM
Penetration rate	10%, 50%, 100%
Run-time	500 s - 1000 s
Transmission range	Path loss ($\approx$ 900 m)
Max. transmit power	26 dBm
EDCA queue length	2
Queue strategy	FIFO with tail drop
DCC measurement interval	1 s
NDL_minDccSampling	500 ms
Maximum CAM age	∞, 50 ms
CAM frequency	10 Hz
CAM AC	AC_VO
CAM size	210 Byte
Certificate size	125 Byte (20 % probability of being attached)

Table 3.4 – Simulation parameters used in our evaluation of ETSI ITS-G5.

density and penetration rates, the DCC state machine oscillated between its states. Figure 3.14 illustrates the effect as observed in the high-density (moving traffic, $\approx$ 170 vehicles/km, 100 % penetration rate) freeway scenario. As can be seen, the state machine continuously changes its states from *Relaxed* to *Active* to *Restrictive* and back, the switching intervals approaching the minimum delay necessary for a state transition, i.e., 1 s and 5 s, respectively. (cf. Figure 2.7, page 16). These transitions instantly affect the channel load causing the system to go into a loop as long as the observed channel load repeatedly exceeds the channel load threshold.

The explanation for this is straight-forward: In the *Relaxed* state the node is allowed to access the channel each time a CAM is created because the minimum packet interval is smaller than the CAM generation interval. It will attempt to transmit the packet with a high transmission power, and so will all vehicles in the vicinity which are also in the *Relaxed* state. This will increase the channel load

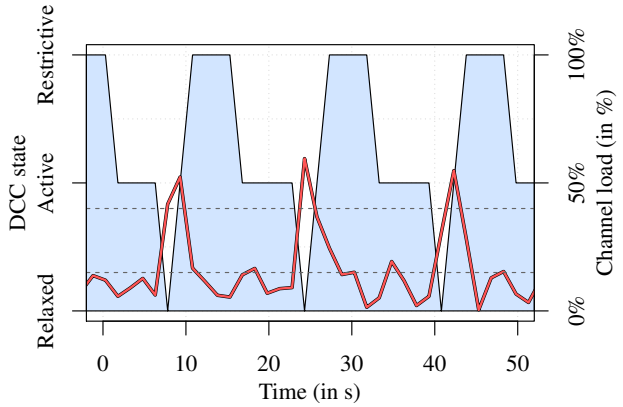
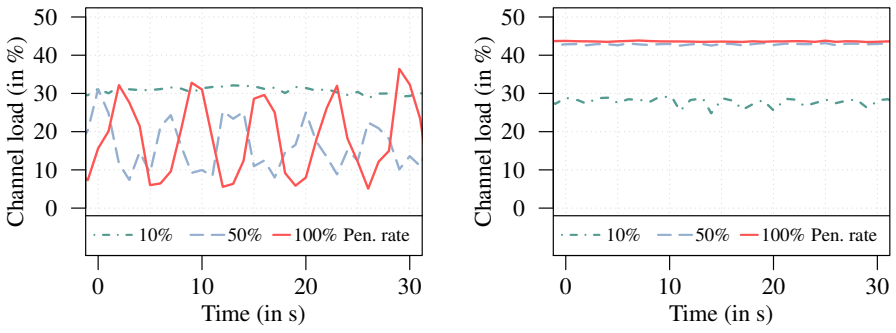


Figure 3.14 – DCC state (black line, blue area) and channel load (red line) of one vehicle on a busy freeway with 100 % penetration rate.

and therefore trigger transitions in the state machine. With a high enough channel load it is possible to reach the *Restrictive* state within 2 s, triggering change of all MAC parameters. Transmit power and, more importantly, the frequency with which packets will be transmitted are set to a minimum, causing nodes to hardly occupy the channel anymore. This in turn will cause the channel load to drop almost instantly (red line in Figure 3.14) and thereby put the state machine back in the *Active* state. Due to the hysteresis of the DCC parametrization, not all MAC settings are changed during this transition. The channel load may therefore drop below the *Relaxed* state threshold (lower dotted line) allowing for another state transition after 5 s. After this happened all MAC parameters will be set to maximum again and the channel load will increase resulting in a loop of DCC state transitions and oscillating channel load.



(a) ETSI ITS-G5 **(b)** IEEE WAVE

Figure 3.15 – Observed average channel load for all vehicles in the motorway junction scenario at medium traffic density with different penetration rates.

When analyzing the overall channel load for all vehicles, we observed that this oscillation does not only take place in a local context but also on a global scale (see Figure 3.15a). Within a cluster of connected vehicles, nodes tend to synchronize their DCC state transitions, causing the globally observed channel load to periodically increase and decrease. As for the local oscillation, the reason for this lies within the parametrization of the different DCC states: we find the *Restrictive* state to have a large influence on the medium access of a node, reducing the number of possible packets to one per second and also changing the transmit power to a value of −10 dBm. This leads to the fact that vehicles in the *Restrictive* state hardly try to access the channel anymore at all, possibly rendering them temporarily invisible to other vehicles. When nodes change their DCC state back to *Relaxed* they will attempt to access the channel almost synchronously. This results in a high packet collision probability, similar to that observed in IEEE WAVE after switching channels (Section 3.2.3). Both effects, that is, the overly restrictive channel access and packet collisions, can be critical issues for safety applications, as IVC-based collision avoidance systems will not work when neighboring vehicles do not send messages. Furthermore, this affects context-aware privacy protection mechanisms due to a vehicle's reduced overall awareness of its surroundings. This is especially problematic as high traffic density scenarios are a potential situation to confuse an adversary.

As a next step, we compared the channel load measurements for both systems to better understand how the mechanisms in the MAC affect the channel conditions. We examined the performance at different penetration rates, that is, the percentage of vehicles equipped with an OBU, in order to vary the channel usage without changing the mobility of vehicles. Figure 3.15 shows our findings for the 3-lane motorway junction scenario with a medium vehicle density (≈ 50 vehicles/km^2).

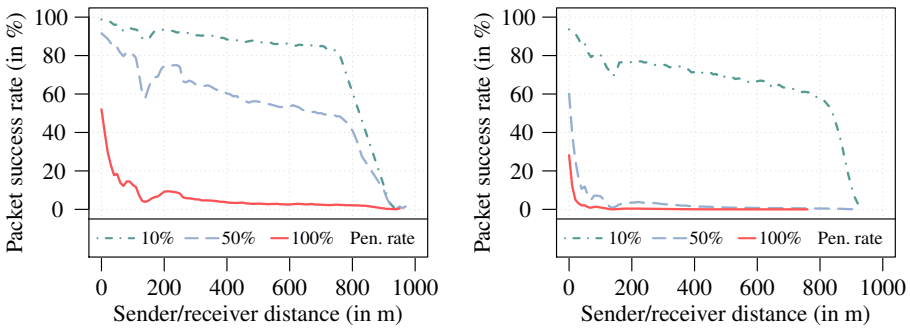
Naturally, the channel load for the WAVE system does not exceed 46 %, caused by radios being tuned to an SCH 50 % of the time and spending an additional 4 % in the CCH guard intervals. However, the remaining channel capacity is almost fully utilized with a high enough node density and remains at a steady level. Observed channel busy times for the ETSI ITS-G5 system show a substantially different behavior (Figure 3.15a). While at a low penetration rate the curve is almost a straight line at about 25 %, the channel load increasingly oscillated with higher penetration rates due to the reasons mentioned above. Although the full channel capacity is available, DCC does not efficiently utilize the available bandwidth in higher penetration rate scenarios. The average channel load observed at high penetration rates was lower than when only 10 % of all vehicles were equipped with On-Board Units.

Of course, this would not pose a problem if it did not affect packet delivery rates and thereby cooperative awareness. However, as we will show in the next section, it potentially does so and is therefore an indicator that DCC has room for improvement in its current version.

3.3.3 Packet Delivery Rates

To study the impacts of DCC and alternating access, respectively, we investigated different metrics based on the number of received packets. In particular, we show their effect on cooperative awareness and the ratio of known neighbors of a vehicle. We show results for the motorway junction scenario with high traffic volume (115 vehicles/km^2) as the observed effects were most prominent in this setting. However, the same effects discussed in this section could be observed whenever a high enough channel load was reached. The road topology and traffic volume influenced whether this was the case, as, for example, in the Manhattan grid scenario with obstacle shadowing enabled, the overall channel load remained on a low level without considerable packet loss.

We plot all these metrics against the sender/receiver distance so that the combined effect of path loss and channel congestion can be evaluated. While it may not be as problematic to not receive location updates from vehicles hundreds of meters away, it is potentially critical not to miss CAMs from nearby vehicles. Safety applications solely based on CAMs then rely on extrapolation, which is error-prone and can give false results when vehicles suddenly change direction or change lanes – the very cases where safety applications are needed to warn and assist drivers.

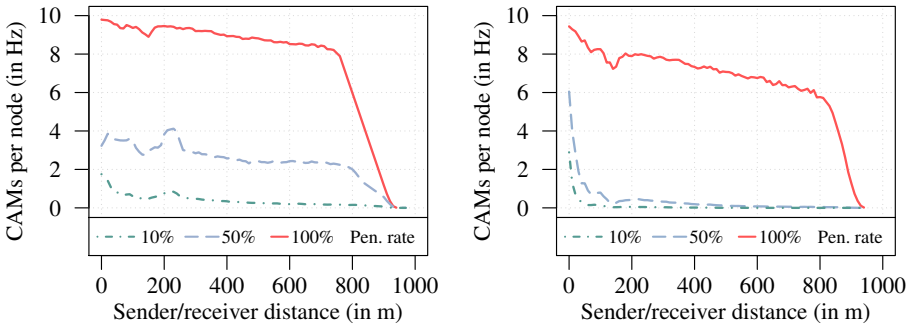


(a) ETSI ITS-G5 **(b)** IEEE WAVE

Figure 3.16 – Packet success rates in the motorway junction scenario, high traffic density.

In Figures 3.16a and 3.16b we compare our findings for both ETSI ITS-G5 and IEEE WAVE. We marked a packet as lost when it was not detected by the PHY or decoded unsuccessfully but could have been successfully received given best-case channel conditions, that is, the SINR would have been high enough without obstacle shadowing or interference from other packets. The decline around the value of 160 m, which can be observed in all plots, is caused by two-ray-interference path loss as a direct result of cancellation (Section 2.3.2). We observe that ETSI ITS-G5 outperformed the IEEE WAVE system at all penetration rates. The busy channel in

IEEE WAVE and the resulting interference resulted in considerable packet loss. Even in the 50 % penetration rate scenario the number of successfully received packets was at an alarmingly low level caused by the oscillation and the synchronous channel access caused by DCC. For ETSI ITS-G5 the packet success ratio was quite high in the lower penetration scenarios but dropped notably when 100 % of all vehicles were equipped with an OBU.



(a) ETSI ITS-G5 **(b)** IEEE WAVE

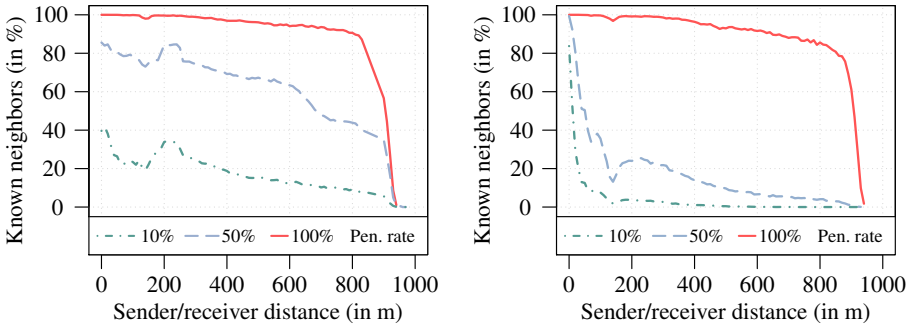
Figure 3.17 – Update frequency (= CAMs received from a node per second) in the motorway junction scenario, high traffic density.

In a next step, we evaluated how many CAMs (or BSMs) a vehicle received from other nodes. This can be seen as the update frequency. Ideally, a vehicle would receive 10 messages per second from each node in its vicinity, as we set the beacon frequency to a constant 10 Hz.

In Figure 3.17a it can be seen that, while ETSI ITS-G5 performs well at a low penetration rate, the number of CAMs received by other nodes per second drops below 50 % for the medium penetration rate. This results in an average update delay, that is, the latency between two consecutive CAMs from one node, of 250 ms and a worst-case update delay of over 1 s. A penetration rate of 100 % amplifies this problem as the channel becomes more and more congested, forcing nodes to go into the *Restrictive* state more often, reducing the effective CAM frequency to 1 Hz due to the minimum allowed interval between packets.

When comparing this to the performance of the WAVE system (Figure 3.17b) we observe a deterioration in performance caused by halving the time vehicles are allowed to send beacons. Already at low penetration rates we observe packet loss, and thus an increasing update delay between vehicles. When we increased the number of sending vehicles (50 % and 100 % penetration rates) the channel became fully congested and transmitting to nodes further away than 100 m was almost impossible. However, at very low distances, more CAMs per neighbor could

be received, mainly caused by the fact that vehicles did not decrease the sending frequency by increasing the minimum packet delay.



(a) ETSI ITS-G5 **(b)** IEEE WAVE

Figure 3.18 – Ratio of known neighbors, i.e., the percentage of vehicles of which a car was aware depending on the distance (motorway junction, high traffic density).

The update frequency is an important metric for many safety applications. However, showing the average values for all vehicles in the vicinity can give a false sense of awareness. We therefore show the ratio of vehicles actually known to another vehicle, that is, vehicles from which it has received at least 1 message in the last second. This metric is important for many PETs as they make use of neighboring vehicles to change pseudonyms [67, 100] or create cryptographic groups [198]. In Figure 3.18a and Figure 3.18b we show our results in a high channel load scenario. We observe that ETSI ITS-G5 performs better than WAVE at higher penetration rates but still has problems even at low distances with only some 40 % of vehicles being visible to the radio receiver.

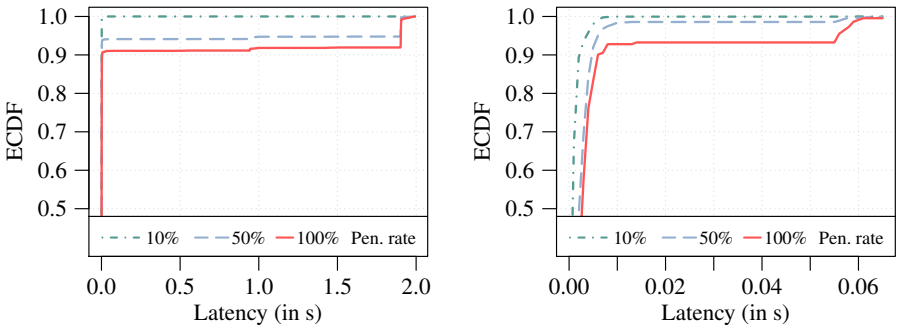
Interestingly, the reason for the low amount of visible neighbors (and received CAMs) is not the same for ETSI ITS-G5 and IEEE WAVE. While Figure 3.16a clearly shows that in IEEE WAVE these effects are caused by packet loss (compare to Figure 3.17b), the ETSI ITS-G5 system still has a high ratio of decodable packets at the medium penetration rate. From this it follows that *if* nodes sent packets, there was a high probability that they could be decoded by the receiver. However, the standard DCC parameters seem overly conservative, forcing nodes to considerably reduce their sending frequency (by increasing the minimum packet interval) even though the wireless channel may still have sufficient capacity.

3.3.4 End-to-End Latency

Safety applications depend on the up-to-dateness of received data in order to function in a reliable and robust way. Even though we believe this metric does not directly

have a significant influence on PETs, we use our simulation model to investigate this important property to contribute to improving future ITS's.

We compared both systems in terms of end-to-end latency, that is the delay between creation of a CAM at the sender and successful decoding at the MAC of the receiver. This delay includes the time the packet spent in the MAC of the sender and the more or less negligible airtime. We did not consider any additional processing time at the receiver end.



(a) ETSI ITS-G5 **(b)** IEEE WAVE

Figure 3.19 – End-to-end delay, measured from the creation of the CAM (or BSM) until the successful reception in the traffic circle scenario (note the different x-axis scale).

We plot the ECDF for the measured latency in the traffic circle at high traffic density (≈ 100 vehicles) and observe that in ETSI ITS-G5 almost 10 % of all packets received in the high penetration rate scenario are older than 1.9 s, a value which can render safety applications ineffective [259]. We chose the traffic circle to illustrate that these effects already occur in simple city scenarios; latencies on clogged freeways were even higher and were also observable in the Manhattan grid and city (Ingolstadt, Germany) scenarios. The main reason for this is the tail drop strategy at the FIFO (First In, First Out) transmit queue: When a CAM waits for transmission in the MAC queue and the DCC state machine is within the *Restrictive* state, the packet may not be sent for another 1 s. The next CAM generated 100 ms after will wait behind in the queue and cannot be sent until 1.9 s after generation, causing the noticeable increase in the ECDF. Assuming a MAC queue size of 2, each CAM after will have to wait for the same duration. Once the DCC state machine enters the *Restrictive* state it will keep a minimum packet interval of 1 s for at least 10 s until it can enter the *Relaxed* state again, which significantly reduces the accuracy and up-to-dateness of the information in a vehicle's CAM. Different policy and scheduling strategies, and also larger queue sizes can be expected to have a decisive impact on the observed end-to-end delay.

While packet loss was much higher in the WAVE system it can be said that successfully transmitted data was substantially more up-to-date (Figure 3.19b). The highest latencies we observed were around 60 ms and thereby still in a range useful to safety applications [259]. These latencies result from a congested wireless channel, sometimes forcing nodes to go into backoff right before the end of a CCH interval. The packet then has to wait at least 54 ms (= the time until the next CCH interval) for a new transmission attempt.

3.3.5 Concluding Remarks

Throughout all scenarios we observed that the percentage of known neighbors and the amount of received beacon messages drops considerably when the penetration rate increases. In IEEE WAVE this is caused by collisions on the channel, while in ETSI ITS-G5 the reason for this is the strict restrictions placed by the state machine, hindering vehicles from trying to access the wireless channel and in some cases also increasing the end-to-end delay. This is problematic as it will give false or inaccurate input to privacy and safety applications. The differences also indicate that a privacy mechanism working well in one system might perform differently in the other. It is therefore highly recommended to use a full-featured model of the envisioned system.

In terms of fairness it can be said that we did not observe that in either system particular nodes used the channel substantially more often than others. In ETSI ITS-G5, when the channel load was high, e.g., in the freeway scenario, most of the nodes dropped about 60 % of all packets within the MAC due to a full transmit queue. This and the fact that packets possibly have to wait for a long time within the MAC makes it clear that the generation of safety messages has to be either aware of MAC and channel conditions, or that policing mechanisms within the MAC have to be tailored for safety messages.

We observed that, in terms of channel conditions, the road topology did not have a substantial effect, it merely affected the number of transmitting vehicles necessary to lead to the observed problems. Throughout all scenarios, even in the Manhattan grid and suburban settings, we discovered that – at higher penetration rates – realistic, common traffic densities were sufficient to cause critical performance issues for both IEEE WAVE and ETSI ITS-G5 when the channel became too congested. Interestingly, signal shadowing caused by obstacles had a positive effect in terms of channel load for both systems as it reduced the number of potential receivers and therefore the interference from nodes located further away.

The employed Decentralized Congestion Control (DCC) mechanism showed improvements at lower penetration rates, but could not ensure proper functioning of safety, privacy, or other applications when nodes were forced to go into the *Restrictive* state. One of the reasons for this is that the difference between the

parameters defined for the different states (packet interval, transmission power, etc.) is large while the amount of states is rather limited. We expect that fine tuning of the parameters or an approach as suggested for the SCHs with its multiple sub-states for the *Active* state can help improve the situation on the CCH. In general, we believe that the smoother the transitions of the state machine are, the better it could actively control the congestion on the channel. A possible direction for future systems would therefore be to transform the state machine into a steady function that – dependent on the observed channel load – parametrizes the MAC accordingly. The possibility to control packet interval with steady functions has already been shown [9, 221] and is currently discussed as an option for IEEE WAVE.

Chapter 4

Privacy-Enhancing Technologies

In this chapter we present several Privacy-Enhancing Technologies (PETs) that we developed to help protect drivers' privacy in future ITS's.

First, we introduce the concept of time-slotted pseudonym pools and discuss how these can be used to prevent tracking even from the system provider or certificate authority (Section 4.1). We then identify privacy issues related to certificate revocation and present a method to preserve backward privacy when revoking the pseudonym pool of a vehicle (Section 4.2). Our method is extremely efficient in terms of network overhead and can be distributed in an epidemic fashion using parked vehicles. To motivate the use of parked vehicles we also show that they can substantially contribute to improving traffic safety by relaying messages from moving vehicles. Lastly, we show a robust and effective fingerprinting technique using scrambler codes and show its implications by means of simulation (Section 4.3). We found that IEEE 802.11p prototype OBUs do not follow the standard correctly and thereby possibly compromise any PET by introducing identifying properties into each sent message.

Parts of this chapter are based on our articles published in *IEEE Transactions on Mobile Computing* [217], *IEEE Communications Magazine* [67], and on papers published at conferences and workshops [21, 56, 66, 68, 218].

4.1 Time-Slotted Pools and Pseudonym Exchange

When designing privacy schemes for vehicular networks, in particular pseudonym changing strategies, important domain-specific constraints have to be kept in mind. Many of the proposed privacy protection mechanisms (including the ones discussed in Section 2.2.3) need a large pool of pseudonyms, so that if the CA is not reachable due to lack of connectivity or a car was not used for a longer time period, the vehicle can still send messages until the CA supplies new pseudonyms. A larger number of pseudonyms stored on each vehicle can therefore decrease the possibility of a car not being able to transmit messages, but the required disk space, transfer volume, and management costs will also increase. As the network grows there will be a considerable computational and network overhead at the CA just to keep all nodes equipped with a sufficient number of pseudonyms.

Second, PETs in vehicular networks usually do not consider the system provider or certificate authority a possible adversary that tries to disclose users' locations. This is worrisome because the CA, which signs the pseudonym certificates, is able to resolve every pseudonym to the static base identity of a vehicle and can hence track every vehicle as long as it is able to overhear beacon messages. The wide area deployment of provider-operated RSUs is a prerequisite for successful operation of many ITS services at low penetration rates, especially in the roll-out phase, where only few vehicles are equipped with wireless devices. In a worst-case scenario, these RSUs could be exploited by the provider to track entities throughout the network, therefore posing a serious threat to the location privacy of participating drivers. This allows an operator (or others setting up access points) to create accurate traces of all participants if the number of observations is high enough [255].

In this section, which is based on our articles "SlotSwap: Strong and Affordable Location Privacy in Intelligent Transportation Systems" [67] published in the *IEEE Communications Magazine* and our conference paper "Strong and Affordable Location Privacy in VANETs: Identity Diffusion Using Time-Slots and Swapping" [66], we present a solution for these problems by introducing reusable time-slotted pseudonym pools and pseudonym exchange amongst vehicles. We will refer to our approach as *SlotSwap*.

A possibility to address the problem of a tracking system provider is to introduce a separation of the CA into a *Privacy Authority* and an *Identity Authority*, both sharing just parts of the identities, hence requiring their collaboration to resolve an identity [94]. This idea of *separation of concerns* is by design susceptible to abuse as it is very hard or even impossible for users to check whether these policies are actually enforced, especially considering current revelations on governmental intelligence programs. In the context of MANETs, the problem of a "Big Brother CA" has been dealt with via the exchange of identifiers between mobile nodes [151]. We adopt

and extend this idea, apply it to vehicular networks, and make it a part of the privacy approach we present in this section.

In general, to overcome the problem of linkability between two messages, a straightforward approach would be to use a different pseudonym for every message to preserve a very high level of privacy, similar to not using source addresses at all. However, many (safety) applications *have* to link two or more successive beacon messages to one vehicle in order to function properly [204]. Simply changing the pseudonym every n seconds has been shown to offer only marginal protection of users' location privacy [198]. Context-aware PETs, that is, strategies that take into account speed, heading, and the number of cars in transmission range, may offer a much higher level of location privacy, as they reduce the chances for an attacker to successfully follow a pseudonym change [100, 151]. Further adding silent periods after a pseudonym change is a promising approach to improve drivers' location privacy, however, this is particularly problematic when performed at high-density spots (mostly found at intersections or traffic lights), as it makes a vehicle invisible to others and causes safety applications to fail [148]. Unfortunately, these are the very situations where IVC-based safety applications are needed the most.

Our approach addresses linkability by both providers and third parties by exchanging pseudonyms between vehicles based on the current situation of a vehicle. The introduction of time-slotted pseudonym pools also substantially reduces network and computational load for the operator, and introduces static upper bounds for disk space usage and communication overhead between vehicles and the CA. When reuse of pseudonyms is implemented, it furthermore makes it impossible for a vehicle to not have a usable pseudonym after it has been equipped with a pseudonym pool once.

4.1.1 System Design

Instead of storing a large amount of pseudonyms, every node maintains a time-slotted pseudonym pool with slot length t. For each time slot i, there exists exactly one assigned pseudonym P_i. The total period length p divided into time slots of length t results in $\frac{p}{t}$ pseudonyms per car with only one valid pseudonym at any given point in time. When a time slot has passed, each node will change its pseudonym simultaneously. This can be achieved by synchronized clocks using the GPS signal.

While the use of non-overlapping pseudonyms, as also proposed in [188], is very similar to time slots, nodes in SlotSwap will reuse pseudonyms. When the $\frac{p}{t}$th time slot has passed, time slot 1 will become active again, meaning that the time period will simply restart from the beginning. This would introduce a privacy problem because it allows for re-identification of vehicles using the same pseudonym again,

however, exchanging pseudonyms beforehand protects against this attack vector, as we will show in this section.

An example choice for those values, t = ten minutes and p = one week, results in a pseudonym being valid for, e.g., Monday from 6:00 a.m. till 6:10 a.m. Note that this pseudonym is then, in fact, valid on *every* Monday for said ten minutes. It can be seen that the only parameter for time-slotted pools which has a direct influence on location privacy during a trip is the time slot length t, which determines how often a node changes its pseudonym.

It has been shown that the exchange of pseudonyms between nodes can increase privacy in mobile networks and complicate tracking for an adversary [151]. If nodes are able to exchange their pseudonyms in secrecy by using encryption and keep third parties from tracking which nodes have swapped pseudonyms, a possible mapping at an authority will also become invalid. One possible method to achieve efficient confidentiality would be to encrypt a symmetric session key using the public key of the candidate. (The design of a secure exchange protocol is not within the scope of this work as methods and protocols for confidential communication are part of the upcoming standards.) Due to the time-slotted pseudonym scheme, only pseudonyms valid for a specific time slot can be exchanged, otherwise it cannot be guaranteed that every vehicle has exactly one pseudonym per time slot. This means that two vehicles must only exchange pseudonyms valid for the same time slot.

Swapping the currently used pseudonym with another node is not trivial, as the exchange partner has to be chosen carefully so that both vehicles can benefit from an exchange in terms of location privacy. For example, two oncoming vehicles will most likely not increase their anonymity by swapping pseudonyms because this action could be easily detected due to the unlikeliness of both cars having turned around at the same time. To effectively gain anonymity from a pseudonym exchange, nodes have to take context information into account [100]. This means that a node evaluates its environment and then decides if changing its pseudonym is profitable, so an adversary cannot simply infer the nodes' pseudonyms after the exchange by extrapolating their expected position based on their last known heading and speed [198].

In our approach, we use the speeds, headings, and positions of other vehicles to determine whether a node A will ask a node B in its vicinity to exchange the currently valid pseudonym. We will refer to all nodes meeting these requirements as *candidates*.

By carefully choosing bounds for similarity we increase the likelihood of both exchange partners being indistinguishable in terms of position. This creates confusion for an overhearing adversary, who then may not be sure whether a pseudonym exchange has taken place or not. The efficiency of this scheme, of course, is highly dependent on the frequency and positional accuracy of the beacons each car emits.

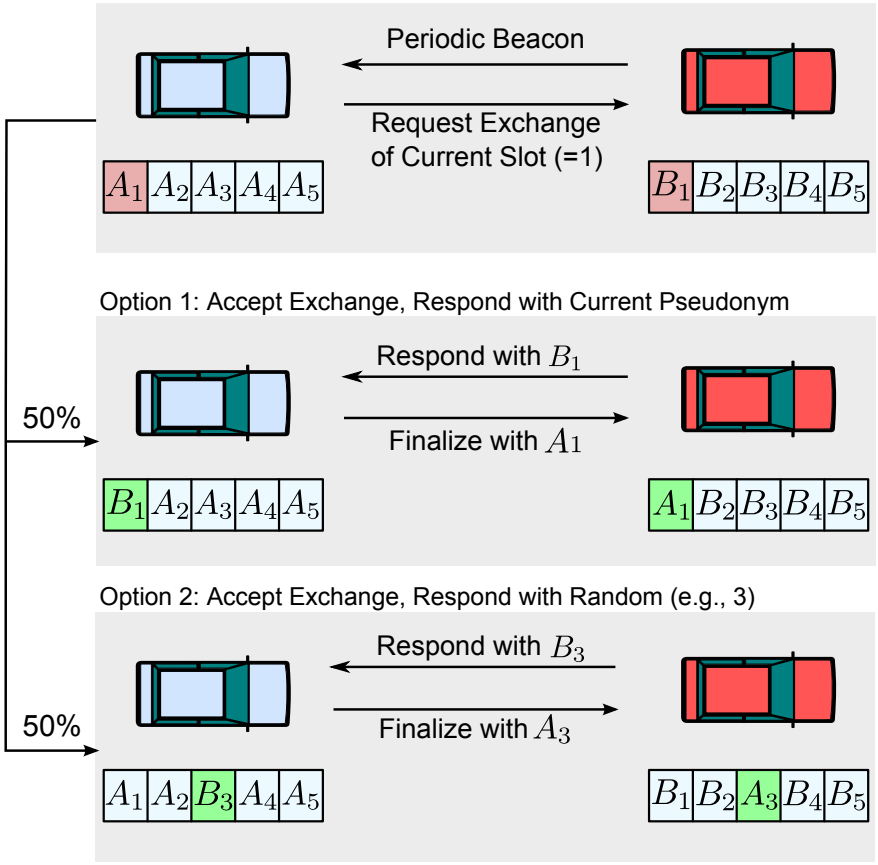


Figure 4.1 – Pseudonym exchange between two cars: The currently valid pseudonym is requested and confirmed in option 1 (resulting in the exchange of the current pseudonym), but rejected and replied with a random pseudonym from the pool in option 2.

The privacy achieved by this approach could thus be amplified by using further privacy-enhancing methods, such as random silent periods [118], where both cars will not send beacons for a certain amount of time after a possible exchange.

However, only exchanging the currently valid pseudonym is problematic: If no other pseudonyms are exchanged, each vehicle will start using the same pseudonym every $\frac{p}{t}$ slots again, because once a new slot has begun, the pseudonym last used in the previous slot would not be exchanged until this slot becomes active again. This way an attacker or the system provider is able to link two locations to one node: the present one (e.g., this Monday 6:00:00 a.m.) and the one from the last time the time slot was active (e.g., last Monday 6:09:59 a.m.). Furthermore, each time a car enters a time slot for the first time (which will happen $\frac{p}{t}$ times after being equipped with the OBU) the operator of the CA could link the first location in these time slots to a base identity and thereby potentially to an individual.

Therefore, cars have to be able to exchange these pseudonyms *before* actually using them. To achieve this, each time a time slot ends, the last used pseudonym is marked as *traceable*. In addition, all pseudonyms that are freshly obtained from the CA are marked *traceable*. When a node encounters another node, it decides to either exchange the current pseudonym (if the other node is a *candidate*), or one marked *traceable*, removing the flag if successful. Preferably the currently active pseudonym is exchanged, as it directly increases the level of location privacy for both users. However, for the exchange of other pseudonyms constraints like speed or heading can be neglected, due to the fact that an attacker is not able to decrypt the transmitted data and determine for which slot pseudonyms were exchanged.

Require: beacon message of v' received **AND** not flagged
  flag beacons of v' for 20 s
  if $v_{\text{speed}}, v_{\text{heading}} \approx v'_{\text{speed}}, v'_{\text{heading}}$ **AND** $swap = true$ **then**
    request exchange of current $P_n = P_{now}$
    $swap \leftarrow false$ for 60 s if succesful
  else if traceable pseudonyms in pool **then**
    request exchange of random traceable P_i
    $P_i \leftarrow$ non-traceable after successful exchange
  else
    request exchange of random $P_i \neq P_{now}$
  end if

Algorithm 4.1 – Pseudo-code for the exchange request for a pseudonym.

Require: v' requests exchange of P_i
  flag beacons of v' for 20 s
  if $P_i \neq P_{now}$ **then**
    exchange pseudonym P_i
    $P_i \leftarrow$ non-traceable
  else if rand() < 0.5 **AND** $swap = true$ **then**
    exchange pseudonym P_{now}
    $swap \leftarrow false$ for 60 s
  else
    exchange (preferably traceable) pseudonym P_j from pool
    $P_j \leftarrow$ non-traceable
  end if

Algorithm 4.2 – Pseudo-code for the response to pseudonym exchange requests.

To reiterate, it has been shown that too frequent pseudonym changes can have a negative impact on safety applications and on geographic routing in vehicular networks [204]. We therefore allow only one pseudonym exchange per car every 60 s. In addition, to avoid overloading the network, node *A* must only contact node *B* every 20 s.

Cooperative Awareness Messages (CAMs) as emitted by vehicles in an ITS will be broadcast unencrypted. Therefore, an overhearing adversary can conclude whether node *A* is a candidate for another node *B* and thus anticipate the exchange of the current pseudonym. To overcome this predictability, we introduce a 50 % probability to decide whether a node will send a positive response. This means that, if a node *A* asks for the exchange of the currently active pseudonym, node *B* will accept *or* reject the request. If the request is rejected, nodes *B* and *A* will exchange another pseudonym instead so that an attacker cannot determine if the nodes have swapped their current pseudonyms simply based on message sizes.

Figure 4.1 depicts possible flows of the pseudonym exchange process, a full description of the algorithm is given in Algorithm 4.1 and Algorithm 4.2. Vehicle *A* requests the encrypted exchange of the currently valid pseudonym from vehicle *B*, because both vehicles happen to have similar values for heading, speed, and position. In half of all cases node *B* will respond with its current pseudonym and *A* will finalize the exchange process by handing over its current pseudonym as well. The vehicles will then use the new pseudonyms. Alternatively, vehicle *B* will not exchange its current identifier but respond with another pseudonym from its pool, preferably one marked as traceable. Vehicle *A* will accept this, and answer with the corresponding pseudonym from its own pool. Both vehicles will replace their old pseudonym for the given slot with the one from the other node and continue using their current identifier.

4.1.2 Evaluation Method

We investigated the effectiveness of our scheme with the help of our Veins simulation framework. We implemented the presented protocol for pseudonym exchange.

We compared SlotSwap with a mechanism that uses random pseudonym changes with subsequent silent periods, as done in some field operational tests at the time of writing. A vehicle will randomly change its pseudonym and enter a random silent period of at most 10 s. The gain of location privacy is then dependent on nearby vehicles also being in such a random silent period.

Exchanging pseudonyms affects various privacy properties: Starting a new slot with an exchanged pseudonym affects anonymity, exchanging the currently used pseudonym tries to achieve unlinkability. Eliminating the mapping of pseudonyms

to base identities allows for plausible deniability. Lastly, the encrypted exchange of key material requires confidentiality.

Attacker Model

The evaluated level of location privacy enjoyed by an individual always depends on the power of an attacker trying to track a specific target in the network. In our simulations, we assume a global passive attacker, that is, an attacker that is able to overhear and decode *every* message sent by any vehicle. The attacker is further able to evaluate the content of all broadcast beacon messages (which we assume to include the speed, position, and heading of a node). As the attacker is well aware of the protocol they are able to conclude which nodes are candidates for another node and therefore might exchange their current pseudonyms. The attacker is, however, not able to actually follow the pseudonym exchange, as these messages are encrypted. What the attacker can gather from observing transmissions in the network is the fact that pseudonym requests and replies have been exchanged.

Our attacker model is based on the strong assumption that at the beginning of the lifetime of a node, the attacker can link an individual to the vehicle. If this was not the case, the individual would already be anonymous from the start and could only be exposed through origin/destination pairs if tracking throughout the network was successful.

When modeling an attacker using tracking algorithms, the strength of the attacker is heavily dependent on the used mobility and driver model. If, for example, nodes do not change lanes or drive in a very predictable manner, tracking algorithms will perform significantly better. As mobility in the Ingolstadt scenario is not based on real traffic demand patterns and the tracking module of the privacy simulation was still in development at the time of this simulation study, we chose to use a probabilistic attacker model.

As we have shown in Sections 2.4.2 and 3.1.2, the entropy is based on the probabilities p_i determining how likely a member i of the anonymity set is to be the target. The distribution of p_i is directly dependent on the strength of an attacker. Here, the attacker strength is defined as the probability with which an attacker is able to follow a pseudonym exchange between two nodes. The weakest possible attacker in our scenario would thus be an attacker which is completely confused by a pseudonym exchange. This means that from the adversary's perspective, an individual I, previously known to be the driver of A, is equally likely to be the driver of A or B after these vehicles have exchanged their current identifiers. The strongest possible attacker cannot be confused by pseudonym exchanges and is therefore able to track every entity throughout the network. The anonymity set for each individual would then only contain the individual itself and the entropy $\mathscr{H}$

would be zero. An attacker strength of n means that the attacker is able to track a pseudonym exchange with $n{\cdot}100\,\%$ certainty. Assuming vehicle A has not exchanged pseudonyms before, then the anonymity set of a node A exchanging pseudonyms with B is $A_A = \{nA, (1-n)B\}$. This probability also affects by how much the level of privacy is increased when a new slot in the pseudonym pool becomes active, that is, when all nodes will start using new pseudonyms. If we assume that two nodes very close to each other could confuse an attacker by exchanging their pseudonyms (the extent being dependent on its strength), this attacker will also be confused when these two nodes both switch to a new pseudonym simultaneously. From this we follow that the level of confusion is based on the amount of candidates directly neighboring a node. As stated, not all cars within transmission range are considered *candidates*, but only those with similar speed, heading, and position.

Simulation Setup

Parameter	Value
Adversary model	External/internal, global, passive, static, domain-specific
Privacy domain	Location privacy
Privacy properties	Anonymity, unlinkability, plausible deniability, confidentiality
Data source	Observable information, re-purposed data
Metrics	Entropy, pseudonym statistics
Scenario	Suburban (Ingolstadt, German), freeway
Technology	IEEE WAVE
Beacon frequency	0.25 Hz
No. of vehicles	25-600
Current pseudonym	max. 1 exchange per 60 s
Other pseudonyms	max. 1 per candidate per 20 s
Angle difference	max. 15°
Position difference	max. 30 m
Speed difference	max. 10 km/h

Table 4.1 – Simulation setup and parameters for the SlotSwap evaluation.

We evaluated SlotSwap in the realistic suburban scenario based on the city of Ingolstadt, Germany as well as in a synthetic four-lane freeway setup (cf. Section 3.1.3). Traffic was created by randomly generating origin/destination pairs and iteratively applying dynamic user assignment, as implemented in SUMO, until the algorithm reported a stable, optimal distribution of flows. In the evaluation, we focus on the

4 km^2 ROI, which contained a typical mix of high- and low-capacity roads, traffic lights, and unregulated intersections, as well as high- and low-density areas. To avoid border effects, traffic is simulated in the whole city of Ingolstadt, while the privacy scheme is only applied to nodes within the ROI. We simulated over 350 h of traffic with a total of over 1 500 000 cars.

The complete simulation setup is given in Table 4.1. The pseudonym pool length p is set to 1 week, the slot length to 10 min. Cars are considered to be eligible for exchange of the current pseudonym, or *candidates*, when their speed difference is at most 10 km/h, the difference in heading is at most 15°, and their distance is no greater than 30 m. The beacon frequency does not affect the achieved level of privacy in our simulation as we used a stochastic attacker model. Based on findings in [255] and our results for an urban scenario in Section 3.1.3, we will assume a strong attacker ($n = 0.95$) that follows pseudonym changes with a certainty of 95 %.

To calculate the communication overhead caused by SlotSwap, we base the amount of data needed for exchanging a pseudonym on the proposed algorithms and certificate lengths in the IEEE WAVE family of standards [122]. We assume a certificate length of 288 B with an asymmetric key length of 1024 bit and a symmetric key length of 128 bit for the *aes_128_ccm* scheme. From this, we conclude that the traffic needed for the exchange of a pseudonym, including IP overhead, is roughly 1 KiB, that is, 0.5 KiB per node. We neglect beacon messages in these calculations, since we consider them to be a prerequisite of ITS deployments in general, not of SlotSwap.

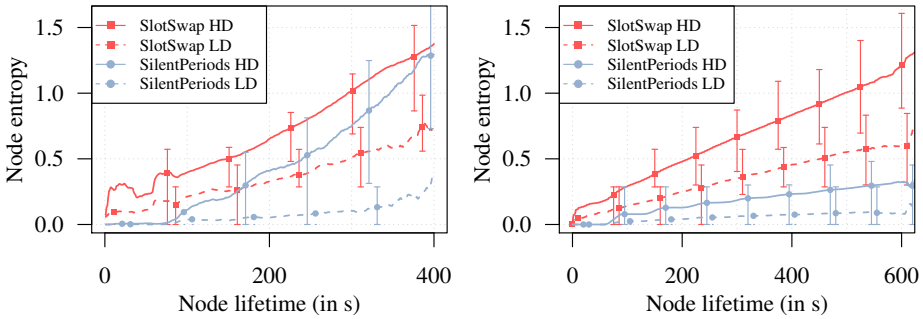
4.1.3 Results

We will now discuss the results of our simulation study both in terms of privacy and overhead caused by pseudonym swapping. To assess the privacy level of drivers we compute the entropy and also give information on the number of exchanged pseudonyms and possible candidate nodes.

Entropy Measurements

In a first set of simulations, we investigated how SlotSwap performs in the urban scenario. We observed nodes moving through the ROI and calculated the entropy resulting from pseudonym exchanges and slot changes. We measured the mean level of privacy in a low-density (LD, 16 cars/km^2) and a high-density (HD, 100 cars/km^2) scenario. The results are shown in Figure 4.2a.

The level of privacy achieved with SlotSwap was higher than with random silent periods, the difference being particularly notable at lower traffic densities. While the effectiveness of randomly changing pseudonyms depends on the coincidence of other vehicles being very close at the time of a pseudonym change, SlotSwap will



(a) Urban scenario **(b)** Freeway scenario

Figure 4.2 – Evaluation of the level of privacy as enjoyed by drivers in the ITS, measured by means of the entropy. Error bars show the second and third quartiles of the data set, lines show the average value for all nodes.

systematically utilize such a situation by exchanging pseudonyms with the nearby vehicle. As a result, drivers in SlotSwap will enjoy a higher level of privacy at the beginning of trips, while an initial delay is apparent before vehicles become anonymous with the random change approach.

As can be concluded from the second and third quartiles (illustrated by error bars), there are more vehicles with a considerably lower level of privacy relative to the average of all vehicles when using the random approach. This is caused by vehicles on less frequented roads hardly having a chance to confuse an attacker by randomly changing their pseudonym. In these scenarios, a concerted approach like SlotSwap (or any other context-aware privacy scheme) is a better choice.

The discontinuities at about 40 s and 90 s are a direct result of the topology of our region of interest: Two highly frequented roads cut the ROI and it took nodes about 40 s and 90 s, respectively, to pass these roads. The set of cars with these lifetimes therefore includes a considerable amount of cars with higher privacy levels, since on busy roads nodes will find potential partners for pseudonym exchanges more easily.

In a second simulation run, we measured the location privacy enjoyed by vehicles on a four-lane freeway in both low (LD, 640 cars/h) and high (HD, 2160 cars/h) traffic volume scenarios (Figure 4.2b).

We found that on a freeway the entropy of nodes increases almost linearly with the lifetime of a car. The cause for this is twofold: Firstly, vehicles almost immediately find a suitable candidate for pseudonym exchange on freeways. Secondly, the lack of intersections and high-density spots negatively influences the level of privacy reached by randomly changing pseudonyms. Even in the HD scenario, vehicles could not reach a similar level of privacy as with SlotSwap in the LD scenario. Our findings suggest that after 10 min on a freeway, even in sparse traffic with a strong

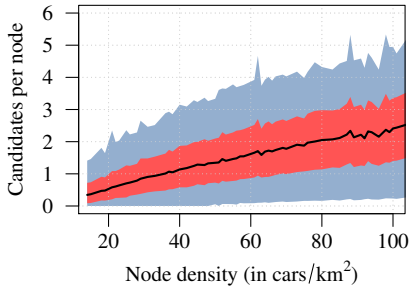
attacker, vehicles implementing the SlotSwap scheme have reached a good level of privacy, under the assumption that pseudonym changes can only be tracked with 95 % certainty.

Pseudonym Exchanges

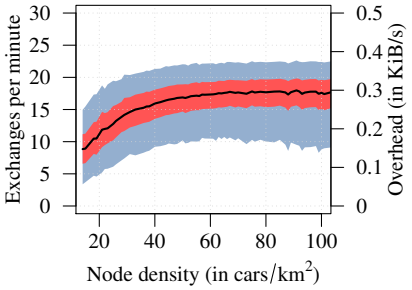
We measured the number of vehicles suitable for exchange of the current pseudonym, the *candidates* of a node, according to our simulation parameters. Our measurements are shown in Figure 4.3a, with the first and second quartile visualized by the red area and the 5 % and 95 % quantiles by the blue area. When node density was $\leq 40\,\text{cars/km}^2$ most of the vehicles were only very infrequently able to find one or more candidates. As expected, the number of candidates rises with the density. With $70\,\text{cars/km}^2$, 75 % of all nodes frequently have one or more nodes suitable for pseudonym exchange nearby. The 5 % quantile is still very low for the $100\,\text{cars/km}^2$ scenario, because there are always nodes traveling on less busy streets, e.g., in residential neighborhoods. It has to be noted that these numbers are of course heavily dependent on the parametrization for a node to be a candidate. Our results in Section 3.1.4 indicate that our choice of values used to determine whether a node is a candidate is rather optimistic, however, finding a candidate node will be more likely in real-world scenarios which frequently exhibit even higher node densities.

Figure 4.3b shows the number of exchanged future pseudonyms per minute, that is, pseudonyms for slots other than the currently active one. One might expect that, with higher density, the amount of pseudonym exchanges also rises. However, as can be seen, exchanges only marginally rise for scenarios with densities higher than $60\,\text{cars/km}^2$. The reason for this is twofold: Firstly, with more nodes in the network, the concurrency of nodes reacting to a beacon message will also increase. That is, new nodes do not only offer more possibilities to exchange a pseudonym, but also compete for requesting exchange from other nodes. Secondly and more importantly, cars preferably exchange their current pseudonym over pseudonyms from other slots. With higher node densities, nodes will find suitable partners for exchanging their current identifier more easily, as previously shown in Figure 4.3a. This also explains the slightly declining 5 % quantile in higher density scenarios. As expected, the traffic overhead caused by SlotSwap is only marginal (Figure 4.3b). It did not exceed an average of 0.5 KiB/s and can therefore be deployed in an ITS without restriction.

Extrapolating our results, the observed pseudonym exchange rates meet a rate of 1200 pseudonyms/h. Assuming that, in a worst-case scenario, traceable pseudonyms are only exchanged when the node carrying it initiates the pseudonym exchange, it would take less than 2 h to exchange the whole pseudonym pool. After this time period a node would only carry untraceable pseudonyms and already be completely



(a) Median candidates per node in the urban scenario



(b) Exchanges of future pseudonyms per node in the urban scenario with resulting traffic overhead

Figure 4.3 – Measurements for neighborship relations and resulting traffic overhead in the urban scenario. Overlaid are the 25 % and 75 % quartiles and the 5 % and 95 % quantiles, respectively.

anonymous when a new slot begins. Deploying a strategy that preferably exchanges pseudonyms for slots in the near future could potentially allow a reduction in the number of pseudonym exchanges per hour without increasing the probability of a vehicle starting a time slot with a traceable pseudonym.

4.1.4 Benefits and Limitations

An advantage of the time-slotted approach over huge pseudonym pools is its ability to ensure that, ideally, a vehicle always has a pseudonym to participate in the ITS as long as it has received its $\frac{p}{t}$ pseudonyms in the setup phase. Even if the CA is not reachable or the car was not used for a longer time period the vehicle will not run out of pseudonyms because it can reuse the old ones. Also, limiting the number of pseudonyms at any point in time to exactly one eliminates the possibility for Sybil attacks, that is, one user pretending to be multiple vehicles at the same time [50].

In addition, our scheme introduces upper limits for disk space and, more importantly, traffic volume. This simplifies the design of on-board units and also reduces the communication costs, making deployment of IVC more affordable. The pseudonym pool size is reduced to a constant value of $\frac{p}{t}$ multiplied by the size of a pseudonym and, more importantly, the workload at the CA is no longer dependent on the number of nodes actually participating in the network but rather on the ones joining it.

Using time slots and GPS-synchronized clocks, every node will change its pseudonym at the same time. Depending on penetration rate and traffic density, this can increase drivers' privacy, as we showed in our evaluation. It also has to be noted that time-based pseudonym changing that does not consider the current situation of

the vehicle may introduce problems in terms of traffic safety as the vehicle could be in a critical situation when a time slot starts. A possibility to circumvent this is to allow vehicles nearby (and thereby also an overhearing adversary) to link new and past pseudonyms [89] and only focus on increasing a user's privacy level when their messages cannot be overheard by an attacker. The notion of 'surrendering' privacy to overhearing adversaries is supported by our findings in Section 3.1.4, where it was almost impossible to confuse an eavesdropping attacker.

However, when a vehicle is currently not in the transmission range of an adversary, simultaneously changing pseudonyms will yield the maximum possible benefit in terms of location privacy as the number of vehicles an adversary is not able to re-identify is maximized. By further applying a pseudonym exchange scheme, the privacy of users can be substantially increased. Allowing the exchange of current and future pseudonyms eliminates the mapping at an authority and allows nodes to start new time slots already anonymously.

Accountability in pseudonym exchange environments remains an open problem. Therefore, the use of our scheme should be limited to non-safety-critical messages to avoid misuse. The class of safety-critical messages includes messages such as accident and emergency brake messages. We argue that for non-critical service messages, but also possibly for periodic beaconing, preservation of unlinkability and privacy is more important than accountability. Each vehicle could then maintain two or more pseudonym pools, one with pseudonyms for safety-critical messages that can be resolved by the CA and one for all non-safety messages. An open challenge in this regard is the revocation of pseudonyms. If there is no mapping from a vehicle's base identity to all of its pseudonyms, revocation of the entire pseudonym pool of a vehicle is a non-trivial task.

While, by design, in our scheme every node has only one valid pseudonym for any point in time, the use of tamper-proof devices is crucial. Manipulated on-board units could be configured to not delete old pseudonyms after exchanging them with another node, allowing an adversary to build up a pool of many pseudonyms, all valid for the same time slot, thus threatening the safety and security of the whole system. A possible approach to ensure that a pseudonym has in fact been deleted are physically uncloneable functions [179]. Secondly, according to the Two Generals' Problem [103], it is theoretically impossible to define a point in time where both parties agree that the exchange has been finalized successfully. It is therefore necessary to reduce the likelihood of an exchange that is considered failed by one vehicle and successful by the other to a minimum. This furthermore requires an error handling protocol, e.g., a vehicle can request a new pseudonym from the CA should it detect such a faulty exchange.

In terms of location privacy, we acknowledge that other context-aware cooperative privacy schemes will offer a similar degree of location privacy. In fact, an

approach in which two nearby cars systematically change pseudonyms at the same time will be as effective as SlotSwap when it comes to complicating tracking for an overhearing adversary. Privacy mechanisms that further take group relations into account instead of only the relation between two vehicles can prevent tracking even more effectively. However, the key strength of SlotSwap is not necessarily the level of privacy obtained during a trip. It is rather the increased efficiency of time-slotted pools and the elimination of the mapping between base identity and pseudonyms at the CA, and, by exchanging future pseudonyms, the anonymous start of a trip.

Due to the discussed technical difficulties that may hinder a real-world deployment of a pseudonym exchange scheme, the main takeaway relevant for the remainder of this thesis is the use of non-overlapping time-slotted pseudonym pools, as they not only increase users' privacy by simultaneous pseudonym changes, but also allow the deployment of a very efficient pseudonym revocation scheme as we will show in the next section.

4.2 SmartRevoc: Efficient and Fast Revocation

If a vehicle sends erroneous data (involuntarily or deliberately), it is desirable to exclude this particular vehicle from the network [17]. This can be done through revocation of all its valid pseudonyms, invalidating messages signed with these pseudonyms. Revocation might also be needed when a vehicle changes ownership [107]. The sooner a vehicle is informed of a misbehaving vehicle the smaller the possibility of damage caused by erroneous data. To be able to trust messages received from other vehicles is a basic requirement for many safety and non-safety applications.

Excluding a vehicle means to invalidate all of its valid pseudonyms. In PKIs, as deployed in the upcoming vehicular networks, this is accomplished by certificate revocation via the distribution of Certificate Revocation Lists (CRLs), which is a list of now-invalidated pseudonyms identified by their public key. The CRL is signed by the CA and can be validated by vehicles using the pre-installed public key of the CA. Messages signed with a revoked certificate must be ignored by receiving vehicles to avoid faulty or forged information being used for traffic safety purposes.

In this section, which is based on our *IEEE Conference on Local Computer Networks* paper "SmartRevoc: An Efficient and Privacy Preserving Revocation System Using Parked Vehicles" [56] and our *IEEE Transactions on Mobile Computing* article "IVC in Cities: Signal Attenuation by Buildings and How Parked Cars Can Improve the Situation" [217],[5] we present *SmartRevoc*, a solution to tackle the challenges that go along with certificate revocation in vehicular networks.

There are several design requirements when developing a certificate revocation mechanism: Firstly, the process has to be as fast as possible, shortening the period of time an attacker may compromise the system for. Thus, a low delay for disseminating new CRLs to participating vehicles is critical to the success of the system.
Secondly, certificate revocation has to be efficient. Assume a scenario where all pseudonyms of vehicles with a certain type of OBU have to be revoked. Distributing a list with millions and millions of pseudonym public keys will result in large amounts of traffic, most likely also influencing the dissemination delay.
Finally, the revocation of vehicles must not compromise drivers' privacy. In particular, the CRL must not reveal and link a vehicle's past pseudonyms as this would allow retrospective association of an individual with a location, e.g., when overheard pseudonyms are stored by an adversary.

The dissemination delay of a CRL is closely tied to the communication technology used for distribution. If all vehicles have cellular internet access, a CRL could be pushed to the vehicles, possibly even via multicast mechanisms, reducing the delay

[5] This journal publication was written in collaboration with the Distributed Embedded Systems Group of the University of Paderborn. It extends the conference paper [68] by the presentation of the obstacle model first presented in [218]. Personal contributions include the use of parked vehicles for safety purposes and the simulation study.

to a minimum. However, traffic over cellular networks is not free. Moreover, even though IEEE 802.11p OBUs are currently discussed to become mandatory, cellular technology will likely remain optional. Thus, a large portion of vehicles are unlikely to be equipped or retrofitted with cellular technology.

Distributing the CRL over a VANET consisting of vehicles and RSUs, i.e., using IEEE WAVE or ETSI ITS-G5, is thus a promising approach; however, this makes both delay and channel load critical properties of the revocation process.

In SmartRevoc, we reduce the channel load incurred by the distribution of CRLs as follows: Our approach makes use of very small CRLs by employing two hash chains and shifting the task of computing which certificates have been revoked to the vehicles. This fulfills both the efficiency and privacy system requirements, because, when used with time-slotted pseudonym pools (see Section 4.1), our system provides backward privacy to all users, that is, it ensures the inability of an attacker to retroactively disclose location information about vehicles with revoked pseudonyms. However, in terms of storage and overhead, our system is just as efficient without the use of time slotted pseudonym pools.

To lower the delay from initial revocation of a pseudonym to wide-area dissemination of a new CRL, we make use of an epidemic dissemination scheme. In the early stages of an ITS, the penetration rate (i.e., the fraction of equipped vehicles) will be small and connectivity will thus be low [7]. However, good connectivity of the network is critical for disseminating new CRLs quickly. We propose the use of parked vehicles to increase connectivity and thereby decrease the delay. As the dissemination of a CRL alone will likely not be a sufficient use case to include parked vehicles in vehicular networks, we also show that they can considerably contribute to road safety by relaying messages from moving vehicles. Particularly in urban scenarios, where radio shadowing can lead to low situational awareness, we show that parked vehicles can offer valuable extra seconds of reaction time.

4.2.1 Related Work

In this section, we discuss relevant articles from the literature. For the sake of readability, this section is divided into three subsections, as our approach touches three different fields, namely the utilization of parked vehicles, certificate revocation, and beacon relaying.

Parked Vehicles

The participation of parked vehicles to support different applications in an ITS has been proposed by several authors [37, 51, 156, 161, 162].

Liu et al. presented a method to use parked vehicles as relay nodes to disseminate information in a DTN fashion [156]. They mainly focus on connectivity and show

that an ITS can greatly benefit from additional nodes. However, they do not provide insights on latency, which is crucial for both traffic safety and the revocation of certificates.

Crepaldi et al. [37], as well as Malandrino et al. [161, 162], expand on the discussion of parked cars as relay nodes to further investigate their usefulness; they propose that parked vehicles can be used to share and provide opportunistic internet access to other vehicles. Subsequently, in [38], they also present an energy management scheme to increase the lifetime of parked vehicles and thus improve the offered service.

Dressler et al. investigate the potentials of using parked vehicles as a temporary network and storage infrastructure [51]. They show that, using protocols from the sensor network domain, they can manage clusters of parked vehicles and enable even routed communication between moving and parked vehicles.

To the best of our knowledge, we are the first to utilize parked vehicles for safety or security related tasks.

Certificate Revocation

Certificate (i.e., pseudonym) revocation in vehicular networks has been widely studied [106, 107, 146, 150, 176, 187], albeit with different goals.

Lequerica et al. propose the dissemination of CRLs using cellular communication [150]. They show that by using multicast mechanisms a CRL can be very quickly distributed. For the aforementioned reasons, we focus on the problem of epidemic dissemination of CRLs using inter-vehicle communication only. Furthermore, we also consider the privacy of users as a feature of the system.

An example of an ad-hoc-only system is the approach presented by Laberteaux et al. [146]; here, CRLs are injected into the VANET by RSUs and then distributed by all moving vehicles. This work showed that the latency substantially decreases if the network density is very high. We show a possible way to also achieve this in sparse scenarios, namely by the participation of parked vehicles.

In order to decrease bandwidth usage when disseminating CRLs, several authors [36, 106] propose that only missing pieces or deltas of the CRL are exchanged between vehicles. *SmartRevoc* uses this method to transmit a CRL once it is discovered that the CRL of another vehicle is not up-to-date.

Another means of reducting the CRL size to increase the efficiency of CRL distribution has been proposed in [176]. The authors recommend splitting the CRL into pieces and to only contain regional revocation information. Even though CRLs are already very small in our approach, this method could be used to further decrease their size.

In [187], the usage of Bloom filters to lower the computational effort was introduced. Bloom filters offer a probabilistic method to check whether a pseudonym is on a CRL. The system design of our approach allows for this method to make pseudonym look-up faster, however, the use of Bloom filters is not within the scope of this thesis.

The work we consider to be the most related to our *SmartRevoc* system has been introduced by Haas et al. [107] (and later by [254]). Contrary to previous schemes, their approach accounts for backward location privacy while using an efficient revocation method. It extends the certificate for slot r by one field, the certificate identifier $C_i = E_{s_i}(r)$, which is the result of a block cipher E or a Cryptographic Pseudorandom Number Generator (CPRNG) that uses elements s_i of a hash chain $s_i = h(s_{i-1})$ as its key. When revoking a certificate, s_i and i are published and vehicles can then compute all subsequent $C_j : i \leq j \leq n$. Our system does not require the use of two different cryptographic functions, but can be based on one cryptographic hash function only, reducing possible security issues, especially when the block cipher is only used with known plain-text, although current ciphers are not believed to be susceptible. Furthermore, their simulation study was not based on packet-level communication, so the specific radio shadowing characteristics of urban environments were not accounted for. Finally, the distribution of CRLs was based on moving vehicles and RSUs only.

Relaying

There have been several publications on safety applications and cooperative awareness using periodic beacon messages in IEEE WAVE or ETSI ITS-G5 [168, 225, 237]. Many of them focus on relaying and retransmitting messages, however they all focus on moving cars only.

Ros et al. propose a beacon-based protocol to increase the reliability of VANETs while minimizing the number of beacon retransmissions [193]. In their approach, local position information is used by cars to determine whether they belong to a connected dominating set and subsequently reduce waiting periods before retransmissions. A similar approach, extended to a 2-hop neighborhood, was presented by Khan et al. [128]. They further exploit geographic location, speed, and direction information. Based on this information, nodes will choose a retransmission strategy for periodic beaconing. 3-hop connectivity was investigated in the scope of the FleetNet project [91]. It has been shown that the available capacity on the wireless channel is sufficient to support safety protocols on these connections.

The idea of placing road side infrastructure – backbone-connected RSUs or autonomous SSUs – in order to strengthen connectivity between moving nodes has also been discussed in the literature. Lochert et al. studied the impact of connected

SSUs to improve the performance of ITS applications in the roll-out phase [157]. They found that those static units can significantly improve connectivity between nodes. Furthermore, Ding et al. presented SADV, an approach that utilizes static nodes at road intersections in order to improve data dissemination in vehicular networks [47]. They use a store-and-forward algorithm to overcome problems in scenarios with low node densities.

Similarly, the protocol presented in [224, 225] integrates beaconing between moving vehicles and available RSUs or SSUs by observing the available channel capacity.

Tang et al. investigated timings for collision avoidance systems [232]. They introduced the *time to avoid collision* metric, which represents the time from detecting a potential collision to the point of just avoiding a collision. We make use of an adaptation of this metric to illustrate the benefit introduced by utilizing parked vehicles.

In general, retransmissions, beaconing, and relaying are well-studied subjects. Including parked vehicles as relay nodes offers a promising possibility to complement these concepts, as parked cars are often placed in advantageous positions – along urban streets. We show that safety applications greatly benefit from this approach, especially in the transitional phase, i.e., when IVC communication devices are being introduced.

4.2.2 Utilizing Parked Vehicles for Safety

As we stated earlier, SmartRevoc can utilize parked vehicles to disseminate CRLs in an epidemic fashion. However, the participation of parked vehicles in an ITS needs a convincing argument. We therefore study how parked vehicles can contribute to road safety, the most important goal for the introduction of IVC.

IVC-based traffic safety relies on the successful reception of packets from other vehicles. The earlier a vehicle knows about the presence of nearby cars, the earlier it can inform the driver about a potentially critical situation. When relied upon, late or missing information (caused, e.g., by lost messages) can lead to accidents.

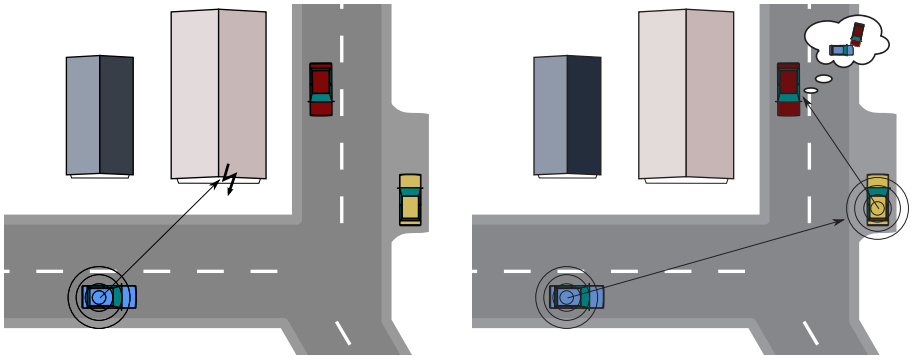
There are several possible causes for the transmission of a safety message to fail. Interference with packets sent by other vehicles can reduce the SINR and make it impossible to successfully decode the packet. Also, even when there is no interference, the received signal strength may be too low to decode the packet. This can be caused by common directionality characteristics of antennas, meaning they will not emit the signal in all directions with uniform strength [145]. More importantly, radio propagation effects such as slow fading or fast fading can significantly reduce the theoretical transmission range of vehicles (see Section 2.3.2).

In this study, we concentrate on signal loss due to radio shadowing caused by buildings. In metropolitan areas the line of sight between vehicles is often blocked by obstacles such as buildings, vegetation, or parked and moving vehicles [22, 44]. This does not necessarily result in the loss of a packet but still leads to a considerable attenuation of the signal (Figure 2.12, p. 39). In a series of real-life experiments, we showed that packet loss caused by radio shadowing is a challenge for safety applications that rely on cooperative awareness [218]. Other nodes, although they are within transmission range, may not sense emitted beacons of a node until both nodes move closer to each other. The time it takes for both nodes to get into communication range constitutes an additional delay that can have a negative influence on safety applications.

A possible approach to maintain high situational awareness despite heavy radio shadowing is multi-hop beaconing or *relaying* [86, 138]. Vehicles will not only periodically broadcast their own beacons but also retransmit received beacons from other nodes. With a high enough penetration rate and traffic density, this approach was shown to improve the cooperative awareness among all nearby nodes. However, in the early stage of ITS deployment, the number of vehicles equipped with OBUs (and, thus, the amount of neighbors with which a vehicle can communicate) will be low [7]. Furthermore, there will always be low traffic density spots in suburban regions where relaying by other moving vehicles is not possible. Finally, during off-peak hours and at night, even in the city center, traffic density can be expected to drop substantially.

To overcome this problem, we argue that parked vehicles should participate as message relays for moving cars. A vehicle is parked for 23 h a day [153]; we believe it would therefore be negligent to not include them in vehicular network communication. Further insights are given by a study on parking behavior in the area of Montreal, Canada [166]: In 2003, out of 61 000 daily parking events, 69.2 % of all parked cars were parked on streets while only 27.1 % were parked on outside parking lots. A minority of 3.7 % were parked in interior parking facilities. On average, the duration of one parking event was about 7 h. The study furthermore shows that parked vehicles were distributed throughout the whole city, which means there is a high probability that a parked car is within transmission range of a moving car.

A benefit of parked cars is their parked position itself. This idea is shown in Figure 4.4, where relaying by the parked yellow vehicle increases situational awareness of the red vehicle, as it was previously unable to receive the beacon sent by the blue vehicle. In these scenarios, parked vehicles function just like an SSU, however, we see a major benefit in the ubiquitous availability of such parked cars in comparison to costly RSUs and SSUs.



(a) Safety message lost due to shadowing **(b)** 2-hop relaying using a parked vehicle

Figure 4.4 – Relaying messages from moving vehicles with the help of parked vehicles can route around obstacles to increase situational awareness and reduce the risk of traffic accidents.

To obtain an upper bound for the safety benefit obtainable by utilizing parked vehicles as relay nodes, we investigate the following two-hop relaying scheme: A parked car will rebroadcast beacon messages from moving vehicles so that other moving cars will then pick up the beacon. Beacons generated by moving vehicles have a Time-to-Live (TTL) value of 1. When another node receives one of these beacons, it decreases the TTL to 0 and retransmits it. Packets with a TTL of 0 are never rebroadcast.

In a final system, a carefully designed relaying algorithm needs to be deployed in order to keep channel usage low but still ensure a benefit close to the upper bound. Possible solutions include the restriction of relaying to only special nodes, for example, nodes that are parked close to intersections [15]. Furthermore, a relaying node could be able to autonomously assess whether packet relaying helps improve cooperative awareness for nearby nodes by observing neighborship relations including movement information such as speed or direction [235]. Also, evaluating current channel conditions in order to determine whether a packet should be relayed seems to be a promising approach [224].

Energy Management

From a networking perspective, the general advantage of vehicles is that they are seemingly energy autonomous: As vehicles move, their battery is continuously recharged. However, parked nodes do not have this virtually unlimited supply of power as their battery does not recharge while the engine is turned off. Powering the transceiver when the engine is turned off is unproblematic as modern vehicles are equipped with dedicated electronics to keep certain devices powered on when the

vehicle is not moving – and to cut power to these devices when the battery charge drops below a certain point.

A typical IEEE WAVE OBU should not drain more than 1 W on average, which is a very generous upper limit. Considering the battery of a small car, providing 480 Wh to 840 Wh [5], the system can run for 20 days, fully draining the battery. Assuming that we allow use of at most 10 % of the battery's capacity for relaying messages, the parked vehicle could still participate for 2 days. Looking at electric or hybrid vehicles, the problem of draining the battery with the OBU becomes negligible. For example, the battery of a Tesla Roadster has a capacity of 53 000 Wh providing energy for several years of constant radio transmission. Still, it is obligatory that the OBUs of parked vehicles do not discharge the battery below a point where the car cannot be started again. There must always be enough power left for the ignition and other mandatory functions of the vehicle. There are two possibilities to overcome this problem: Either power to the on-board device is cut (or it switches itself off) when the battery is drained below a threshold, or the DSRC device is equipped with a dedicated battery that is also recharged when the car moves again.

In conclusion, we can say that the use of such a relay system for a parking time of less than one day is without any critical impact on the usability of the vehicle. For the remainder of this study, we therefore assume that, without loss of generality, all cars always have enough energy left to operate their OBU.

4.2.3 Evaluation of Relaying-Based Safety

We performed extensive simulations using our Veins framework to show the effectiveness of using parked cars to support safety applications in vehicular networks. We demonstrate that traffic safety can be considerably improved, justifying the participation of parked vehicles in an ITS. This in turn allows other (non-safety) applications such as SmartRevoc to also make use of these stationary nodes.

Simulation Setup

We investigate two different scenarios where obstacles can heavily influence radio transmission: a synthetic grid scenario and the more realistic Ingolstadt scenario (see Section 3.1.3).

We allowed vehicles to park anywhere along the street in the grid scenario. For the Ingolstadt setting, we added parking areas based on satellite data and distributed vehicles corresponding to the size of the parking area. Vehicles were allowed to park anywhere in these areas, their locations following a random uniform distribution according to the findings presented in [166]. As before, traffic was generated for the entire scenario but results were only obtained from vehicles in a 1.5 km^2 ROI.

Serving as a baseline for comparison, we also deploy RSUs that are acting as relays as well. In order to provide optimal conditions for message dissemination, RSUs are deployed on the busiest intersections in the exact center of a junction to maximize their coverage. For the examined 2-hop (i.e., 1 relay) forwarding scenario, RSUs do not need to make use of a backbone connection and are thus functionally equivalent to SSUs.

In both scenarios, all moving vehicles (but no parked vehicles) emit beacon messages (representing CAMs messages) with 1 Hz. The beacons could then be relayed in a 2-hop fashion by nodes in the immediate neighborhood. We configured these relaying nodes to re-transmit beacons only after a short random processing time of 1 ms to 10 ms. Depending on the simulated configuration we enabled a different subset of relays: either moving cars only, parked cars only, RSUs only, or a combination of moving vehicles and one type of stationary node.

Table 4.2 gives an overview of parameters and terminology used in the following discussion.

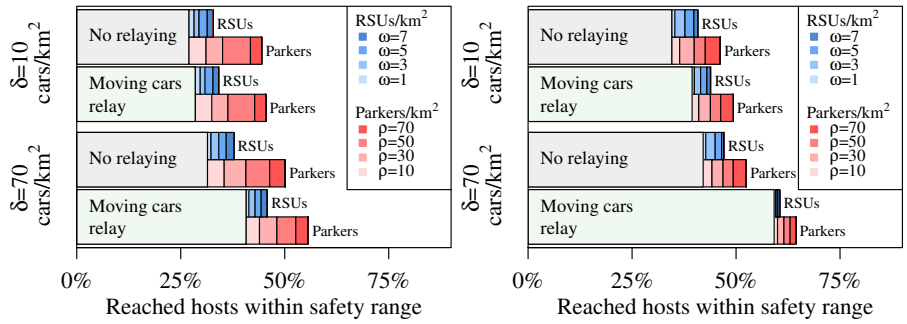
Parameter	**Value**
Scenario	Manhattan grid, suburban (Ingolstadt, Germany)
ROI size	1.5 km^2
Metrics	Reachability, time benefit
Technology	IEEE WAVE
Path loss model	Obstacle path loss (Section 2.3.2)
Obstacle parameters	$\beta = 9\,\mathrm{dB}$, $\gamma = 0.4\,\mathrm{dB/m}$
Beacon Frequency	1 Hz
Amount of moving (equipped) cars δ	0 cars/km^2 to 90 cars/km^2
Amount of parked (equipped) cars ρ	0 parkers/km^2 to 90 parkers/km^2
Amount of Roadside Units (RSUs) ω	0 RSUs/km^2 to 7 RSUs/km^2
t_{org}	Reaction time without relaying
t_{rel}	Reaction time with relaying

Table 4.2 – Simulation setup and parameters for the parked cars study.

In order to accurately measure the safety benefit provided by a system, one would have to identify certain classes of constellations between vehicles, obstacles, and parked cars. The classification of these cases, however, is an open challenge and a 100 % coverage of all cases cannot be guaranteed [126]. Therefore, we chose as the primary metric in our simulation the ratio between the number of vehicles in a theoretical maximum safety range that could be reached with beacon messages and the total amount of vehicles in this range.

We thus obtain a ratio describing the *reachability* of nodes in the network. The safety range (which corresponds to the maximum unobstructed transmission range of a node) was configured to be 400 m, as we believe that nodes further away do not play an important role for safety applications in urban environments.

In order to be able to obtain baseline measurements, we employ a modified IEEE 802.11p medium access scheme that is idealized (collision free). This allows us to abstract away from the effects that real-world protocols would necessarily need to introduce in order to coordinate fair and scalable medium use. We can thus give an upper bound on the number of possible data transmissions that is independent of the used protocol. A medium access scheme for parked vehicles is not within the scope of this work as we focus on giving insights on the upper bound of such a system. For possible channel access mechanisms we refer the reader to [135] where we show the applicability of *p*-persistence-based protocols incorporating the amount of neighboring vehicles.



(a) Manhattan grid **(b)** Suburban scenario

Figure 4.5 – Comparison of beacon relay approaches in the suburban and Manhattan grid scenarios.

Relaying: Parked Cars vs. RSUs

As a first step, we examined how relaying to route around obstacles can improve the amount of reachable moving vehicles within a safety range.

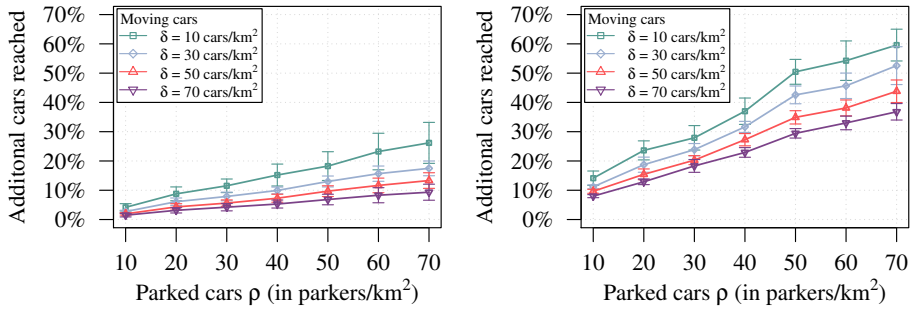
Figure 4.5 shows our findings. It can be seen that the effect of obstacle shadowing in urban environments is prominent, causing the percentage of reached hosts to drop to values of 25 % to 40 % in both the low node density (LD, $\delta = 10\,\text{cars/km}^2$) and high node density (HD, $\delta = 70\,\text{cars/km}^2$) scenarios (Figure 4.5a, gray bars 'No relaying'). Results were slightly better in the suburban scenario due to the lower density of obstacles (Figure 4.5b). By enabling relaying on moving nodes only (light green bars 'Moving cars relay'), percentages just marginally improved when the

traffic volume was low, while helping reach a considerably higher percentage of nodes when the traffic density was high.

We then enabled relaying on either RSUs or parked vehicles with and without the help of relaying moving cars. In the Manhattan grid scenario, exclusive relaying by parked cars (red stacked bars) clearly outperformed regular VANET broadcasting. The reason for this is that, in the synthetic Manhattan scenario, every parked car is a good relay node candidate, as it parks on the curbside next to a building. There were no ineffective parked vehicles in contrast to the suburban scenario, where parking spaces are not necessarily located next to a building, but also in areas not suitable for relaying around obstacles; here, the amount of newly reached vehicles is accordingly lower.

Comparing the performance of parked vehicles to optimally placed RSUs, we observe that, in both the suburban and the Manhattan grid scenarios, as little as $\rho = 30\,\text{parkers/km}^2$ and $\rho = 15\,\text{parkers/km}^2$, respectively, yield the same level of cooperative awareness as $\omega = 7\,\text{RSUs/km}^2$ – yet with zero deployment cost. We conclude that, in areas with heavy obstacle shadowing, the aid of parked cars can considerably boost cooperative awareness and therefore reduce the number of needed RSUs. In a more open area, RSUs can only marginally increase the amount of reachable hosts, unless deployed in an excessive number. Parked cars, however, can be expected to be available in high numbers in these suburban residential neighborhoods. We also observe that the set of vehicles additionally reached with the aid of stationary nodes is not a subset of the nodes reached with moving vehicles, as the number of reached hosts still increased.

To obtain a more detailed insight on the benefit of relaying by parked cars, we parametrized both node density and the amount of stationary nodes. We therefore



(a) Parked cars, suburban scenario

(b) Parked cars, Manhattan grid

Figure 4.6 – Increase in message delivery success when additionally using parked cars as relay nodes. Error bars show the 95 % confidence intervals.

measure the amount of additionally reached cars compared to the number of cars reached with moving vehicle relaying only.

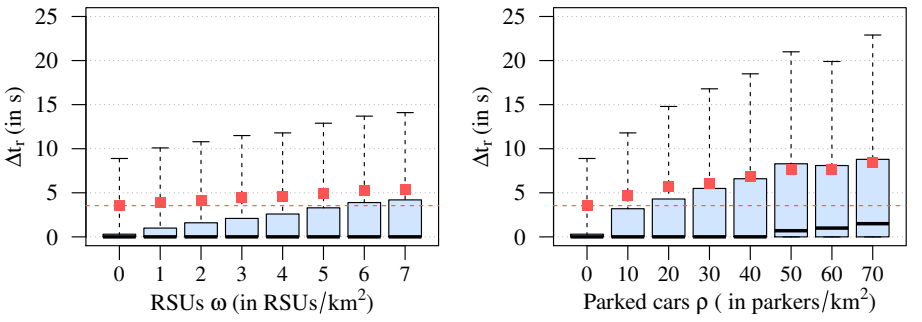
In all setups presented in Figure 4.6, we observe that the benefit of parked vehicles (and RSUs, not shown) is higher when the traffic volume δ was lower. With more moving vehicles in the network, the probability of a vehicle previously being unreachable due to a blocked radio signal – but reachable through an intermediate moving vehicle – increases, reducing the benefit of stationary nodes.

As expected, the benefit is higher in the Manhattan grid scenario due to the many obstacles potentially blocking radio communication. With a large amount of parked vehicles relative to moving vehicles we observe that over 50 % additional vehicles could be reached. Even the highest simulated density of parked cars ($\rho = 70\,\text{cars/km}^2$) in both scenarios can be seen as realistic [166]. In conclusion, we observed that the benefit seems to grow somewhat linearly with the number of parked vehicles.

Improved Reaction Times

When investigating safety applications, not only is the amount of cars reached by a broadcast message relevant, but also how early vehicles become aware of the existence of a nearby car. The earlier an in-car safety system knows of the presence of another vehicle, the earlier it can notify the driver or prepare active and passive safety systems.

For each pair of vehicles we therefore tracked when they first became aware of one another, taking their time t_{known} of receiving the first periodic beacon. We also



(a) Additional afforded reaction time through RSU relaying.

(b) Additional afforded reaction time through parked car relaying.

Figure 4.7 – Additional afforded reaction time Δt_r in critical situations ($t_{\text{org}} \leq 3\,\text{s}$): time difference achieved when enabling relaying via moving vehicles only ($\omega = \rho = 0$) as well as when supported by either RSUs ($\omega > 0$) or parked cars ($\rho > 0$). Plotted are results for the low-density Manhattan grid scenario ($\delta = 20$ cars/km^2).

tracked how much later these two vehicles actually met at an intersection, taking their time of closest distance t_{met}. We term this time difference ($t_{\text{met}} - t_{\text{known}}$) the reaction time. We then compare t_{org}, that is, the original reaction time without the help of stationary nodes (RSUs or parked cars) and only relaying of moving vehicles, with t_{rel}, the reaction time with relaying on stationary nodes enabled. We isolate critical situations where $t_{\text{org}} \leq 3\,\text{s}$ [31] and investigate to what extent such situations could be mitigated by support of either parked cars or RSUs. We quantify the achieved safety benefit as the difference in afforded reaction time $\Delta t_{\text{r}} = t_{\text{rel}} - t_{\text{org}}$.

The results are displayed in Figure 4.7 in the form of a box plot. The boxes reach from the 25 % to the 75 % quantile while the whiskers extend from the 5 % quantile to the 95 % quantile. The bold line marks the median. As the distribution of the recorded data was heavy-tailed, we also plot the mean of recorded values (small red squares); for the sake of easier comparison we plot the mean for the cases where only moving cars relayed ($\omega = \rho = 0\,\text{nodes/km}^2$) as a red dashed line.

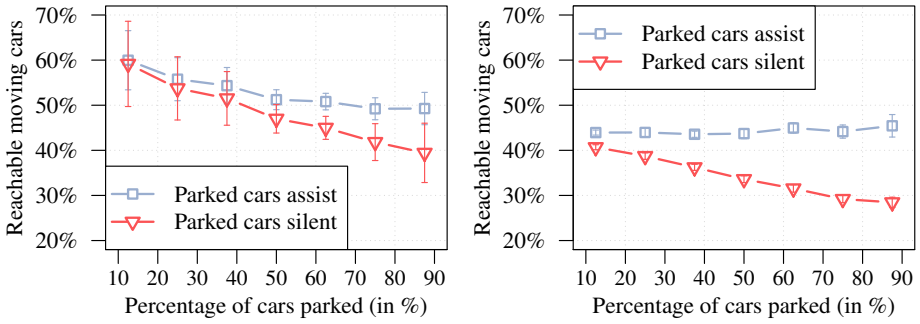
We investigate the Manhattan grid scenario with a fixed low traffic density of $\delta = 30\,\text{cars/km}^2$. We observe that for 50 % of all cars there is no improvement when allowing only moving nodes to relay messages ($\omega = \rho = 0$). Adding RSUs to support relaying (cf. Figure 4.7a) can add valuable extra seconds to the afforded reaction time, albeit only for a small portion of drivers. This stems from the fact that an RSU can only help improve safety at the particular junction at which it is placed. Thus, even very high deployment densities of $7\,\text{RSUs/km}^2$ do not suffice to noticeably increase the median time benefit above 0 s.

This is in contrast to results obtained by enabling relaying via parked vehicles, of which a much higher number is available (Figure 4.7b). Even a small portion of parked vehicles ($\rho \geq 40\,\text{parkers/km}^2$) results in a clear improvement, giving at least 50 % of drivers valuable extra seconds to react. Finally, if the full number of $\rho = 70\,\text{parkers/km}^2$ are available, afforded additional reaction times increase to levels that might allow at least half of all critical situations to be defused.

Day-to-Night Transition

As a last step, we examine the effect of moving vehicles becoming stationary. This scenario can be understood as the typical day-to-night transition where the traffic volume decreases and reaches a minimum sometime in the night. The previously moving vehicles, however, may still be parked along the street and can therefore be used as relay nodes in a vehicular network. In our setup, we considered the amount of total vehicles $\delta + \rho$ as invariant, but varied the ratio $\frac{\delta}{\rho}$.

Our findings are presented in Figure 4.8. For the suburban scenario, we observe that the level of situational awareness could not be fully maintained when the number of moving nodes decreases (Figure 4.8a, blue line). However, without parked cars



(a) Suburban scenario **(b)** Manhattan grid

Figure 4.8 – Percentage of hosts reached within the safety range plotted for a typical day-to-night transition. While the density of active nodes in the scenario is invariant ($\delta + \rho = 80$ cars/km^2), the ratio of moving to parked vehicles ($\frac{\delta}{\rho}$) decreases. Error bars visualize the 95 % confidence intervals.

as relay nodes, this curve drops considerably faster (red line). In the synthetic Manhattan grid scenario, the effect of aiding parked cars is again clearer. Although the density of moving nodes δ became lower and lower, the lack of relaying moving nodes could be completely compensated for by parked cars. The percentage of reachable hosts within the safety range only varied marginally (blue line) when moving vehicles further participated in the network as parked ones. In contrast to that, the level of awareness considerably dropped when this was not the case (red line).

IVC-based traffic safety is particularly important at night, when lighting conditions are worse and drivers might be more inclined to drive faster on the now almost empty streets. With the help of parked cars, vehicles can experience the same level of cooperative awareness at night, as if there were still many more moving vehicles on the street.

In conclusion, we showed that parked cars can substantially improve situational awareness in urban and suburban scenarios and thereby improve the effectiveness of IVC-based traffic safety. The fact that they come with no additional deployment cost and are readily available at promising positions is a strong argument for their consideration as relay nodes as it would require an excessive number of RSUs to reach the same level of situational awareness. Independent from day and night, we furthermore conclude that with parked cars we can achieve the same level of cooperative awareness in sparsely populated areas as we would have in those with many moving vehicles. Considering these advantages, we deem the participation of parked vehicles in future ITS's a realistic assumption and will make use of the possibility to also use them for other applications.

4.2.4 Backward Privacy-Preserving Revocation

IEEE WAVE and ETSI ITS-G5 are similar in many aspects, however, while certificate revocation is not planned in ETSI ITS-G5, it is an inherent and important part of the IEEE WAVE family of standards. The National Highway Traffic Safety Administration published a technical report in August 2014 listing use cases where certificate revocation is needed [111]: These include administrative revocation (e.g., retired vehicles) and revocation of vehicles sending obviously erroneous data as identified by other vehicles, the vehicle itself, or a regulation authority. Each pseudonym of the to-be-revoked vehicle will then be published on a CRL. Alternative methods, some very similar to our SmartRevoc concept are also discussed. However, it needs to be noted that the document was published later than our original SmartRevoc publication, and could therefore not be considered for the development of our revocation scheme. They also state that "the revocation process has not yet been finalized" [111] and give information on possible methods on how a CRL could be designed.

In order to keep the size of a CRL low, we include an additional identifier (or linkage value) C_v^i in each certificate C_v of a vehicle v. The purpose of this identifier it to allow each vehicle to compute all revoked pseudonyms in a single vehicle's v pseudonym pool $P_v^\star$ after a secret key (or seed) κ_v has been released by the CA. The identifiers C_v^i $\forall C_v \in P_v^\star$ are linked through κ_v which allows the computation of C_v^{i+1} from C_v^i. Identifiers and secret keys are generated and stored by the CA which also adds the identifiers to the certificates during the signing process. When receiving a message from another vehicle, a vehicle can then simply check whether a certificate contains a revoked identifier C^i. An overview of the used notation and symbols is given in Table 4.3.

Notation	Description
N	Size of certificate pool
C_v^i	Identifier of certificate i of vehicle v
$\mathrm{CRL}_{(v,i)}$	CRL entry for vehicle v, starting with certificate i
$h(\cdot, k)$	Keyed cryptographic hash function, using key k
$h(\cdot, k)^{(i)}$	$h(\cdot, k)$ applied to its own result i times
$s(\cdot)$	Cryptographic hash function
κ_v	Key for vehicle v stored at CA
κ_v^i	Key for vehicle v, hashed $i-1$ times using $s(\cdot)$
$P^\star$	Pseudonym pool of a vehicle

Table 4.3 – Overview of notation used in this section.

To allow decentralized computation of revoked certificates the CA uses a known, keyed cryptographic hash function $h(\cdot, \kappa_v)$ to generate a set of linkage values $C_v^1, \ldots, C_v^i, C_v^{i+1}, \ldots, C_v^N$ as

$$\begin{aligned} C_v^1 &= \text{rand}() \\ C_v^{i+1} &= h(C_v^i, \kappa_v) = h(C_v^1, \kappa_v)^{(i)} \end{aligned} \tag{4.1}$$

This leads to a hash chain of identifiers

$$C_v^1 \xrightarrow{h(C_v^1, \kappa_v)} C_v^2 \xrightarrow{h(C_v^2, \kappa_v)} C_v^3 \xrightarrow{h(C_v^3, \kappa_v)} \cdots \tag{4.2}$$

A CRL entry to revoke n certificates after C_v^i (typically $n = N - i$ to revoke all valid remaining pseudonyms of a pool of size N) would be:

$$\text{CRL}_{(v,i)} = \left(C_v^i, \kappa_v, n\right) \tag{4.3}$$

The idea behind such a system is to achieve backward privacy because it is not possible to compute C_v^{i-1} using C_v^i if the cryptographic hash function $h()$ is not broken. If pseudonyms can be reused, the CRL entry must always be (C_v^1, κ_v, n) to revoke all pseudonyms of a vehicle; backward privacy can then not be achieved as it links all passed used pseudonyms of a vehicle.

Require: C^j, $\text{CRL}_{(v,i)} = \left(C^i, \kappa_v, n\right)$, N
1: **for all** $m \in [1, N-n]$ **do**
2: **if** $h(C^j, \kappa_v)^{(m)} = C^i$ **then**
3: **return true** ▷ C^j is an identifier linked to C_v^i
4: **end if**
5: **end for**
6: **return false** ▷ C^j is not an identifier linked to C_v^i

Algorithm 4.3 – Attack to reveal whether any C^j is a past identifier of a pseudonym pool revoked in $\text{CRL}_{(v,i)}$.

However, even if pseudonyms are not reused, this scheme allows an attacker to run a feasible brute-force attack to reveal whether a past observed identifier C^j belongs to the same pseudonym pool as a currently revoked C_v^i: Assume an adversary has stored a large set of pseudonym certificates including identifiers C^j along with possibly privacy-invading information such as the location of the vehicle at the time. Given the CRL entry (C_v^i, κ_v, n) they can now try to use $h(\cdot, \kappa_v)$ on every stored identifier C^j. If C^j is in fact a past pseudonym from the same pseudonym pool as the revoked one, executing $h(C^j, \kappa_v)^x$ would yield C_v^i, considering the distance between C^j and C_v^i in the pseudonym pool is x. The adversary cannot know the distance x

and therefore has to execute $h(C^j, \kappa_v)$ at most $N-n$ times. If after $N-n$ iterations the cryptographic hash function has not yielded C_v^i then C^j can be ruled out as a possible past pseudonym of vehicle v. The algorithm is given in Algorithm 4.3. When a match is found, not just the current and future privacy of vehicle v is compromised, but also part of its past privacy (i.e., its identity during the time between using C^j and the publication of the CRL).

Our *SmartRevoc* scheme therefore makes use of a second, unkeyed[6] cryptographic hash function $s(\cdot)$. This function is used to construct a new secret κ_v^i for each step of generating C_v^i, yielding

$$\begin{aligned} \kappa_v^1 &= \kappa_v \\ \kappa_v^{i+1} &= s(\kappa_v^i) \\ C_1 &= \text{rand}() \\ C_{i+1} &= h(C_v^i,\, \kappa_v^{i+1}) \end{aligned} \tag{4.4}$$

This results in two hash chains, one for the keys κ and one for the identifiers C_v^i.

$$\begin{array}{ccccccc} \kappa_v & \xrightarrow{s(\kappa_v)} & \kappa_v^2 & \xrightarrow{s(\kappa_v^2)} & \kappa_v^3 & \xrightarrow{s(\kappa_v^3)} & \cdots \\ & & \downarrow & & \downarrow & & \\ C_v^1 & \xrightarrow[h(C_v^1,\kappa_v^2)]{} & C_v^2 & \xrightarrow[h(C_v^2,\kappa_v^3)]{} & C_v^3 & \xrightarrow[h(C_v^3,\kappa_v^4)]{} & \cdots \end{array} \tag{4.5}$$

A CRL entry to revoke n certificates after C_i is then given by:

$$\text{CRL}_{(v,i)} = \left(C_v^i,\, \kappa_v^{i+1},\, n\right) \tag{4.6}$$

In order to execute a similar brute-force attack as shown before, an attacker needs to now also know a previous key κ_v^j with $j < i$ to reveal the relation of any C^j with C_v^i. However, this key is never disclosed and it is not possible to compute κ_v^{i-1} from κ_v^i due to the nature of the cryptographic hash function s. A brute-force attack would then consist of guessing κ_v^{i-1} until $s(\kappa_v^{i-1}) = \kappa_v^i$, an attack that can be considered infeasible when κ has a length of 128 bit or more. SmartRevoc therefore provides backward privacy as it does not allow linking of past pseudonym certificates to revoked ones.

An alternative design has been proposed by Whyte at al. [254] after our original SmartRevoc publication. Instead of two hash chains, they propose the use of one hash chain to link all κ_v^i and an additional known encryption function e (see Equation 4.7). Certificate identifiers are then directly computed from κ_v^i, and a CRL entry consists only of this key and the number n of revoked pseudonyms. This potentially makes the CRL entries even smaller. An in-depth cryptanalysis and comparison of this

[6] For ease of implementation, the same $h(\cdot,\cdot)$ can be used with a fixed dummy key to supply $s(\cdot)$.

and our design is not within the scope of this thesis, as it depends not only on the identifier and key lengths, but, more importantly, on the deployed functions h and e and on the pseudonym pool size.

$$\begin{array}{ccccccc} \kappa_v & \xrightarrow{h(\kappa_v)} & \kappa_v^2 & \xrightarrow{h(\kappa_v^2)} & \kappa_v^3 & \xrightarrow{h(\kappa_v^3)} & \cdots \\ \downarrow e(\kappa_v) & & \downarrow e(\kappa_v^2) & & \downarrow e(\kappa_v^3) & & \\ C_v^1 & & C_v^2 & & C_v^3 & & \cdots \end{array} \tag{4.7}$$

4.2.5 Overhead and Distribution

Our approach requires certificates to be extended by only one field, the certificate identifier C^i, which is the output of a cryptographic hash function, such as the SHA-256. A typical [107] size for this value is 16 B. Depending on the size of the pseudonym pool and the resulting hash collision probability, a longer or smaller value can be chosen, either by truncating the hash-output or by using a hash function with a longer output. Assuming a pseudonym is valid for 600 s [67, 107] and pseudonyms are not re-used, a vehicle needs to store 52 560 pseudonym certificates for one year. Thus, *SmartRevoc* will require an additional 800 kB of storage space on the vehicle. Considering recent recommendations of storing only a week's worth of pseudonyms [111], this value would decrease to approximately 16 kB.

The certificate authority only needs to store an additional κ_v for each participating vehicle v, resulting in an overheard of 16 B per vehicle. The probability of hash collisions, that is, two pseudonyms having the same certificate identifier, depends on the size of C^i. As an alternative, if hash collisions must be completely avoided, the CA could save all issued certificate identifiers, checking against collisions before signing pseudonyms, and, in case of a collision, choose a different κ_v. Storing all identifiers would then take up an additional $N \cdot 16$ B for each vehicle, N being the size of the pseudonym pool.

In terms of message overhead, the extra field in the certificate introduces an additional 16 B per transmitted message. The possibility of including linkage values into certificates has also been discussed and recommend by the US DOT [111]. To decrease message overhead when the channel becomes congested, certificate omission schemes can be applied [88]. For the distribution of the CRL, the currently known CRL version is piggybacked on periodic safety beacon messages which are sent with a frequency of up to 10 Hz [79, 196]. The CRL version can be an integer of 4 B and is only attached to a beacon message once every second, resulting in a negligible overhead of 4 B/s per vehicle.

When used with 16 B identifiers, a SmartRevoc CRL uses 36 B for one revoked vehicle, consisting of $C_v^i = 16$ B, $\kappa_v^i = 16$ B, and $n = 4$ B. Considering certificate

overhead (every update of the CRL has to be signed by the certificate authority), more than 30 vehicles can be revoked within one packet. The CRL is injected by RSUs (or possibly a vehicle with cellular internet access) and then distributed in an epidemic fashion using both moving and parked vehicles.

When a node detects that its own version of the CRL is higher than the one piggybacked in a safety beacon received from another vehicle it will try to broadcast the delta of the CRL. However, to keep channel load and packet collisions low, vehicles do not broadcast CRL updates immediately (avoiding what is commonly known as a broadcast storm). Instead, similar to common broadcast suppression schemes [258], CRL broadcasts are delayed by a random time (up to 1 s and 10 s for moving and parked vehicles, respectively). The broadcast will only be performed if no other vehicle broadcasts a CRL during this time. Alternatively, persistence-based access schemes can be applied to further reduce channel congestion caused by CRL distribution [135].

4.2.6 Simulation Study

In addition to the analytic study to show the correctness of our approach, we investigated the behavior and effects of our CRL distribution scheme in an extensive simulation study using our Veins simulation framework. Radio propagation calculations make use of our obstacle model (Section 2.3.2) to accurately model signal attenuation by buildings in a computationally efficient way. PHY and MAC simulation employs a fully-featured IEEE 802.11p model (Section 3.2), configured to operate

Parameter	**Value**
Scenario	Manhattan grid (ROI = 16 km^2)
	Ingolstadt (ROI = 8 km^2)
Metrics	Coverage, delay till 95 % coverage
Technology	IEEE WAVE
Transmit power	20 mW
Radio sensitivity	−92 dBm
Path loss model	Obstacle path loss (Section 2.3.2)
Obstacle parameters	$\beta = 9$ dB, $\gamma = 0.4$ dB/m
Beacon frequency	1 Hz
Amount of moving (equipped) cars δ	3 cars/km^2 to 35 cars/km^2
Amount of parked (equipped) cars ρ	0 parkers/km^2 to 35 parkers/km^2
Amount of Roadside Units (RSUs) ω	1 RSUs to 16 RSUs

Table 4.4 – Simulation setup and parameters for the SmartRevoc CRL dissemination study.

in single radio / single channel mode. The application layer, piggybacking CRL identifiers on messages and disseminating new CRLs, is implemented as described. A summary of all relevant configuration parameters is given in Table 4.4.

We simulated the dissemination of the CRL in two different scenarios: the suburban Ingolstadt, Germany scenario and the Manhattan grid setting (see Section 3.1.3). As before, parking in the Ingolstadt scenario was based on satellite data, while parking in the grid scenario was allowed alongside any street. In order to study the dissemination speed, we defined ROIs considerably larger than before, using a 16 km^2 ROI for the grid and an 8 km^2 one for Ingolstadt.

According to [257], different penetration rates can be characterized by simulating different traffic densities. Traffic densities were therefore chosen to reflect an early stage of an ITS deployment, where a penetration rate of higher than 10 % cannot be assumed after one year of operation [7]. Each scenario was simulated with low, medium, and high traffic density considering a penetration rate of 10 %.

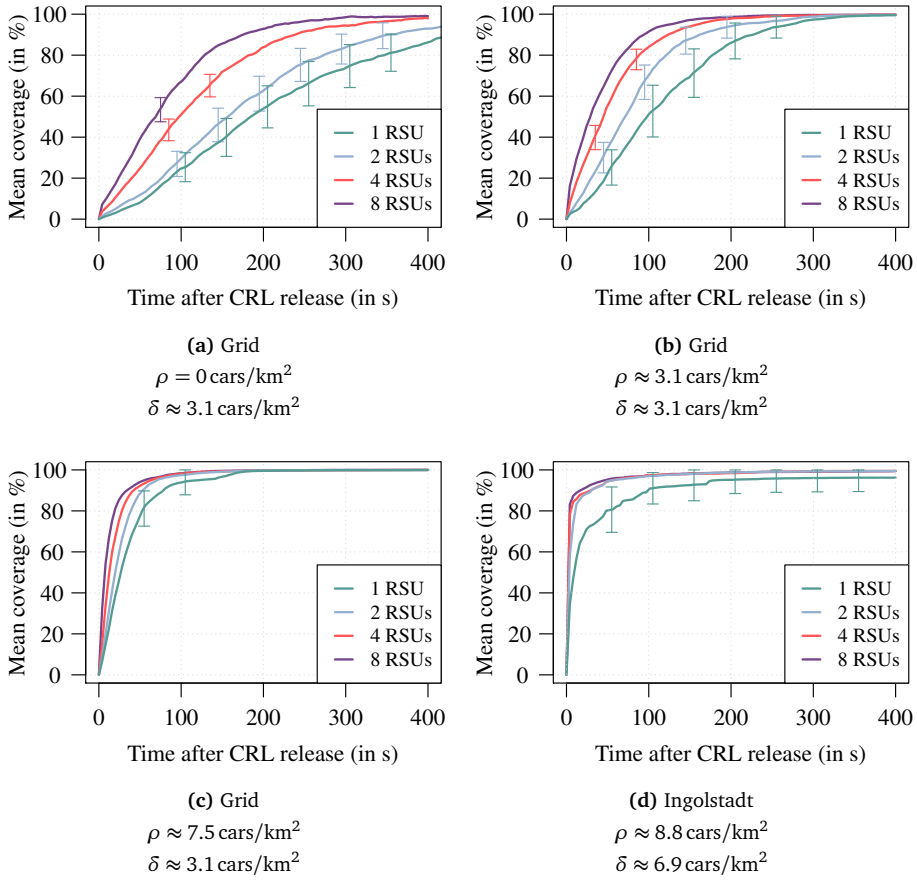
To provide optimal conditions for message dissemination and injection via RSUs, we place them at the exact center of intersections. This way transmission ranges could be maximized, as signal shadowing caused by buildings had a lesser impact.

In order to obtain statistically sound results, we performed 30 (differently seeded) repetitions of each simulation scenario and parameter set.

CRL Dissemination Speed

As a first step, we investigate how the number of parked cars and RSUs affects the CRL dissemination progress. We triggered a new CRL to be released at an arbitrary point in time after the transient phase of the simulation (labeled $t = 0\,s$) and measured how the CRL coverage, that is the percentage of moving and parked vehicles with the most recent version of the CRL, changed as time progresses.

We illustrate the results for the Manhattan grid scenario at a static traffic volume of $\delta \approx 3.1\,cars/km^2$ and different densities of parked vehicles in Figures 4.9a to 4.9c. Without the participation of parked cars (Figure 4.9a), even a high number of RSUs were not sufficient for timely CRL dissemination: reaching just 50 % of vehicles took on the order of minutes. Injecting the CRL using 8 RSUs improves the situation, however it is still considerably slower than when a small number of parked vehicles are combined with 2 RSUs (Figure 4.9b). Using only 7.5 parked vehicles per km^2, distribution can be sped up considerably (Figure 4.9c), further reducing the benefit of RSUs deployment for message dissemination. From this we determine that CRL distribution using parked vehicles is effective and can reduce the necessity of costly RSUs. In this context, equipping some vehicles with cellular devices can even completely replace provider-run RSUs without experiencing slower message dissemination.



(a) Grid
$\rho = 0\,\text{cars/km}^2$
$\delta \approx 3.1\,\text{cars/km}^2$

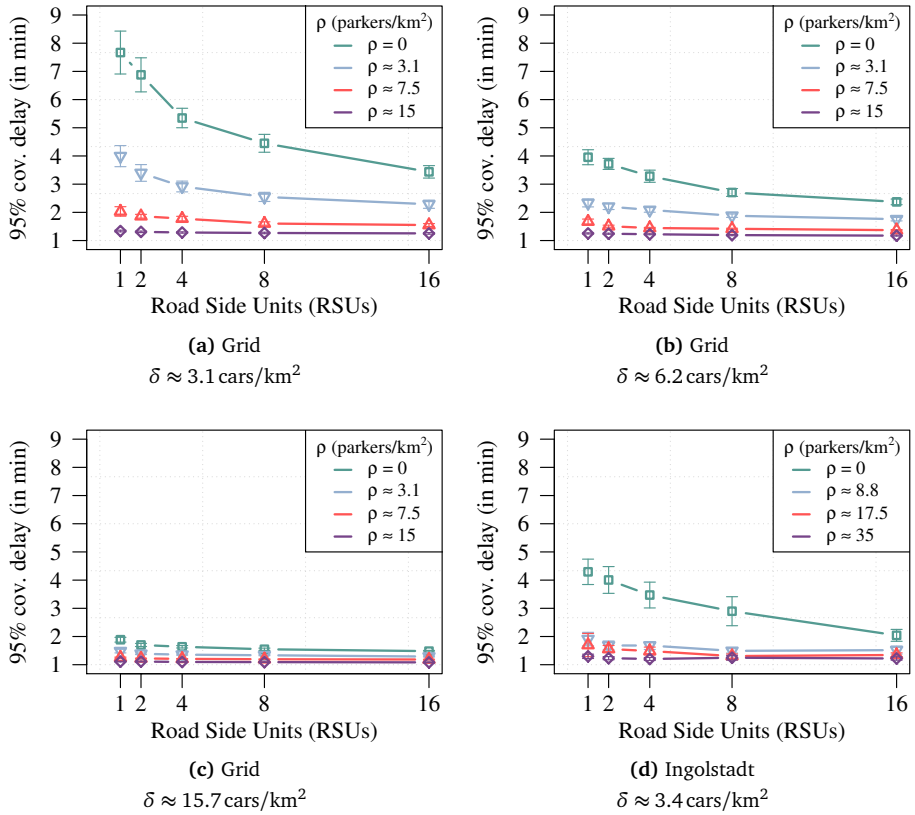
(b) Grid
$\rho \approx 3.1\,\text{cars/km}^2$
$\delta \approx 3.1\,\text{cars/km}^2$

(c) Grid
$\rho \approx 7.5\,\text{cars/km}^2$
$\delta \approx 3.1\,\text{cars/km}^2$

(d) Ingolstadt
$\rho \approx 8.8\,\text{cars/km}^2$
$\delta \approx 6.9\,\text{cars/km}^2$

Figure 4.9 – CRL dissemination coverage (including both moving and parked vehicles) over time in the low traffic density ($\delta \approx 3.1\,\text{cars/km}^2$) grid scenario, depending on the number of RSUs deployed for CRL injection and the number of parked vehicles available for supporting CRL dissemination. Ingolstadt scenario ($\delta \approx 6.9\,\text{cars/km}^2$) for comparison. Error bars show the 95 % confidence interval where greater than 5 %.

For comparison, Figure 4.9d shows CRL distribution in the Ingolstadt scenario. Higher traffic inhomogeneity compared to the grid scenario sometimes led to a very fragmented network, temporarily stopping message dissemination when a vehicle had no other vehicles close-by. This effect was more prominent when using only 1 RSU, however, with the aid of parked vehicles, the CRL could still be disseminated in a timely manner. The plots also reveal that CRL coverage increases smoothly, with no sudden jumps or discontinuities. This motivates us to choose the delay it took to reach 95 % CRL coverage as the primary metric for the following comparisons.



(a) Grid
$\delta \approx 3.1\,\text{cars/km}^2$

(b) Grid
$\delta \approx 6.2\,\text{cars/km}^2$

(c) Grid
$\delta \approx 15.7\,\text{cars/km}^2$

(d) Ingolstadt
$\delta \approx 3.4\,\text{cars/km}^2$

Figure 4.10 – Delay until 95 % coverage was reached in the grid and Ingolstadt scenarios with different traffic densities. Error bars show the 95 % confidence intervals.

Delay Until Reaching 95 % CRL Coverage

Figures 4.10a to 4.10d show how the delay until 95 % of vehicles were up-to-date changed with traffic density, the number of deployed RSUs, and depending on the number of available parked vehicles.

In the Manhattan grid scenario, which is dominated by huge building blocks and therefore strong signal shadowing, the benefit of parked cars to disseminate CRLs is clearly visible when looking at low traffic densities (or penetration rates) (Figure 4.10a). The latency of previous approaches (green line, $\rho = 0$ parked vehicles) is considerably higher even with 8 deployed RSUs. By adding just as many parked vehicles as moving ones, delays can be more than halved, substantially reducing the need for more than 1 RSU. Further increasing the number of parked vehicles (and thereby almost establishing full connectivity of the vehicular network) reduced dissemination times further, informing all vehicles in the 16 km^2 ROI in a little over 1 minute (purple line).

Doubling the traffic volume (Figure 4.10b, $\delta \approx 6.2\,\text{cars/km}^2$) lowered the absolute benefit of parked vehicles support, but still shows a substantial improvement in terms of latency. In our highest density setup (Figure 4.10c, $\delta \approx 15.7\,\text{cars/km}^2$) network connectivity was already at a level that resulted in low latency. Adding parked vehicles helped improve the situation even more, while additional RSUs had only a smaller impact. This means that, as soon as the CRL is injected into the network, no additional effort of the provider has to be undertaken to disseminate the CRL with low delay.

Lastly, we investigated coverage delays in the Ingolstadt scenario with only sparse traffic (Figure 4.10d, $\delta \approx 3.4\,\text{cars/km}^2$). Compared to the grid scenario, delays in the suburban scenario are lower, because fewer vehicles are needed to reach high network connectivity due to the smaller impact of signal shadowing caused by buildings. Nevertheless, we observe that a realistic number of parked vehicles equipped with OBUs reduced the update latency to well below what could be reached with RSUs only – even when deployed in what we believe to be an unreasonably high number for such a suburban area. In general, we note that the saturation point was indeed reached earlier than in the Manhattan grid scenario, but the results clearly show how a vehicular network can benefit from the help of parked vehicles, especially in the early stages of deployment.

4.2.7 Concluding Remarks

The process of revocation is a challenging task in terms of efficiency, distribution, and privacy. SmartRevoc successfully addresses these three issues and therefore provides a solution for certificate revocation in future ITS's.

Efficiency is achieved by moving the computation of the actual revoked certificates to the vehicles, reducing the size of a CRL entry to 36 B for the revocation of a vehicle's pseudonym pool. This way, revoking even a large number of vehicles, e.g., due to a security problem with a certain type of OBU, can be done with low effort compared to traditional CRLs where each entry represents a pseudonym.

Distribution is achieved by the help of parked cars. We showed that parked cars can substantially contribute to improving traffic safety by relaying messages from moving cars and thereby routing around obstacles that would otherwise block radio communication. We therefore propose to disseminate CRLs in an epidemic fashion including parked vehicles, should cellular communication not be available in future on-board systems. CRLs are injected by the help of provider-run RSUs. In terms of coverage, a reasonable number of parked vehicles could outperform a comparatively high number of strategically placed RSUs in both the urban and suburban simulation scenarios. This is especially helpful in the early stages of an ITS and in low traffic

density areas where network connectivity can be low, which would lead to higher dissemination delays.

The last goal, the preservation of privacy, is one of the core features of SmartRevoc. Through the usage of two hash chains, our system preserves full backward location privacy of drivers, i.e., when a vehicle is revoked, linking past pseudonyms is not possible as only future pseudonyms are affected. If pseudonyms should be reused, SmartRevoc can still be applied, however, backward privacy can then not supported, as there are no 'past' pseudonyms. When used with time-slotted pools as recommended in Section 4.1, SmartRevoc can be used to revoke vehicles efficiently and with strong privacy protection without reservation.

4.3 The Scrambler Attack

When adopting existing technology such as IEEE 802.11 from a previous context where privacy was not critical to a new privacy-critical one, all parts of the old system that potentially affect privacy have to be adjusted. This also holds for seemingly privacy-neutral parts of the system, e.g., the encoding or transmitter modules, because one susceptible component can annul privacy protection in other layers. As privacy is always a cross-layer mechanism, it is therefore important to include all layers and components in privacy research.

In this section, which is based on our paper "The Scrambler Attack: A Robust Physical Layer Attack on Location Privacy in Vehicular Networks" [21],[7] we investigate the susceptibility of IEEE 802.11p prototype hardware to physical layer fingerprinting attacks.

The concept of fingerprinting attacks, along with examples that affect IEEE 802.11 hardware, is explained in Section 2.2.3. The general idea is to analyze, e.g., the electromagnetic waveform of a packet to find special distinctive patterns that allow (re-)identification of a certain device. If such an attack is possible, an adversary can be able to link two messages even though they were sent with different pseudonyms. Fortunately, most of the known attacks require expensive hardware such as spectrum analyzers or controlled laboratory environments with little background noise and are therefore unlikely to be feasible for vehicular networks. However, with the use of an Software Defined Radio (SDR) we were able to look into parts of a packet which are usually inaccessible when using Commercial Off-The-Shelf (COTS) hardware, as the transceiver chips do not disclose them. This attack is extremely robust as it does not require stable, controlled environments and can be carried out with affordable hardware ($<$ 400 EUR).

Fingerprinting attacks on IEEE 802.11 are not only relevant in the context of vehicular networks, but also in the general WLAN domain, where privacy protection seems to become more and more of interest. For example, some newer notebooks and smartphones support randomizing their source MAC addresses to prevent tracking when probing for new access points. With the growing market for wearable devices, this trend can be expected to continue. Our main focus will however be vehicular networks even though the scrambler attack may also be feasible on other devices.

The basic approach for our fingerprinting attack is an SDR. SDRs can be divided into two parts: the hardware components and the software components. The electromagnetic spectrum is sampled by general-purpose hardware, and, contrary

[7] This publication was written in collaboration with the Distributed Embedded Systems Group of the University of Paderborn, Germany. Personal contributions include the general approach of using Software Defined Radios (SDRs) for fingerprinting attacks and privacy-related research, and also the simulation study to quantify the impact of the attack. The SDR test bed, analysis and reverse engineering of the scrambler seeds were provided by the co-authors.

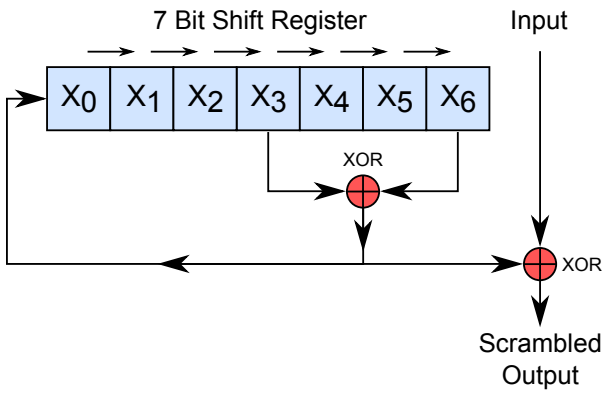


Figure 4.11 – Schematic overview of the IEEE 802.11 scrambling algorithm.

to traditional COTS transceivers, the actual processing and decoding of frames is done using a programmable component such as a Field-Programmable Gate Array (FPGA), a Digital Signal Processor (DSP), or a normal CPU. Bloessl et. al developed an IEEE 802.11a/g/p receiver for GNU Radio that allows analysis and decoding of IEEE 802.11p packets sent by prototype OBUs [19]. As a hardware component we used the N210 from Ettus Research.[8] This setup can also be used to transmit IEEE 802.11p compatible packets [20], however, for this work we only needed to receive messages, making this attack passive and undetectable.

4.3.1 Scrambling as a Vulnerability

In IEEE 802.11a/g/p, before a frame is sent, the input data needs to be encoded to Orthogonal Frequency Division Multiplexing (OFDM) symbols. This process includes several steps, such as puncturing, interleaving, and modulation, however, before these are carried out, the input data is scrambled. Despite its name, scrambling is not a security feature but a performance feature for the physical layer. Its purpose is to encode possibly correlated input bits of the payload to seemingly random output bits of the same length. Thereby long sequences of same-value bits and other correlated sequences are replaced, maximizing the entropy and thus information content.

In OFDM systems, the scrambler has another advantage, as without it certain bit patterns map to disadvantageous waveforms with a very high Peak to Average Power Ratio (PAPR) [152]. Sending the same data payload twice will not lead to sending the same OFDM symbols as the scrambler is seeded with a pseudorandom sequence per frame which leads to different OFDM symbols. Without a scrambler, the same payload would always generate the same physical waveform, possibly causing specific payloads to systematically experience higher packet error rates.

[8] `http://www.ettus.com/`

An IEEE 802.11 scrambler is shown in Figure 4.11: The 7-bit linear feedback shift register is filled with a pseudorandom sequence before scrambling starts. Next, each bit of the input data is xor'd with bits X_3 and X_6 of the register. Before the next bit of the input data is processed, the first bit X_0 of the register will be set to the xor value of bit X_3 and bit X_6, and the register is shifted to the right. As can be seen, the content of the input data does not affect the register and only the length of the input determines how often the register is shifted. How the register changes during the scrambling process only depends on its initial state before the process began. We refer to this initial state as the scrambler seed.

To maximize performance, and also for privacy reasons, the IEEE 802.11 standard [121] states the scrambler should be seeded randomly with a nonzero value for every frame. More precisely, it states: "When transmitting, the initial state of the scrambler shall be set to a pseudorandom nonzero state" [121, Section 18.3.5.5]. To descramble a packet, the initial state is required at the receiver side. For this, the seven least significant bits of the service field are set to zeros prior to scrambling. This allows receivers to reconstruct the initial scrambler state and consecutively descramble the remainder of the packet.

Attack Vector

The initial state of the scrambler should be set to a pseudorandom state using a Pseudorandom Number Generator (PRNG). If the used pseudorandom number sequence can be guessed or even predicted by an adversary, they could be able to relate two transmissions to the same OBU and thereby link two messages. This can completely annul the effect of pseudonym changing strategies and bypass all higher layer privacy mechanisms.

Using an SDR, we investigate how current IEEE 802.11p hardware generates scrambler seeds and whether they are susceptible to fingerprinting attacks. Since vehicular networks are not yet deployed, we looked at prototype IEEE 802.11p devices and adapted WLAN cards commonly used in field operational tests. We examined the **Cohda Wireless MK2**[9] prototype (as used in major FOTs in Europe and North America) and the **Unex DCMA-86P2**[10] miniPCI card (as used in the Grand Cooperative Driving Challenge [184]). Using the SDR, we analyzed and reverse engineered the scrambler PRNG for both devices.

Cohda Wireless is one of the leading suppliers of IEEE 802.11p prototypes. The MK2 is an ARM-based PC with an IEEE 802.11p radio implemented on an FPGA that ships with all the firmware (rev. 4.0.14615) and software of a complete IEEE WAVE-enabled OBU or RSU. We found that the Cohda device uses a freewheeling scrambler,

[9] http://www.cohdawireless.com/
[10] http://www.unex.com.tw/

meaning the scrambler register is not re-initialized with a random seed prior to scrambling, but the value the register had after scrambling the last packet is the seed for the next packet. An adversary can therefore pre-compute the next seed using the overheard scrambler seed and the length of the current packet. Identifying the next packet from the same on-board device can then be simply done by comparing this pre-computed value with those in received packets. As a side effect, using a freewheeling scrambler also means that if the packet size is a multiple of the cycle length of the scrambler, the seed does not change at all.

For the Unex card, our findings were similarly worrisome. The card is based on the Atheros AR5413 transceiver chip, which can be regarded as very representative of Commercial Off-The-Shelf (COTS) hardware. We found that the Unex card simply increases the scrambler seed by one after each frame, making it trivial to relate two consecutive messages to the same device. In order to make sure that this is always the case, we tested the card in different settings, such as in monitor and ad-hoc mode or with cross-traffic overheard by the device. None of this affected the PRNG.

Obviously, both algorithms allow for a trivial re-identification of consecutive frames from one card when overheard by an adversary. In this context, the changing of pseudonyms and other re-identification protection mechanisms are rendered useless. Consider the following case for the Unex card: The adversary overhears four consecutive frames $\binom{P}{n}$ with pseudonym P and scrambler seed n: $\binom{A}{15}, \binom{A}{16}, \binom{B}{17}, \binom{B}{18}$. It is trivial to guess that (with high probability) pseudonyms A and B are the same entity. The same attack is possible for the Cohda device with the additional consideration of the message length. However, this does not increase the complexity for a listening attacker as the length of a packet can be easily derived by receiving it.

This fingerprinting technique is not only effective when an adversary overhears all communication, but also when only parts of the network are covered by adversary receivers. If the vehicle leaves one part of the network that was covered by an adversary and enters another covered area some time later, the adversary can predict the expected scrambler seed of the vehicles by estimating the amount of messages the vehicle sent in the meantime. This attack becomes especially feasible when vehicles use static (or a discrete set of) beaconing frequencies, and in the case of Cohda devices, use messages of the same length. Both beacon frequencies and message length can potentially be guessed by an adversary in IEEE WAVE and ETSI ITS-G5. We will show the feasibility of this approach in the simulation study in Section 4.3.2.

Countermeasures

The IEEE 802.11 standard's requirement for the scrambler seed is insufficient in terms of privacy as it does not clearly state how the pseudorandom sequence should be generated. From a physical layer performance perspective it is sufficient to

predictably change the scrambler seed from frame to frame. This seems to have caused most vendors to employ very simple algorithms, not considering possible implications on location privacy.

The most straightforward solution is the use of a cryptographic PRNG, possibly seeded by the large number of entropy sources in a vehicle (e.g., time when the vehicle was started or values from fuel level and tire pressure sensors). Using a constant scrambler value for *all* IEEE 802.11p devices in the network would prevent the presented fingerprinting attack, however, it would likely degrade network communication performance as same inputs would always be mapped to the same OFDM symbols [152].

Repairing the scrambler for the Cohda prototype platform is not complicated as it does not rely on a transceiver chip but implements all logic on reconfigurable FPGAs. Therefore, it should be possible to address the scrambling algorithm with a firmware update of the prototype.

For COTS hardware, the situation is different. Because vendors do not provide detailed information about their hardware design, it is difficult to tell where certain functionality is implemented and if a solution can be achieved with a driver or firmware update. In the worst case, the scrambling algorithm is implemented in hardware and hence cannot be changed, but instead the chip would have to be replaced.

4.3.2 Evaluation of Impact

In order to study the quantitative impact of scrambler-based fingerprinting we carried out an extensive set of simulations using our privacy simulator for the Veins framework (Section 3.1). We implemented the scrambler algorithms from the Cohda and Unex devices and extended the tracking algorithm to also make use of them. During the simulation, the adversary tries to learn what kind of device a certain vehicle is using. This can be achieved by first tracking a vehicle without consideration of scrambler seed values as explained in Section 3.1.1. After a track contains three observations with a high enough probability, the adversary can analyze the scrambler seeds of these observations and check whether they follow a known pattern, i.e., that of the investigated Cohda and Unex scramblers. If a device type could be identified, the track is marked with that information along with a confidence value which rises with the number of observations made that follow the expected scrambler pattern. In the next gating phase, all observations that would not continue the track's scrambler are assigned a considerably higher statistical distance (based on the confidence) so observations that do follow the pattern are preferably used for the next track-to-observation assignment.

Parameter	Value
Adversary model	External, local, passive, adaptive, scenario-specific
Privacy domain	Location privacy
Privacy property	Unlinkability
Data source	Observable information
Metrics	Adversary's success rate
Scenario	Intersection, blind spot freeway (800 m gap)
Technology	IEEE WAVE
Beacon frequency	1 Hz
Pseudonym changing	New for every message
No. of vehicles	25-400
Perc. of Cohda OBUs	33 % to 50 %
Perc. of Unex OBUs	33 % to 50 %
Perc. of non-affected OBUs	0 % to 33 %

Table 4.5 – Simulation setup and parameters for the evaluation of the scrambler attack.

If a track could not be continued in the last interval, e.g., due to packet loss or the vehicle leaving the covered area, it is not deleted. The adversary tries to continue this track in the next observation period by extrapolating the scrambler seed values using the amount of potentially missed packets and their size. For example, knowing the vehicle emits beacons with a frequency of 1 Hz and the last known observation from this vehicle was four seconds ago, the next scrambler seed value has to be increased by 4 in the case of the Unex card. This way, vehicles can possibly be tracked through blind spots not covered by adversary radio receivers.

Simulation Setup

Analogously to Section 3.1.4, we investigate two challenging scenarios in which the adversary deployed radio receivers along the road to track vehicles: the intersection scenario and the blind spot freeway scenario (Section 3.1.3).

In the intersection scenario, the theoretical transmission range allowed the attacker to receive packets from vehicles approaching and leaving the intersection and on the intersection itself. A vehicle is considered tracked when it was possible to fully recreate the path of a vehicle over the intersection from receiving the first packet until receiving the last packet. Note that, as in our simulation the attacker is not omniscient but uses a radio receiver, they can experience packet loss and therefore lose track of a vehicle or associate an overheard beacon with the wrong vehicle.

In the investigated freeway setting, the attacker is not able to fully overhear all messages in the scenario, but placed two receivers along a 3 km stretch of freeway with a blind spot of 800 m between them. Here, a vehicle is considered tracked if it was possible to track its path from entering the transmission range of the first receiver to leaving the transmission range of the second one.

In both scenarios the number of vehicles was kept constant throughout the simulation: For every vehicle that left the scenario a new one of a random preset type (e.g., truck, van, or car) with a new, random route was inserted. For each scenario we investigated two different configurations: In the first, each vehicle was randomly assigned a Cohda or a Unex device so the adversary could always exploit the weak scrambler PRNG. In the second run, we additionally introduced a good scrambling device which was not susceptible to fingerprinting so that the distributions of Unex, Cohda, and this device were an equal 33 %.

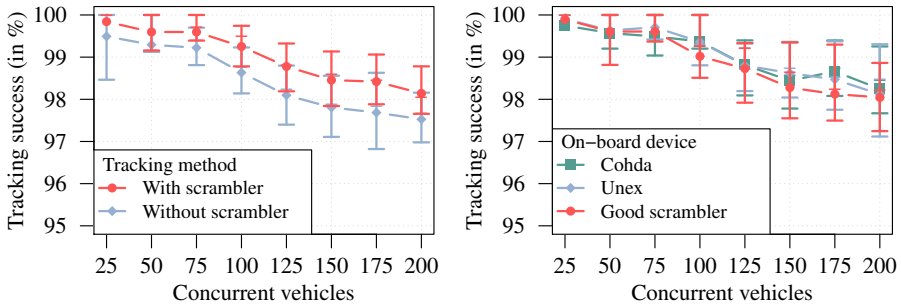
Because the scrambler attack is able to annul privacy protection mechanisms on the MAC and higher layers, we investigated a best-case scenario for privacy: Vehicles used a new pseudonym for each message, making it impossible for the adversary to link messages based on any upper layer identifier. Also, vehicles emitted beacons with a frequency of only 1 Hz, which is the lowest possible beacon frequency according to the ETSI family of standards [84], further complicating tracking. The physical layer was simulated using two-ray-interference path loss with a transmission power of 20 mW, leading to a theoretical transmission range of about 600 m. All relevant simulation parameters are given in Table 4.5.

Tracking Success

Figure 4.12 shows our results for the intersection scenario. We plot the mean tracking success over all simulation runs. The error bars represent the 25 % and 75 % quantiles, the line in between the whiskers shows the median.

Similar to the previous results (Section 3.1.4), the adversary was able to track almost every vehicle passing the intersection (Figure 4.12a). Looking into the cases where tracking failed, we found that it was caused by packet loss due to coincidental synchronization of nodes and very sharp turn maneuvers in SUMO. Even though tracking success is already an alarmingly high level without physical layer fingerprinting (blue line), additionally using the scrambler seed as input can further increase the adversary's success (red line).

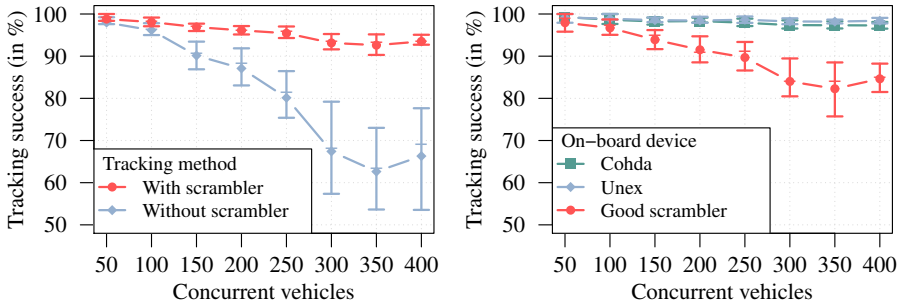
Introducing the non-susceptible device, we witnessed almost no difference in the tracking success (Figure 4.12b). As expected, successfully tracking vehicles over the busy intersection did not rely on the scrambler seed. Our results suggest that vehicles with exploitable scramblers are easier to track, however, the overlapping



(a) Intersection scenario **(b)** Intersection scenario

Figure 4.12 – Impact of the Scrambler Attack in the intersection scenario. Plotted are the mean values of the simulation runs, errors bars show the 25 % and 75 % quantiles. The line in-between indicates the median of the distribution.

error bars do not allow for statements regarding the significance of the results in the intersection scenario.



(a) Freeway gap scenario **(b)** Freeway gap scenario

Figure 4.13 – Impact of the scrambler attack in the blind spot freeway scenario. Error bars show quantiles as in Figure 4.12.

The results for the freeway scenario are shown in Figure 4.13. First, we examined the difference between either all vehicles using an exploitable scrambler, or none (Figure 4.13a). The blind spot between the two receivers made it considerably harder (blue line, diamond markers) for the attacker to track vehicles. Dynamics in the mobility of vehicles such as lane changing, overtaking, or varying velocities led to wrong associations of beacons to vehicles on the attacker side. We observed that the mobility generated by SUMO seemed to be more dynamic than one would expect; we believe that to confuse an attacker in real life assuming 'normal' mobility,

the gap between the receivers would likely have to be wider. Congestion setting in at the highest vehicle density caused a slight increase in tracking probability, due to fewer lane changes and passing maneuvers.

When the attacker used additional scrambler information to track vehicles, the situation completely changed (red line, circle markers): We observe that the gap between the two attacker radios only marginally influenced the capability to track vehicles. Approximating the number of beacons presumably sent by a vehicle while driving in the uncovered section of the freeway, the attacker is able to estimate a number of possible scrambler values. Using this technique, we obtained tracking probabilities of over 95 %, almost reaching the level of the fully covered intersection scenario. This shows that, even on a busy freeway with interrupted radio coverage, the scrambler attack allows an attacker to effectively circumvent any higher layer privacy protection and track a large portion of vehicles.

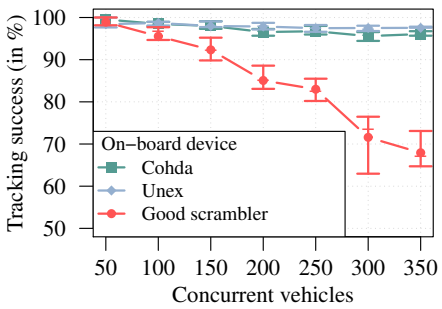


Figure 4.14 – Results for the freeway with one additional lane per direction, differentiated by the type of OBU used.

To fully illustrate the crucial requirement of unpredictable scrambler values we further analyzed the results for the freeway scenario, showing the tracking probability differentiated by the type of IEEE 802.11p radio (Figure 4.13b). As can be seen, location privacy cannot be achieved using a predictable scrambler – the attacker was able to track almost every vehicle using the Cohda or the Unex radio. Even the vehicles using a good scrambler (red line, circle markers) suffer from the now smaller number of vehicles possibly confusing an attacker. This is a case of interdependent privacy where actions (or in this case, devices) of some users affect the privacy of others [245]. Their probability of being tracked is considerably higher than it was when no vehicle used exploitable devices (Figure 4.13a, blue line). This further underlines the necessity to address this problem and not allow for a circumvention of higher layer privacy measures.

As a last step, we wanted to investigate the effect of mobility more. For that we increased the lanes per direction on the freeway to four, allowing easier overtaking. We observe that, while the level of privacy enjoyed by users with predictable

scramblers only marginally improved (Figure 4.14, blue and green line), vehicles with non-predictable scramblers are now harder to track, as vehicles can move more freely due to the additional lane, which complicates tracking. This again emphasizes the importance of synchronized pseudonym changing as proposed in Section 4.1: Privacy protection is most effective when all vehicles currently not in the transmission range of an eavesdropping attacker change pseudonyms simultaneously to maximize the adversary's confusion and complicate re-identification.

4.3.3 Concluding Remarks

Adapting and developing technology for privacy-preserving operation is a challenging task. We illustrated that even seemingly unaffected components of the system can be critical when it comes to the preservation of privacy. The scrambler of IEEE 802.11 devices is such a component: We showed that, when not implemented in a privacy-preserving way, it can completely annul the effect of many other privacy protection mechanisms in higher layers such as the changing of pseudonyms – the most important PET in vehicular networks. For current prototype hardware, we were able to identify a problem within the PRNG that did not generate unpredictable random numbers, enabling adversaries to correlate the scrambler seeds of different frames to link and relate them to the same device. In contrast to existing lower layer attacks, this attack is extremely robust, as it makes use of data rather than signal characteristics.

We used our privacy simulation framework to give detailed insights on the quantitative impact in terms of tracking probabilities. Investigating both intersection and freeway scenarios we showed that, even when vehicles traveled through sections where an adversary was not able to overhear messages, it was possible to reliably track drivers. We also observed effects of interdependent privacy, where the privacy of all drivers was reduced by the presence of those with exploitable scramblers. The results highlight the importance of using cryptographic PRNGs in IEEE 802.11p scramblers, not to increase the performance of the system, but to preserve the location privacy of drivers.

The simulative evaluation of the scrambler attack illustrates how effective our privacy simulation framework is. By extending the gating process and changing the statistical distance between tracks and observations based on the estimated scrambler values, we were able to assess the privacy loss caused by this physical layer fingerprinting attack. Our framework can be used to study all kinds of data association and correlation attacks by simply adjusting or extending the subcomponents described throughout Section 3.1, making it a powerful tool for the evaluation of privacy-enhancing technologies in vehicular networks.

Chapter 5

Conclusion

Privacy is important. Although the main motivation of IVC is to improve traffic safety, privacy protection must not be neglected in future ITS's. These systems will become a reality in the near future, as standards in North America and Europe are being finalized. Privacy protection must be considered an inherent part from the very beginning, as retrofitting privacy into existing systems is bound to fail.

Privacy is still a nebulous concept that is hard to grasp and to put a number on. In this thesis, we alleviated this problem by highlighting the concrete privacy risks in vehicular networks. We extended the taxonomy and identified shortcomings of current privacy research that need to be overcome in order to make privacy concepts easier to understand. We showed how this situation can be improved and how simulations can be used to evaluate location privacy in vehicular networks.

We developed a privacy simulation framework for our Veins simulator that makes use of a state-of-the-art tracking system. To further allow for the detailed analysis of packet-level privacy protection and attacks, we also implemented models for IEEE WAVE and ETSI ITS-G5. We identified deficits in both standards: IEEE WAVE exhibited synchronization effects with resulting packet loss, while the ETSI system introduced channel load oscillation by the use of DCC and thereby suboptimal channel usage. We believe that our research helped raise awareness of these issues and contributed to improving vehicular network technology in general.

We used our privacy simulation framework to develop and evaluate different privacy-enhancing technologies:

SlotSwap introduces the concept of exchanging pseudonyms between vehicles based on time-slotted pseudonym pools. While pseudonym exchange requires tamper-proof hardware and protocols and should therefore only be considered with caution, the use of non-overlapping time-slotted pools and the resulting synchronized pseudonym change of all vehicles in the network has decisive benefits: It eliminates the possibility of Sybil attacks as each car has only one valid pseudonym

for any point in time, and it increases the privacy gained by pseudonym changes as such a coordinated change of identifiers will considerably complicate tracking of vehicles.

Time-slotted pools have one further advantage: They can be revoked efficiently. We presented SmartRevoc, an efficient, backward privacy-preserving certificate revocation scheme that also utilizes parked vehicles for the distribution of the CRL. To motivate the inclusion of parked vehicles, we showed how they can considerably improve traffic safety by relaying beacons from moving vehicles and thereby routing around obstacles such as buildings. SmartRevoc uses only 36 B per revoked vehicle and completely shifts the computational effort to the vehicles. By using two hash chains it ensures backward privacy of revoked vehicles, as only future pseudonyms are affected. We showed the effectiveness of distributing a CRL using parked vehicles in both urban and suburban environments, outperforming even a large number of optimally placed roadside units.

Lastly, we showed that, even when well thought-out PETs are deployed, a single component may compromise the privacy offered by the system. We identified a weakness in the implementation of the IEEE 802.11 scrambler in two common IEEE 802.11p prototype devices that allowed correlation of scrambler seed values for different packets. This allows an adversary to link messages even when they are not able to completely overhear all parts of the road network. Using our simulation we showed the feasibility and the effectiveness of the attack and illustrated the principle that each subcomponent of a cross-layer privacy system needs to be aware of the context it is used in.

A Proposal for Privacy Protection in Vehicular Networks

The lessons learned throughout this thesis allow us to derive an overview of the potentials and limitations of privacy protection in vehicular networks.

One of the biggest challenges when changing pseudonyms is the privacy–safety trade-off: Confusing a tracking adversary can also mean confusing the OBUs of other vehicles. Never must privacy protection in vehicular networks cause a traffic accident or, even worse, injury or death. The fact that many people will likely value safety much higher than privacy in a potentially critical situation has to be accounted for when developing and deploying pseudonym changing strategies. Our results showed that it is difficult (or almost impossible) to prevent an eavesdropping adversary from tracking, even with beacon frequencies as low as 1 Hz and new pseudonyms for each message. In these situations, privacy mechanisms can be potentially dangerous in terms of traffic safety while their benefit for privacy is at best negligible.

We therefore propose to consider sacrificing privacy to nearby vehicles (and eavesdropping adversaries) and instead fully concentrate on privacy protection

when an adversary is not eavesdropping. The time in which a vehicle's transmissions cannot be overheard by an adversary must therefore be used effectively to increase the level of location privacy. To confuse an attacker, not only must a vehicle change its own pseudonym before re-entering an area covered by an adversary, but, ideally, many vehicles will have done the same to maximize confusion for the adversary. The use of time-slotted pseudonym pools achieves exactly that by making all vehicles change their address at the same time, complicating the re-identification of *any* vehicle not currently in the vicinity of an adversary. The level of privacy then depends on the time slot length, which has to be chosen carefully to balance location privacy and certificate overhead, as smaller time slots lead to larger pseudonym pools and higher load at the CA.

To prevent time-slotted pseudonym pools from having a negative impact on traffic safety (as vehicles can no longer postpone pseudonym changes) vehicles could announce their pseudonym change by temporarily adding their old pseudonym identifier to new packets. Depending on the requirements of the running safety applications, the duration in which old pseudonym identifiers are included in new messages can vary but should typically not exceed a few seconds. Even though this makes tracking trivial for eavesdropping adversaries and also virtually increases adversaries' coverage (as pseudonym changes just prior to entering a covered area can now be tracked), it completely eliminates every negative effect pseudonym changing can have on safety applications. We believe that this way future ITS's can achieve a good balance between traffic safety and privacy protection, as it further allows the installment of other privacy-enhancing technologies such as the presented efficient and privacy-preserving revocation system.

Future Research

The developed privacy simulation framework offers the possibility for a detailed analysis of privacy protection in vehicular networks and opens several future research directions: For the parametrization of the time slot length it would be beneficial to fully understand the influence of adversary network coverage on the level of privacy. Particularly in urban and suburban scenarios where (possibly collaborating) adversaries can easily set up new access points (or use existing ones) to cover large areas of the network, the slot length must be chosen in a way so that the location privacy of drivers is still sufficiently protected.

Another open research direction is the analysis of the impact of possibly identifying information included in transmitted messages. Studying the consequences of application data on drivers' location privacy should be a requirement for every future IVC application. If messages sent by applications include identifying information they could compromise any privacy protection mechanism. For example, an IVC-

based traffic information system that periodically reports similar traffic situations can be exploited by an adversary to re-identify vehicles even after an unobserved pseudonym change. The only way to guarantee privacy protection in vehicular networks is to analyse every piece of information sent by the vehicles and investigate whether this information constitutes a danger for the location privacy of drivers.

In terms of vehicular network simulation in general, it can be said that, although the quality and availability of mobility models and network models has steadily increased over the last decade, there is still potential for improvement. Realistic mobility is the basis for the meaningful simulation of vehicular networks and should therefore receive further attention. This does not only include the microscopic mobility of vehicles, but also ranges from the inclusion of pedestrians and cyclists to adopting new macroscopic mobility patterns, possibly induced by large-scale car sharing platforms or autonomous vehicles.

On the networking side, simulations can be improved by a more detailed modeling of the wireless channel, as many assumptions of today's simulations are optimistic. For example, including non-omnidirectional antennas could introduce interesting effects previously neglected in simulations. Receiving a packet from another node does then no longer mean that this node can also be potentially reached. Also the simulated network layers and protocols need to represent the real world as accurately as possible. Changes in the upcoming standards and technology have to be accounted for and simulation models have to be continuously maintained. By developing and maintaining models for our Veins simulation framework we provide a powerful simulation tool to the vehicular network research community and try to contribute to shaping the road traffic of tomorrow.

Abbreviations

AIFS	Arbitration Interframe Spacing
AIFSN	Arbitration Interframe Spacing Number
API	Application Programming Interface
ASS	Anonymity Set Size
AWGN	Additive White Gaussian Noise
BSM	Basic Safety Message
CA	Certificate Authority
CAM	Cooperative Awareness Message
CCA	Clear Channel Assessment
CCH	Control Channel
COTS	Commercial Off-The-Shelf
CPRNG	Cryptographic Pseudorandom Number Generator
CPU	Central Processing Unit
CRL	Certificate Revocation List
DCC	Decentralized Congestion Control
DENM	Decentralized Environmental Notification Message
DLR	German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt)
DSC	DCC Sensitivity Control
DSP	Digital Signal Processor
DSRC	Dedicated Short-Range Communication
DTN	Delay Tolerant Network
ECDF	Empirical Cumulative Distribution Function
ECU	Electronic Control Unit
EDCA	Enhanced Distributed Channel Access
EDCAF	Enhanced Distributed Channel Access Function
FIFO	First In, First Out
FOT	Field Operational Test
FPGA	Field-Programmable Gate Array
GLOSA	Green Light Optimal Speed Advisory

GPS	Global Positioning System
IDM	Intelligent Driver Model
ITS	Intelligent Transportation System
IVC	Inter-Vehicle Communication
JPDA	Joint Probabilistic Data Association
LLC	Logical Link Control
LTE	Long-Term Evolution
MAC	Medium Access Control Layer
MANET	Mobile Ad-Hoc Network
MOBIL	Minimizing Overall Braking Induced by Lane change
NIC	Network Interface Controller
OBU	On-Board Unit
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak to Average Power Ratio
PET	Privacy-Enhancing Technology
PHY	Physical Layer
PKI	Public Key Infrastructure
PRNG	Pseudorandom Number Generator
QoS	Quality of Service
QPSK	Quadrature Phase-shift Keying
ROI	Region of Interest
RSU	Roadside Unit
SCH	Service Channel
SDR	Software Defined Radio
SIFS	Short Interframe Spacing
SINR	Signal-to-Interference-plus-Noise Ratio
SPAT	Signal Phase and Timing Message
SSU	Stationary Support Unit
SUMO	Simulation of Urban Mobility
TAC	Transmit Access Control
TDC	Transmit Data Rate Control
TOPO	Road Topology Message
TPC	Transmit Power Control
TraCI	Traffic Control Interface
TRC	Transmit Rate Control
TTL	Time-to-Live
UML	Unified Modeling Language
UMTS	Universal Mobile Telecommunications System
VANET	Vehicular Ad-Hoc Network
WSMP	Wave Short Message Protocol

List of Figures

List of Tables

Bibliography

[1] N. AN, T. GAUGEL, and H. HARTENSTEIN, "VANET: Is 95% Probability of Packet Reception Safe?" in *11th International Conference on ITS Telecommunications (ITST 2011)*. Saint Petersburg, Russia: IEEE, August 2011, pp. 113–119.

[2] N. AN, M. MAILE, D. JIANG, J. MITTAG, and H. HARTENSTEIN, "Balancing the Requirements for a Zero False Positive/Negative Forward Collision Warnings," in *10th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2013)*. Banff, Canada: IEEE, March 2013, pp. 191–195.

[3] N. AN, J. MITTAG, and H. HARTENSTEIN, "Designing Fail-Safe and Traffic Efficient 802.11p-based Rear-End Collision Avoidance," in *6th IEEE Vehicular Networking Conference (VNC 2014)*. Paderborn, Germany: IEEE, December 2014, pp. 9–16.

[4] M. ASEFI, J. MARK, and X. SHEN, "A Cross-Layer Path Selection Scheme for Video Streaming over Vehicular Ad-Hoc Networks," in *72nd IEEE Vehicular Technology Conference (VTC2010-Fall)*. Ottawa: IEEE, September 2010, pp. 1–5.

[5] V. BADESCU, "Dynamic model of a complex system including PV cells, electric battery, electrical motor and water pump," *Elsevier Energy*, vol. 28, no. 12, pp. 1165–1181, October 2003.

[6] F. BAI and B. KRISHNAMACHARI, "Exploiting the Wisdom of the Crowd: Localized, Distributed Information-Centric VANETs," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 138–146, May 2010.

[7] F. BAI and B. KRISHNAMACHARI, "Spatio-Temporal Variations of Vehicle Traffic in VANETs: Facts and Implications," in *6th ACM International Workshop on Vehicular Inter-Networking (VANET 2009)*. Beijing, China: ACM, September 2009, pp. 43–52.

[8] M. BALMER, K. MEISTER, M. RIESER, K. NAGEL, and K. AXHAUSEN, "Agent-based Simulation of Travel Demand: Structure and Computational Perfor-

mance of MATSim-T," in *2nd TRB Conference on Innovations in Travel Modeling*, Portland, OR, USA, June 2008, pp. 1–37.

[9] G. BANSAL, J. KENNEY, and C. ROHRS, "LIMERIC: A Linear Adaptive Message Rate Algorithm for DSRC Congestion Control," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4182–4197, November 2013.

[10] R. BARR, Z. J. HAAS, and R. VAN RENESSE, "JiST: an efficient approach to simulation using virtual machines," *Software: Practice and Experience*, vol. 35, no. 6, pp. 539–576, 2005.

[11] R. BAUMANN, S. HEIMLICHER, and M. MAY, "Towards Realistic Mobility Models for Vehicular Ad-hoc Networks," in *26th IEEE Conference on Computer Communications (INFOCOM 2007): IEEE Workshop on Mobile Networking for Vehicular Environments (MOVE 2007)*. Anchorage, AK: IEEE, May 2007, pp. 73–78.

[12] R. BAUMANN, F. LEGENDRE, and P. SOMMER, "Generic Mobility Simulation Framework (GMSF)," in *9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc 2008), 1st ACM SIGMOBILE Workshop on Mobility Models (MobilityModels'08)*. Hong Kong, China: ACM, May 2008, pp. 49–56.

[13] BBC, "Telefonica hopes 'big data' arm will revive fortunes," 9 October 2012. [Online]. Available: http://www.bbc.co.uk/news/technology-19882647

[14] N. BEL GEDDES, *Magic Motorways*. Random House, 1940.

[15] A. BENSLIMANE, "Optimized Dissemination of Alarm Messages in Vehicular Ad-Hoc Networks (VANET)," in *IEEE International Conference on High Speed Networks and Multimedia Communications (HSNMC 2004)*, vol. LNCS 3079. Toulouse, France: Springer, June 2004, pp. 655–666.

[16] C. BEWERMEYER, R. BERNDT, S. SCHELLENBERG, R. GERMAN, and D. ECKHOFF, "Poster: cOSMetic - Towards Reliable OSM to SUMO Network Conversion," in *7th IEEE Vehicular Networking Conference (VNC 2015)*. Kyoto, Japan: IEEE, December 2015, pp. 155–156.

[17] N. BISSMEYER, C. STRESING, and K. M. BAYAROU, "Intrusion Detection in Vanets Through Verification of Vehicle Movement Data," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*. Jersey City, NJ: IEEE, December 2010, pp. 166–173.

[18] S. BLACKMAN and R. POPOLI, *Design and Analysis of Modern Tracking Systems*. Artech House Boston, 1999.

[19] B. BLOESSL, M. SEGATA, C. SOMMER, and F. DRESSLER, "An IEEE 802.11a/g/p OFDM Receiver for GNU Radio," in *ACM SIGCOMM 2013, 2nd ACM SIGCOMM Workshop of Software Radio Implementation Forum (SRIF 2013)*. Hong Kong, China: ACM, August 2013, pp. 9–16.

[20] B. BLOESSL, M. SEGATA, C. SOMMER, and F. DRESSLER, "Towards an Open Source IEEE 802.11p Stack: A Full SDR-based Transceiver in GNURadio," in *5th IEEE Vehicular Networking Conference (VNC 2013)*. Boston, MA: IEEE, December 2013, pp. 143–149.

[21] B. BLOESSL, C. SOMMER, F. DRESSLER, and D. ECKHOFF, "The Scrambler Attack: A Robust Physical Layer Attack on Location Privacy in Vehicular Networks," in *4th IEEE International Conference on Computing, Networking and Communications (ICNC 2015), CNC Workshop*. Anaheim, CA: IEEE, February 2015, pp. 395–400.

[22] M. BOBAN, T. VINHOSA, J. BARROS, M. FERREIRA, and O. K. TONGUZ, "Impact of Vehicles as Obstacles in Vehicular Networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 29, no. 1, pp. 15–28, January 2011.

[23] M. BOBAN, W. VIRIYASITAVAT, and O. TONGUZ, "Modeling Vehicle-to-Vehicle Line of Sight Channels and its Impact on Application-Layer Performance," in *10th ACM International Workshop on Vehicular Internetworking (VANET 2013)*. Taipei, Taiwan: ACM, June 2013, pp. 91–93.

[24] R. BODENHEIMER, A. BRAUER, D. ECKHOFF, and R. GERMAN, "Enabling GLOSA for Adaptive Traffic Lights," in *6th IEEE Vehicular Networking Conference (VNC 2014)*. Paderborn, Germany: IEEE, December 2014, pp. 167–174.

[25] R. BODENHEIMER, D. ECKHOFF, and R. GERMAN, "GLOSA for Adaptive Traffic Lights: Methods and Evaluation," in *9th International Workshop on Communication Technologies for Vehicles (Nets4Cars-2015-Fall)*. Munich, Germany: IEEE, October 2015, pp. 320–328.

[26] R. BRAUN, F. BUSCH, C. KEMPER, R. HILDEBRANDT, F. WEICHENMEIER, C. MENIG, I. PAULUS, and R. PRESSLEIN-LEHLE, "TRAVOLUTION – Netzweite Optimierung der Lichtsignalsteuerung und LSA-Fahrzeug-Kommunikation," *Strassenverkehrstechnik*, vol. 53, pp. 365–374, June 2009.

[27] V. BRIK, S. BANERJEE, M. GRUTESER, and S. OH, "Wireless Device Identification with Radiometric Signatures," in *13th ACM International Conference on Mobile Computing and Networking (MobiCom 2008)*. San Francisco, CA, USA: ACM, September 2008, pp. 116–127.

[28] L. BUTTYÁN, T. HOLCZER, and I. VAJDA, "On the Effectiveness of Changing Pseudonyms to Provide Location Privacy in VANETs," in *4th European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS 2007)*. Cambridge, UK: Springer, July 2007, pp. 129–141.

[29] C. CAMPOLO, A. MOLINARO, and A. VINEL, "Understanding the Performance of Short-lived Control Broadcast Packets in 802.11p/WAVE Vehicular Networks," in *3rd IEEE Vehicular Networking Conference (VNC 2011)*. Amsterdam, Netherlands: IEEE, November 2011, pp. 102–108.

[30] C. CAMPOLO, A. MOLINARO, and A. VINEL, "Understanding Adjacent Channel Interference in Multi-channel VANETs," in *6th IEEE Vehicular Networking Conference (VNC 2014)*. Paderborn, Germany: IEEE, December 2014, pp. 101–104.

[31] S.-H. CHANG, C.-Y. LIN, C.-C. HSU, C.-P. FUNG, and J.-R. HWANG, "The effect of a collision warning system on the driving performance of young drivers at intersections," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 5, pp. 371–380, 2009.

[32] Q. CHEN, D. JIANG, and L. DELGROSSI, "IEEE 1609.4 DSRC Multi-Channel Operations and Its Implications on Vehicle Safety Communications," in *1st IEEE Vehicular Networking Conference (VNC 2009)*. Tokyo, Japan: IEEE, October 2009, pp. 1–8.

[33] X. CHEN and J. PANG, "Measuring Query Privacy in Location-Based Services," in *Proceedings of the 2nd ACM Conference on Data and Application Security and Privacy*. San Antonio, TX, USA: ACM, 2012, pp. 49–60.

[34] C. CHRYSANTHOU and H. L. BERTONI, "Variability of Sector Averaged Signals for UHF Propagation in Cities," *IEEE Transactions on Vehicular Technology*, vol. 39, no. 4, pp. 352–358, November 1990.

[35] L. CODECA, R. FRANK, and T. ENGEL, "Luxembourg SUMO Traffic (LuST) Scenario: 24 Hours of Mobility for Vehicular Networking Research," in *7th IEEE Vehicular Networking Conference (VNC 2015)*. Kyoto, Japan: IEEE, December 2015, pp. 1–8.

[36] D. COOPER, "A More Efficient Use of Delta-CRLs," in *IEEE Symposium on Security and Privacy (S&P 2000)*. Oakland, CA, USA: IEEE, May 2000, pp. 190–202.

[37] R. CREPALDI, R. BEAVERS, B. EHRAT, M. JAEGER, S. BIERSTEKER, and R. KRAVETS, "LoadingZones: Leveraging Street Parking to Enable Vehicular

Internet Access," in *7th ACM International Workshop on Challenged Networks (CHANTS 2012)*. Istanbul, Turkey: ACM, August 2012, pp. 23–30.

[38] R. CREPALDI and R. KRAVETS, "Governing Energy for Parked Cars," in *10th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2013)*. Banff, AB, Canada: IEEE, March 2013, pp. 87–94.

[39] D. CVRCEK, M. KUMPOST, V. MATYAS, and G. DANEZIS, "A Study on The Value of Location Privacy," in *13th ACM Conference on Computer and Communications Security (CCS '06). 5th Workshop on Privacy in Electronic Society (WPES)*. Alexandria, VA, USA: ACM, October 2006, pp. 109–118.

[40] B. DANEV, D. ZANETTI, and S. ČAPKUN, "On Physical-Layer Identification of Wireless Devices," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1–28, November 2012.

[41] G. DANEZIS, S. LEWIS, and R. ANDERSON, "How Much is Location Privacy Worth?" in *Fourth Workshop on the Economics of Information Security (WEIS'05)*, Cambridge, MA, USA, June 2005, pp. 1–13.

[42] M. DENG, K. WUYTS, R. SCANDARIATO, B. PRENEEL, and W. JOOSEN, "A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements," *Requirements Engineering*, vol. 16, no. 1, pp. 3–32, March 2011.

[43] B. DEZSŐ, A. JÜTTNER, and P. KOVÁCS, "LEMON - an Open Source C++ Graph Template Library," *Electronic Notes in Theoretical Computer Science*, vol. 264, no. 5, pp. 23–45, July 2011.

[44] D. DHOUTAUT, A. RÉGIS, and F. SPIES, "Impact of Radio Propagation Models in Vehicular Ad Hoc Networks Simulations," in *3rd ACM International Workshop on Vehicular Ad Hoc Networks (VANET 2006)*. Los Angeles, CA, USA: ACM, September 2006, pp. 40–49.

[45] C. DÍAZ, "Anonymity metrics revisited," in *Anonymous Communication and its Applications*. Dagstuhl, Germany: IBFI, Schloss Dagstuhl, Germany, October 2005, pp. 1–6.

[46] C. DÍAZ, S. SEYS, J. CLAESSENS, and B. PRENEEL, "Towards Measuring Anonymity," in *Second International Workshop on Privacy Enhancing Technologies (PET 2002)*, vol. LNCS 2482. Springer, April 2002, pp. 54–68.

[47] Y. DING, C. WANG, and L. XIAO, "A Static-Node Assisted Adaptive Routing Protocol in Vehicular Networks," in *4th ACM International Workshop on Vehicular*

Ad Hoc Networks (VANET 2007). Montréal, QC, Canada: ACM, September 2007, pp. 59–68.

[48] A. DJANATLIEV, P. KOLOMINSKY-RABAS, B. M. HOFMANN, A. AISENBREY, and R. GERMAN, "Hybrid Simulation Approach for Prospective Assessment of Mobile Stroke Units," in *2nd International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH 2012)*, Rome, Italy, 2012, pp. 357–366.

[49] F. DÖTZER, "Privacy Issues in Vehicular Ad Hoc Networks," in *5th International Workshop on Privacy Enhancing Technologies (PET 2005)*, vol. LNCS 3856. Cavtat, Croatia: Springer, May 2005, pp. 197–209.

[50] J. DOUCEUR, "The Sybil Attack," in *Peer-To-Peer Systems: First International Workshop (IPTPS 2002)*, vol. LNCS 2429. Cambridge, MA, USA: Springer, March 2002, pp. 251–260.

[51] F. DRESSLER, P. HANDLE, and C. SOMMER, "Towards a Vehicular Cloud - Using Parked Vehicles as a Temporary Network and Storage Infrastructure," in *ACM International Workshop on Wireless and Mobile Technologies for Smart Cities (WiMobCity 2014)*. Philadelphia, PA, USA: ACM, August 2014, pp. 11–18.

[52] F. DRESSLER and C. SOMMER, "On the Impact of Human Driver Behavior on Intelligent Transportation Systems," in *71st IEEE Vehicular Technology Conference (VTC2010-Spring)*. Taipei, Taiwan: IEEE, May 2010, pp. 1–5.

[53] F. DRESSLER, C. SOMMER, D. ECKHOFF, and O. K. TONGUZ, "Towards Realistic Simulation of Inter-Vehicle Communication: Models, Techniques and Pitfalls," *IEEE Vehicular Technology Magazine*, vol. 6, no. 3, pp. 43–51, September 2011.

[54] C. DWORK, "Differential Privacy: A Survey of Results," in *Theory and Applications of Models of Computation*, ser. Lecture Notes in Computer Science, M. AGRAWAL, D. DU, Z. DUAN, and A. LI, Eds. Springer Berlin Heidelberg, 2008, vol. 4978, pp. 1–19.

[55] D. ECKHOFF, "Privacy and Surveillance: Concerns About a Future Transportation System," in *1st GI/ITG KuVS Fachgespräch Inter-Vehicle Communication (FG-IVC 2013)*, Innsbruck, Austria, February 2013, pp. 15–18.

[56] D. ECKHOFF, F. DRESSLER, and C. SOMMER, "SmartRevoc: An Efficient and Privacy Preserving Revocation System Using Parked Vehicles," in *38th IEEE Conference on Local Computer Networks (LCN 2013)*. Sydney, Australia: IEEE, October 2013, pp. 855–862.

[57] D. ECKHOFF, A. FESTAG, M. GRUTESER, F. SCHIMANDL, M. SEGATA, and E. UHLEMANN, "Working Group on Best Practices for Field Operational Testing," in *Dagstuhl Seminar 13392 - Inter-Vehicular Communication - Quo Vadis*. Schloss Dagstuhl, Wadern, Germany: Schloss Dagstuhl, September 2013, pp. 206–209.

[58] D. ECKHOFF, T. GANSEN, R. MÄNZ, D. THUM, O. KLAGES, and C. SOMMER, "Simulative Performance Evaluation of the simTD Self Organizing Traffic Information System," in *10th IFIP/IEEE Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2011)*. Favignana Island, Sicily, Italy: IEEE, June 2011, pp. 79–86.

[59] D. ECKHOFF, B. HALMOS, and R. GERMAN, "Potentials and Limitations of Green Light Optimal Speed Advisory Systems," in *5th IEEE Vehicular Networking Conference (VNC 2013)*. Boston, MA: IEEE, December 2013, pp. 103–110.

[60] D. ECKHOFF, M. PROTSENKO, and R. GERMAN, "Towards an Open Source Location Privacy Evaluation Framework for Vehicular Networks," in *80th IEEE Vehicular Technology Conference Fall (VTC 2014-Fall)*. Vancouver, Canada: IEEE, September 2014, pp. 1–2, to appear.

[61] D. ECKHOFF, N. SOFRA, and R. GERMAN, "A Performance Study of Cooperative Awareness in ETSI ITS G5 and IEEE WAVE," in *10th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2013)*. Banff, Canada: IEEE, March 2013, pp. 196–200.

[62] D. ECKHOFF and C. SOMMER, "A Multi-Channel IEEE 1609.4 and 802.11p EDCA Model for the Veins Framework," in *5th ACM/ICST International Workshop on OMNeT++ (OMNeT++ 2012), Poster Session*. Desenzano, Italy: ACM, March 2012.

[63] D. ECKHOFF and C. SOMMER, "Driving for Big Data? Privacy Concerns in Vehicular Networking," *IEEE Security and Privacy*, vol. 12, no. 1, pp. 77–79, February 2014.

[64] D. ECKHOFF and C. SOMMER, "Simulative Performance Evaluation of Vehicular Networks," in *Vehicular Communications and Networks: Architectures, Protocols, Operation and Deployment*, W. CHEN, Ed. Elsevier, March 2015, pp. 255–274.

[65] D. ECKHOFF, C. SOMMER, and F. DRESSLER, "On the Necessity of Accurate IEEE 802.11p Models for IVC Protocol Simulation," in *75th IEEE Vehicular Technology Conference (VTC2012-Spring)*. Yokohama, Japan: IEEE, May 2012, pp. 1–5.

[66] D. ECKHOFF, C. SOMMER, T. GANSEN, R. GERMAN, and F. DRESSLER, "Strong and Affordable Location Privacy in VANETs: Identity Diffusion Using Time-Slots and Swapping," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*. Jersey City, NJ, USA: IEEE, December 2010, pp. 174–181.

[67] D. ECKHOFF, C. SOMMER, T. GANSEN, R. GERMAN, and F. DRESSLER, "SlotSwap: Strong and Affordable Location Privacy in Intelligent Transportation Systems," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 126–133, November 2011.

[68] D. ECKHOFF, C. SOMMER, R. GERMAN, and F. DRESSLER, "Cooperative Awareness At Low Vehicle Densities: How Parked Cars Can Help See Through Buildings," in *IEEE Global Telecommunications Conference (GLOBECOM 2011)*. IEEE, December 2011.

[69] M. EDMAN and B. YENER, "Active Attacks Against Modulation-based Radiometric Identification," Rensselaer Polytechnic Institute, Department of Computer Science, Tech. Rep. 09-02, August 2009.

[70] J. EDMONDS, "Maximum Matching and a Polyhedron with 0, l-vertices," *J. Res. Bur. Stand*, vol. 69B, no. 1-2, pp. 125–130, January 1965.

[71] S. EICHLER, "Performance Evaluation of the IEEE 802.11p WAVE Communication Standard," in *66th IEEE Vehicular Technology Conference (VTC2007-Fall)*, Baltimore, MD, USA, October 2007, pp. 2199–2203.

[72] J. ERDMANN, "SUMO's Lane-Changing Model," in *2nd SUMO Conference – Modeling Mobility with Open Data (SUMO2014)*, vol. Lecture Notes in Mobility. Berlin, Germany: Springer, May 2014, pp. 105–123.

[73] EUROPEAN COMMISSION, MOBILITY AND TRANSPORT, "Road safety in the European Union," European Commission, Brussels, Belgium, Tech. Rep., March 2015.

[74] EUROPEAN PARLIAMENT, "Directive 95/46/EC," *Official Journal*, vol. L 281, pp. 0031–0050, November 1995.

[75] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Security; Security Services and Architecture," ETSI, TS 102 731 V1.1.1, September 2010.

[76] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Security; Threat, Vulnerability and Risk Analysis (TVRA)," ETSI, TS 102 893 V1.1.1, March 2010.

[77] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service," ETSI, TS 102 637-3 V1.1.1, September 2010.

[78] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 1: Requirements," ETSI, TS 102 636-1 V1.1.1, March 2010.

[79] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Decentralized Congestion Control Mechanisms for Intelligent Transport Systems operating in the 5 GHz range; Access layer part," ETSI, TS 102 687 V1.1.1, July 2011.

[80] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Security; ITS Communications Security Architecture and Security Management," ETSI, TS 102 940 V1.1.1, June 2012.

[81] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Security; Trust and Privacy Management," ETSI, TS 102 941 V1.1.1, June 2012.

[82] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band ," ETSI, EN 302 663 V1.2.1, July 2013.

[83] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Users and applications requirements; Part 1: Facility layer structure, functional requirements and specifications ," ETSI, EN 102 894-1 V1.1.1, August 2013.

[84] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service," ETSI, EN 302 637-2 V1.3.0, August 2013.

[85] K. EVENSEN, H.-J. FISCHER, W. HOEFS, J. MORING, R. ROY, and S. SILL, "EU-US Standards Harmonization Task Group Report: Status of ITS Communication (Document HTG3-1)," Federal Highway Administration, Tech. Rep. HWA-JPO-13-080, November 2012.

[86] F. FARNOUD and S. VALAEE, "Reliable Broadcast of Safety Messages in Vehicular Ad Hoc Networks," in *28th IEEE Conference on Computer Communications (INFOCOM 2009)*. Rio de Janeiro, Brazil: IEEE, April 2009, pp. 226–234.

[87] L. M. FEENEY, "Towards Trustworthy Simulation of Wireless MAC/PHY Layers: a Comparison Framework," in *15th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM 2012)*. Paphos, Cyprus: ACM, October 2012, pp. 295–304.

[88] M. FEIRI, J. PETIT, and F. KARGL, "Congestion-based Certificate Omission in VANETs," in *9th ACM International Workshop on Vehicular Internetworking (VANET 2012)*. Ambleside, United Kingdom: ACM, June 2012, pp. 135–138.

[89] M. FEIRI, J. PETIT, and F. KARGL, "The Case for Announcing Pseudonym Changes," in *3rd GI/ITG KuVS Fachgespräch Inter-Vehicle Communication (FG-IVC 2015)*, Ulm, Germany, March 2015.

[90] M. FELLENDORF and P. VORTISCH, "Microscopic Traffic Flow Simulator VISSIM," in *Fundamentals of Traffic Simulation*, ser. International Series in Operations Research & Management Science, J. BARCELÓ, Ed. Springer, 2010, vol. 145, pp. 63–93.

[91] A. FESTAG, H. FÜSSLER, H. HARTENSTEIN, A. SARMA, and R. SCHMITZ, "FleetNet: Bringing Car-to-Car Communication into the Real World," in *11th ITS World Congress and Exhibtion*, vol. 4, Nagoya, Japan, October 2004, p. 16.

[92] R. L. FINN, D. WRIGHT, and M. FRIEDEWALD, *Seven Types of Privacy*. Springer, 2013, pp. 3–32.

[93] M. FIORE and J. HÄRRI, "The Networking Shape of Vehicular Mobility," in *9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc 2008)*. Hong Kong, China: ACM, May 2008, pp. 261–272.

[94] L. FISCHER, A. AIJAZ, C. ECKERT, and D. VOGT, "Secure Revocable Anonymous Authenticated Inter-Vehicle Communication (SRAAC)," in *4th Conference on Embedded Security in Cars (ESCAR 2006)*. Berlin, Germany: PSU, November 2006, pp. 1–9.

[95] L. FISCHER, S. KATZENBEISSER, and C. ECKERT, "Measuring unlinkability revisited," in *Proceedings of the 7th ACM Workshop on Privacy in the Electronic Society*. Alexandria, VA, USA: ACM, 2008, pp. 105–110.

[96] J. FONT, P. IÑIGO, M. DOMÍNGUEZ, J. L. SEVILLANO, and C. AMAYA, "Architecture, design and source code comparison of ns-2 and ns-3 network simulators," in *2010 Spring Simulation Multiconference (SpringSim 2010)*. Orlando, FL, USA: SCS, April 2010, pp. 1–8.

[97] J. FRANKLIN, D. MCCOY, P. TABRIZ, V. NEAGOE, J. V. RANDWYK, and D. SICKER, "Passive Data Link Layer 802.11 Wireless Device Driver Fingerprinting," in

15th USENIX Security Symposium. Vancouver, Canada: USENIX, July 2006, pp. 167–178.

[98] J. FREUDIGER, M. RAYA, M. FÉLEGYHÁZI, P. PAPADIMITRATOS, and J.-P. HUBAUX, "Mix-Zones for Location Privacy in Vehicular Networks," in *First Int. Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS 2007)*. Vancouver, Canada: ACM, August 2007, pp. 1–7.

[99] P. FUXJÄGER, A. COSTANTINI, D. VALERIO, P. CASTIGLIONE, G. ZACHEO, T. ZEMEN, and F. RICCIATO, "IEEE 802.11p Transmission Using GNURadio," in *6th Karlsruhe Workshop on Software Radios (WSR'10)*. Karlsruhe, Germany: KIT, March 2010, pp. 83–86.

[100] M. GERLACH and F. GÜTTLER, "Privacy in VANETs Using Changing Pseudonyms - Ideal and Real," in *65th IEEE Vehicular Technology Conference (VTC2007-Spring)*. Dublin, Ireland: IEEE, April 2007, pp. 2521–2525.

[101] E. GIORDANO, R. FRANK, G. PAU, and M. GERLA, "CORNER: a Realistic Urban Propagation Model for VANET," in *7th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2010), Poster Session*. Kranjska Gora, Slovenia: IEEE, February 2010, pp. 57–60.

[102] P. G. GIPPS, "A Behavioural Car-Following Model for Computer Simulation," *Transportation Research Part B: Methodological*, vol. 15, no. 2, pp. 105–111, April 1981.

[103] P. J. GMYTRASIEWICZ and E. H. DURFEE, "Decision-Theoretic Recursive Modeling and the Coordinated Attack Problem," in *Proceedings of the First International Conference on Artificial Intelligence Planning Systems (AIPS92)*. College Park, MD, USA: Morgan Kaufmann, June 1992, pp. 88–95.

[104] P. GOLLE and K. PARTRIDGE, "On the Anonymity of Home/Work Location Pairs," in *7th International Conference on Pervasive Computing*, vol. LNCS 5538. Nara, Japan: Springer, May 2009, pp. 390–397.

[105] B. GUKHOOL and S. CHERKAOUI, "IEEE 802.11p Modeling in ns-2," in *33rd IEEE Conference on Local Computer Networks (LCN 2008)*, Montreal, Canada, October 2008, pp. 622–626.

[106] J. J. HAAS, Y.-C. HU, and K. P. LABERTEAUX, "Design and Analysis of a Lightweight Certificate Revocation Mechanism for VANET," in *6th ACM International Workshop on Vehicular Inter-Networking (VANET 2009)*. Beijing, China: ACM, September 2009, pp. 89–98.

[107] J. J. HAAS, Y.-C. HU, and K. P. LABERTEAUX, "Efficient Certificate Revocation List Organization and Distribution," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 3, pp. 595–604, March 2011.

[108] F. HAGENAUER, F. DRESSLER, and C. SOMMER, "A Simulator for Heterogeneous Vehicular Networks," in *6th IEEE Vehicular Networking Conference (VNC 2014), Poster Session*. Paderborn, Germany: IEEE, December 2014, pp. 185–186.

[109] M. HAKLAY and P. WEBER, "OpenStreetMap: User-Generated Street Maps," *IEEE Pervasive Computing*, vol. 7, no. 4, pp. 12–18, October 2008.

[110] F. HANSEN and F. MENO, "Mobile fading – Rayleigh and Lognormal Superimposed," *IEEE Transactions on Vehicular Technology*, vol. 26, no. 4, pp. 332–335, November 1977.

[111] J. HARDING, G. POWELL, R. YOON, J. FIKENTSCHER, C. DOYLE, D. SADE, M. LUKUC, J. SIMONS, and J. WANG, "Vehicle-to-Vehicle Communications: Readiness of V2V Technology for Application," National Highway Traffic Safety Administration, Tech. Rep. DOT HS 812 014, August 2014.

[112] J. HÄRRI, F. FILALI, and C. BONNET, "A Framework for Mobility Models Generation and its Application to Inter-Vehicular Networks," in *2005 International Conference on Wireless Networks, Communications and Mobile Computing*. Maui, HI, USA: IEEE, June 2005, pp. 42–47.

[113] H. HARTENSTEIN and K. LABERTEAUX, Eds., *VANET - Vehicular Applications and Inter-Networking Technologies*, ser. Intelligent Transport Systems. John Wiley & Sons, December 2010.

[114] T. HECKER, J. ZECH, B. SCHÄUFELE, R. GRÄFE, and I. RADUSCH, "Model Car Testbed for Development of V2X Applications," *Journal of Communications*, vol. 6, no. 1, February 2011.

[115] B. HOH and M. GRUTESER, "Protecting Location Privacy Through Path Confusion," in *1st International Conference on Security and Privacy for Emerging Areas in Communications Networks (SecureComm 2005)*, Athens, Greece, September 2005, pp. 194–205.

[116] B. HOH, M. GRUTESER, H. XIONG, and A. ALRABADY, "Preserving Privacy in GPS Traces via Uncertainty-Aware Path Cloaking," in *14th ACM Conference on Computer and Communications Security (CCS'07)*. Alexandria, VA, USA: ACM, October 2007, pp. 161–171.

[117] H.-Y. HUANG, P.-E. LUO, M. LI, D. LI, X. LI, W. SHU, and M.-Y. WU, "Performance Evaluation of SUVnet With Real-Time Traffic Data," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3381–3396, November 2007.

[118] L. HUANG, K. MATSUURA, H. YAMANE, and K. SEZAKI, "Enhancing Wireless Location Privacy Using Silent Period," in *IEEE Wireless Communications and Networking Conference (WCNC 2005)*. New Orleans, LA, USA: IEEE, March 2005, pp. 1187–1192.

[119] J.-P. HUBAUX, S. ČAPKUN, and J. LUO, "The Security and Privacy of Smart Vehicles," *IEEE Security and Privacy*, vol. 2, no. 3, pp. 49–55, May 2004.

[120] IEEE, "IEEE Standard for Wireless Access in Vehicular Environments (WAVE) - Multi-channel Operation," IEEE, Std 1609.4, February 2011.

[121] IEEE, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE, Std 802.11-2012, 2012.

[122] IEEE, "IEEE Standard for Wireless Access in Vehicular Environments - Security Services for Applications and Management Messages," IEEE, Tech. Rep. 1609.2, April 2013.

[123] J. T. ISAAC, J. S. CAMARA, S. ZEADALLY, and J. T. MARQUEZ, "A Secure Vehicle-to-Roadside Communication Payment Protocol in Vehicular ad hoc Networks," *Computer Communications*, vol. 31, no. 10, pp. 2478–2484, June 2008.

[124] D. JIANG, Q. CHEN, and L. DELGROSSI, "Communication Density: A Channel Load Metric for Vehicular Communications Research," in *4th IEEE International Conference on Mobile Ad Hoc and Sensor Systems (MASS 2007)*. Pisa, Italy: IEEE, October 2007, pp. 1–8.

[125] S. JOERER, F. DRESSLER, and C. SOMMER, "Comparing Apples and Oranges? Trends in IVC Simulations," in *9th ACM International Workshop on Vehicular Internetworking (VANET 2012)*. Low Wood Bay, UK: ACM, June 2012, pp. 27–32.

[126] S. JOERER, M. SEGATA, B. BLOESSL, R. LO CIGNO, C. SOMMER, and F. DRESSLER, "To Crash or Not to Crash: Estimating its Likelihood and Potentials of Beacon-based IVC Systems," in *4th IEEE Vehicular Networking Conference (VNC 2012)*. Seoul, Korea: IEEE, November 2012, pp. 25–32.

[127] R. KALMAN, "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME Journal of Basic Engineering*, vol. D, no. 82, pp. 35–45, 1960.

[128] A. KHAN, I. STOJMENOVIC, and N. ZAGUIA, "Parameterless broadcasting in static to highly mobile wireless ad hoc, sensor and actuator networks," in *22nd*

International Conference on Advanced Information Networking and Applications. Okinawa, Japan: IEEE, March 2008, pp. 620–627.

[129] B. KITCHENHAM, "Procedures for Performing Systematic Reviews," Keele University, Tech. Rep. TR/SE-0401, July 2004.

[130] R. W. KLEIN, M. A. TEMPLE, and M. J. MENDENHALL, "Application of Wavelet-Based RF Fingerprinting to Enhance Wireless Network Security," *Journal of Communications and Networks*, vol. 11, no. 6, pp. 544–555, December 2009.

[131] F. KLINGLER, F. DRESSLER, J. CAO, and C. SOMMER, "Use Both Lanes: Multi-Channel Beaconing for Message Dissemination in Vehicular Networks," in *10th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2013)*. Banff, Canada: IEEE, March 2013, pp. 162–169.

[132] B. KLOIBER, T. STRANG, F. DE PONTE-MÜLLER, C. GARCIA, and M. RÖCKL, "An Approach for Performance Analysis of ETSI ITS-G5A MAC for Safety Applications," in *10th International Conference on Intelligent Transport Systems Telecommunications (ITST'10)*. Kyoto, Japan: IEICE, November 2010, pp. 1–7.

[133] B. KLOIBER, T. STRANG, M. RÖCKL, and F. DE PONTE-MÜLLER, "Performance of CAM Based Safety Applications Using ITS-G5A MAC in High Dense Scenarios," in *Intelligent Vehicles Symposium (IV'11)*, Baden-Baden, Germany, June 2011, pp. 654–660.

[134] P. KNAPIK, E. SCHOCH, M. MÜLLER, and F. KARGL, "Understanding Vehicle Related Crime to Elaborate on Countermeasures based on ADAS and V2X Communication," in *4th IEEE Vehicular Networking Conference (VNC 2012)*. Seoul, Korea: IEEE, November 2012, pp. 86–93.

[135] F. KOERFER, "Channel Access Mechanisms for Clusters of Parking Vehicles," Master's Thesis, University of Erlangen-Nürnberg, Erlangen, Germany, September 2015.

[136] T. KOHNO, A. BROIDO, and K. CLAFFY, "Remote Physical Device Fingerprinting," *IEEE Transactions on Dependable and Secure Computing*, vol. 2, no. 2, pp. 93–108, May 2005.

[137] A. KÖPKE, M. SWIGULSKI, K. WESSEL, D. WILLKOMM, P. K. HANEVELD, T. PARKER, O. VISSER, H. S. LICHTE, and S. VALENTIN, "Simulating Wireless and Mobile Networks in OMNeT++ – The MiXiM Vision," in *1st ACM/ICST International Workshop on OMNeT++ (OMNeT++ 2008)*. Marseille, France: ACM, March 2008, pp. 1–8.

[138] G. KORKMAZ, E. EKICI, F. ÖZGÜNER, and Ü. ÖZGÜNER, "Urban Multi-Hop Broadcast Protocol for Inter-Vehicle Communication Systems," in *1st ACM Workshop on Vehicular Ad Hoc Networks (VANET 2004)*. Philadelphia, PA, USA: ACM, October 2004, pp. 76–85.

[139] D. KOTZ and T. HENDERSON, "Crawdad: A Community Resource for Archiving Wireless Data at Dartmouth," *IEEE Pervasive Computing*, vol. 4, no. 4, pp. 12–14, October 2005.

[140] D. KRAJZEWICZ, R. BLOKPOEL, F. CARTOLANO, P. CATALDI, A. GONZALEZ, O. LAZARO, J. LEGUAY, L. LIN, J. MANEROS, and M. RONDINONE, "iTETRIS - A System for the Evaluation of Cooperative Traffic Management Solutions," in *14th International Forum on Advanced Microsystems for Automotive Applications (AMAA 2010)*. Berlin, Germany: Springer, May 2010, pp. 399–410.

[141] D. KRAJZEWICZ, G. HERTKORN, C. RÖSSEL, and P. WAGNER, "SUMO (Simulation of Urban MObility); An Open-source Traffic Simulation," in *4th Middle East Symposium on Simulation and Modelling (MESM 2002)*. Sharjah, United Arab Emirates: EUROSIS, September 2002, pp. 183–187.

[142] S. KRAUSS, "Microscopic Modeling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics," PhD Thesis, University of Cologne, 1998.

[143] J. KRUMM, "Inference Attacks on Location Tracks," in *5th International Conference on Pervasive Computing (PERVASIVE 2007)*, ser. LNCS, vol. 4480. Toronto, Canada: Springer, May 2007, pp. 127–143.

[144] A. KUNTZ, F. SCHMIDT-EISENLOHR, O. GRAUTE, H. HARTENSTEIN, and M. ZITTERBART, "Introducing Probabilistic Radio Propagation Models in OMNET++ Mobility Framework and Cross Validation Check with NS-2," in *1st ACM/ICST International Workshop on OMNeT++ (OMNeT++ 2008)*. Marseille, France: ICST, Mar. 2008, pp. 1–7.

[145] A. KWOCZEK, Z. RAIDA, J. LÁČÍK, M. POKORNÝ, J. PUSKELY, and P. VÁGNER, "Influence of Car Panorama Glass Roofs on Car2car Communication," in *3rd IEEE Vehicular Networking Conference (VNC 2011), Poster Session*. Amsterdam, Netherlands: IEEE, November 2011, pp. 246–251.

[146] K. P. LABERTEAUX, J. J. HAAS, and Y.-C. HU, "Security Certificate Revocation List Distribution for VANET," in *5th ACM International Workshop on Vehicular Inter-Networking (VANET 2008)*. San Francisco, CA, USA: ACM, September 2008, pp. 88–89.

[147] A. M. LAW, *Simulation, Modeling and Analysis*, 4th ed. McGraw-Hill, 2007.

[148] S. LEFÈVRE, J. PETIT, R. BAJCSY, C. LAUGIER, and F. KARGL, "Impact of V2X Privacy Strategies on Intersection Collision Avoidance Systems," in *5th IEEE Vehicular Networking Conference (VNC 2013)*. Boston, MA: IEEE, December 2013, pp. 71–78.

[149] S. LEMAN-LANGLOIS, "Privacy as Currency: Crime, Information and Control in Cyberspace," in *Technocrime: Technology, Crime and Social Control*. Portland, OR, USA: Willan Publishing, July 2008, pp. 112–138.

[150] I. LEQUERICA, J. MARTINEZ, and P. RUIZ, "Efficient Certificate Revocation in Vehicular Networks using NGN Capabilities," in *72nd IEEE Vehicular Technology Conference Fall (VTC2010-Fall)*. Ottawa, Canada: IEEE, September 2010, pp. 1–5.

[151] M. LI, K. SAMPIGETHAYA, L. HUANG, and R. POOVENDRAN, "Swing & Swap: User-Centric Approaches Towards Maximizing Location Privacy," in *5th ACM Workshop On Privacy In The Electronic Society*. Alexandria, VA, USA: ACM, October 2006, pp. 19–28.

[152] D.-W. LIM, S.-J. HEO, and J.-S. NO, "An Overview of Peak-to-Average Power Ratio Reduction Schemes for OFDM Signals," *Journal of Communications and Networks*, vol. 11, no. 3, pp. 229–239, June 2009.

[153] T. LITMAN, "Parking Management: Strategies, Evaluation and Planning," Victoria Transport Policy Institute, Tech. Rep., November 2013.

[154] B. LIU, B. KHORASHADI, D. GHOSAL, C.-N. CHUAH, and M. ZHANG, "Assessing the VANET's Local Information Storage Capability under Different Traffic Mobility," in *29th IEEE Conference on Computer Communications (INFOCOM 2010)*. San Diego, CA, USA: IEEE, March 2010, pp. 1–5.

[155] B. LIU, B. KHORASHADI, H. DU, D. GHOSAL, C.-N. CHUAH, and M. ZHANG, "VGSim: An Integrated Networking and Microscopic Vehicular Mobility Simulation Platform," *IEEE Communications Magazine*, vol. 47, no. 5, pp. 134–141, May 2009.

[156] N. LIU, M. LIU, W. LOU, G. CHEN, and J. CAO, "PVA in VANETs: Stopped Cars Are Not Silent," in *30th IEEE Conference on Computer Communications (INFOCOM 2011), Mini-Conference*. Shanghai, China: IEEE, April 2011, pp. 431–435.

[157] C. LOCHERT, B. SCHEUERMANN, M. CALISKAN, and M. MAUVE, "The Feasibility of Information Dissemination in Vehicular Ad-Hoc Networks," in *4th Annual Conference on Wireless On demand Network Systems and Services (WONS 2007)*. Obergurgl, Austria: IEEE, January 2007, pp. 92–99.

[158] N. E. LOWNES and R. B. MACHEMEHL, "VISSIM: A Multi-parameter Sensitivity Analysis," in *38th Winter Simulation Conference (WSC '06)*. Monterey, CA, USA: IEEE, Dec. 2006, pp. 1406–1413.

[159] Z. MA, F. KARGL, and M. WEBER, "Measuring long-term location privacy in vehicular communication systems," *Elsevier Computer Communications*, vol. 33, no. 12, pp. 1414–1427, March 2010.

[160] P. C. MAHALANOBIS, "On the Generalized Distance in Statistics," *Proceedings of the National Institute of Sciences of India*, vol. 2, no. 1, pp. 49–55, April 1936.

[161] F. MALANDRINO, C. CASETTI, C.-F. CHIASSERINI, C. SOMMER, and F. DRESSLER, "Content Downloading in Vehicular Networks: Bringing Parked Cars Into the Picture," in *23rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2012)*. Sydney, Australia: IEEE, September 2012, pp. 1534–1539.

[162] F. MALANDRINO, C. CASETTI, C.-F. CHIASSERINI, C. SOMMER, and F. DRESSLER, "The Role of Parked Cars in Content Downloading for Vehicular Networks," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4606–4617, November 2014.

[163] T. MANGEL, O. KLEMP, and H. HARTENSTEIN, "A Validated 5.9 GHz Non-Line-of-Sight Path-Loss and Fading Model for Inter-Vehicle Communication," in *11th International Conference on ITS Telecommunications (ITST 2011)*. St. Petersburg, Russia: IEEE, August 2011, pp. 75–80.

[164] T. MANGEL, T. KOSCH, and H. HARTENSTEIN, "A Comparison of UMTS and LTE for Vehicular Safety Communication at Intersections," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*. Jersey City, NJ, USA: IEEE, December 2010, pp. 293–300.

[165] R. MANGHARAM, D. WELLER, R. RAJKUMAR, P. MUDALIGE, and F. BAI, "GrooveNet: A Hybrid Simulator for Vehicle-to-Vehicle Networks (Invited Paper)," in *2nd International Workshop on Vehicle-to-Vehicle Communications (V2VCOM 2006)*. San Jose, CA, USA: IEEE, July 2006, pp. 1–8.

[166] C. MORENCY and M. TRÉPANIER, "Characterizing Parking Spaces Using Travel Survey Data," CIRRELT, TR 2008-15, May 2008.

[167] R. MORRIS, J. JANNOTTI, F. KAASHOEK, J. LI, and D. DECOUTO, "CarNet: A Scalable Ad hoc Wireless Network System," in *9th ACM SIGOPS European Workshop*. Kolding, Denmark: ACM, Sep. 2000, pp. 61–65.

[168] P. MUHLETHALER, A. LAOUITI, and Y. TOOR, "Comparison of Flooding Techniques for Safety Applications in VANETs," in *7th International Conference on ITS Telecommunications (ITST'07)*. Sophia Antipolis, France: IEEE, June 2007, pp. 1–6.

[169] K. G. MURTY, "An Algorithm for Ranking all the Assignments in Order of Increasing Cost," *Operations Research*, vol. 16, no. 3, pp. 682–687, May 1968.

[170] K. NAGEL and M. SCHRECKENBERG, "A cellular automaton model for freeway traffic," *Journal de Physique I France*, vol. 2, pp. 2221–2229, December 1992.

[171] N. NAKAGAMI, "The m-Distribution, a General Formula for Intensity Distribution of Rapid Fading," in *Statistical Methods in Radio Wave Propagation*, W. G. HOFFMAN, Ed. Oxford, UK: Pergamon, 1960.

[172] T. NEUDECKER, N. AN, O. K. TONGUZ, T. GAUGEL, and J. MITTAG, "Feasibility of Virtual Traffic Lights in Non-Line-of-Sight Environments," in *9th ACM International Workshop on Vehicular Internetworking (VANET 2012)*. Low Wood Bay, UK: ACM, Jun. 2012, pp. 103–106.

[173] H. NISSENBAUM, "Privacy as Contextual Integrity," *Washington Law Review*, vol. 79, no. 1, pp. 119–158, June 2004.

[174] Y. PAN, J. LI, L. FENG, and B. XU, "An Analytical Model for Random Changing Pseudonyms Scheme in VANETs," in *International Conference on Network Computing and Information (NCIS 2011)*. Guilin, China: IEEE, May 2011, pp. 141–145.

[175] R. PANAYAPPAN, J. M. TRIVEDI, A. STUDER, and A. PERRIG, "VANET-based Approach for Parking Space Availability," in *4th ACM International Workshop on Vehicular Ad Hoc Networks (VANET 2007)*. Québec, Canada: ACM, September 2007, pp. 75–76.

[176] P. PAPADIMITRATOS, G. MEZZOUR, and J.-P. HUBAUX, "Certificate Revocation List Distribution in Vehicular Communication Systems," in *5th ACM International Workshop on Vehicular Inter-Networking (VANET 2008)*. San Francisco, CA, USA: ACM, September 2008, pp. 86–87.

[177] N. PATWARI and S. K. KASERA, "Robust Location Distinction Using Temporal Link Signatures," in *13th ACM International Conference on Mobile Computing and Networking (MobiCom 2007)*. Montreal, Canada: ACM, September 2007, pp. 111–122.

[178] K. PAWLIKOWSKI, H.-D. JEONG, and J.-S. R. LEE, "On Credibility of Simulation Studies of Telecommunication Networks," *IEEE Communications Magazine*, vol. 40, no. 1, pp. 132–139, January 2002.

[179] J. PETIT, C. BOSCH, M. FEIRI, and F. KARGL, "On the Potential of PUF for Pseudonym Generation in Vehicular Networks," in *4th IEEE Vehicular Networking Conference (VNC 2012)*. Seoul, Korea: IEEEE, November 2012, pp. 94–100.

[180] J. PETIT, F. SCHAUB, M. FEIRI, and F. KARGL, "Pseudonym Schemes in Vehicular Networks: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 228–255, March 2015.

[181] A. PFITZMANN and M. HANSEN, "A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management," August 2010, v0.34.

[182] M. PIÓRKOWSKI, M. RAYA, A. L. LUGO, P. PAPADIMITRATOS, M. GROSSGLAUSER, and J.-P. HUBAUX, "TraNS: Realistic Joint Traffic and Network Simulator for VANETs," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 12, pp. 31–33, January 2008.

[183] J. PLOEG, B. SCHEEPERS, E. VAN NUNEN, N. VAN DE WOUW, and H. NIJMEIJER, "Design and Experimental Evaluation of Cooperative Adaptive Cruise Control," in *IEEE International Conference on Intelligent Transportation Systems (ITSC 2011)*. Washington, DC: IEEE, October 2011, pp. 260–265.

[184] J. PLOEG, S. SHLADOVER, H. NIJMEIJER, and N. VAN DE WOUW, "Introduction to the Special Issue on the 2011 Grand Cooperative Driving Challenge," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 3, pp. 989–993, September 2012.

[185] M. PROTSENKO, "A Framework for Performance Analysis of Tracking Algorithms in Vehicular Networks," Pre-Master's Thesis (Studienarbeit), University of Erlangen, August 2011.

[186] T. S. RAPPAPORT, *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall, 2009.

[187] M. RAYA, P. PAPADIMITRATOS, I. AAD, D. JUNGELS, and J.-P. HUBAUX, "Eviction of Misbehaving and Faulty Nodes in Vehicular Networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 25, no. 8, pp. 1557–1568, October 2007.

[188] M. RAYA, R. SHOKRI, and J.-P. HUBAUX, "On the Tradeoff Between Trust and Privacy in Wireless Ad Hoc Networks," in *3rd ACM Conference on Wireless Network Security (WiSec 2010)*. Hoboken, NJ, USA: ACM, March 2010, pp. 75–80.

[189] RESEARCH and I. T. A. (RITA), "Connected Vehicle Safety Pilot Program," U.S. Department of Transportation (US DOT), ITS Research Fact Sheet FHWA-JPO-11-031 v3, April 2013.

[190] S. O. RICE, "Statistical Properties of a Sine Wave Plus Random Noise," *Bell System Technical Journal*, vol. 27, no. 1, pp. 109–157, January 1948.

[191] P. RIEKERT and T.-E. SCHUNCK, "Zur Fahrmechanik des gummibereiften Kraftfahrzeugs," *Ingenieur-Archiv*, vol. 11, no. 3, pp. 210–224, 1940.

[192] M. RONDINONE, J. MANEROS, D. KRAJZEWICZ, R. BAUZA, P. CATALDI, F. HRIZI, J. GOZALVEZ, V. KUMAR, M. RÖCKL, L. LIN, O. LAZARO, J. LEGUAY, J. HÄRRI, S. VAZ, Y. LOPEZ, M. SEPULCRE, M. WETTERWALD, R. BLOKPOEL, and F. CARTOLANO, "iTETRIS: A Modular Simulation Platform for the Large Scale Evaluation of Cooperative ITS Applications," *Simulation Modelling Practice and Theory*, vol. 34, pp. 99–125, May 2013.

[193] F. J. ROS, P. M. RUIZ, and I. STOJMENOVIC, "Reliable and Efficient Broadcasting in Vehicular Ad Hoc Networks," in *69th IEEE Vehicular Technology Conference (VTC2009-Spring)*. Barcelona, Spain: IEEE, April 2009, pp. 1–5.

[194] J. RYBICKI, B. SCHEUERMANN, M. KOEGEL, and M. MAUVE, "PeerTIS - A Peer-to-Peer Traffic Information System," in *6th ACM International Workshop on Vehicular Inter-Networking (VANET 2009)*. Beijing, China: ACM, September 2009, pp. 23–32.

[195] SAE INT., "Dedicated Short Range Communications (DSRC) Message Set Dictionary," SAE, Tech. Rep. J2735-201509, September 2015.

[196] SAE INT. DSRC COMMITTEE, "On-board Minimum Performance Requirements for V2V Safety Communications," SAE, Draft Std. J2945/1, February 2015.

[197] S. SAI, T. OSHIDA, R. ONISHI, A. YOSHIOKA, and H. TANAKA, "Comparisons of Non-Line-Of-Sight Inter-Vehicle Communications in the Urban Environment Between 5.9GHz and 700MHz Bands," in *4th IEEE Vehicular Networking Conference (VNC 2012), Poster Session*. Seoul, Korea: IEEE, November 2012, pp. 144–151.

[198] K. SAMPIGETHAYA, L. HUANG, M. LI, R. POOVENDRAN, K. MATSUURA, and K. SEZAKI, "CARAVAN: Providing location privacy for VANET," in *Embedded Security in Cars (ESCAR 2005)*, Tallinn, Estonia, July 2005, pp. 1–15.

[199] R. G. SARGENT, "Verification and Validation of Simulation Models," in *39th Winter Simulation Conference (WSC 2007)*. Piscataway, NJ, USA: IEEE, December 2007, pp. 124–137.

[200] S. R. SAUNDERS and A. ARAGÓN-ZAVALA, *Antennas and Propagation for Wireless Communication Systems*, 2nd ed. Wiley, 2007.

[201] S. SCHELLENBERG, R. BERNDT, D. ECKHOFF, and R. GERMAN, "A Computationally Inexpensive Battery Model for the Microscopic Simulation of Electric Vehicles," in *80th IEEE Vehicular Technology Conference Fall (VTC 2014-Fall)*. Vancouver, Canada: IEEE, September 2014, pp. 1–6.

[202] B. SCHEUERMANN, C. LOCHERT, J. RYBICKI, and M. MAUVE, "A Fundamental Scalability Criterion for Data Aggregation in VANETs," in *15th ACM International Conference on Mobile Computing and Networking (MobiCom 2009)*. Beijing, China: ACM, September 2009, pp. 285–296.

[203] B. SCHILIT, J. HONG, and M. GRUTESER, "Wireless Location Privacy Protection," *Computer*, vol. 36, no. 12, pp. 135–137, December 2003.

[204] E. SCHOCH, F. KARGL, T. LEINMÜLLER, S. SCHLOTT, and P. PAPADIMITRATOS, "Impact of Pseudonym Changes on Geographic Routing in VANETs," in *3rd European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS 2006)*, vol. LNCS 4357. Hamburg, Germany: Springer, September 2006.

[205] B. SCHÜNEMANN, "V2X Simulation Runtime Infrastructure VSimRTI: An Assessment Tool to Design Smart Traffic Management Systems," *Elsevier Computer Networks*, vol. 55, no. 14, pp. 3189–3198, October 2011.

[206] M. SEGATA, S. JOERER, B. BLOESSL, C. SOMMER, F. DRESSLER, and R. LO CIGNO, "PLEXE: A Platooning Extension for Veins," in *6th IEEE Vehicular Networking Conference (VNC 2014)*. Paderborn, Germany: IEEE, December 2014, pp. 53–60.

[207] A. SERJANTOV and G. DANEZIS, "Towards an Information Theoretic Metric for Anonymity," in *2nd International Workshop on Privacy Enhancing Technologies (PET 2002)*, R. DINGLEDINE and P. SYVERSON, Eds., vol. LNCS 2482. Springer, 2003, pp. 41–53.

[208] S. SHLADOVER, "PATH at 20 – History and Major Milestones," in *IEEE Intelligent Transportation Systems Conference (ITSC 2006)*, Toronto, Canada, September 2006, pp. 22–29.

[209] R. SHOKRI, J. FREUDIGER, and J.-P. HUBAUX, "A Unified Framework for Location Privacy," in *Proceedings of the 3rd Symposium on Hot Topics in Privacy Enhancing Technologies (HotPETS 2010)*, Berlin, Germany, July 2010, pp. 203–214.

[210] R. SHOKRI, C. TRONCOSO, C. DÍAZ, J. FREUDIGER, and J.-P. HUBAUX, "Unraveling an Old Cloak: k-anonymity for Location Privacy," in *Proceedings of the 9th Annual ACM Workshop on Privacy in the Electronic Society*. Chicago, IL, USA: ACM, 2010, pp. 115–118.

[211] H. SODEIKAT, *Co-operative Transport Management with Euro-Scout*. Artech Print on Demand, 1993, pp. 65–78.

[212] C. SOMMER, "Car-to-X Communication in Heterogeneous Environments," PhD Thesis (Dissertation), University of Erlangen, June 2011.

[213] C. SOMMER and F. DRESSLER, "The DYMO Routing Protocol in VANET Scenarios," in *66th IEEE Vehicular Technology Conference (VTC2007-Fall)*. Baltimore, MD, USA: IEEE, September 2007, pp. 16–20.

[214] C. SOMMER and F. DRESSLER, "Using the Right Two-Ray Model? A Measurement based Evaluation of PHY Models in VANETs," in *17th ACM International Conference on Mobile Computing and Networking (MobiCom 2011), Poster Session*. Las Vegas, NV, USA: ACM, September 2011, pp. 1–3.

[215] C. SOMMER and F. DRESSLER, *Vehicular Networking*. Cambridge University Press, November 2014.

[216] C. SOMMER, D. ECKHOFF, and F. DRESSLER, "Improving the Accuracy of IVC Simulation using Crowd-sourced Geodata," *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, vol. 33, no. 4, pp. 278–283, December 2010.

[217] C. SOMMER, D. ECKHOFF, and F. DRESSLER, "IVC in Cities: Signal Attenuation by Buildings and How Parked Cars Can Improve the Situation," *IEEE Transactions on Mobile Computing*, vol. 13, no. 8, pp. 1733–1745, August 2014.

[218] C. SOMMER, D. ECKHOFF, R. GERMAN, and F. DRESSLER, "A Computationally Inexpensive Empirical Model of IEEE 802.11p Radio Shadowing in Urban Environments," in *8th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2011)*. Bardonecchia, Italy: IEEE, January 2011, pp. 84–90.

[219] C. SOMMER, R. GERMAN, and F. DRESSLER, "Bidirectionally Coupled Network and Road Traffic Simulation for Improved IVC Analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, January 2011.

[220] C. SOMMER, S. JOERER, and F. DRESSLER, "On the Applicability of Two-Ray Path Loss Models for Vehicular Network Simulation," in *4th IEEE Vehicular*

Networking Conference (VNC 2012). Seoul, Korea: IEEE, November 2012, pp. 64–69.

[221] C. SOMMER, S. JOERER, M. SEGATA, O. K. TONGUZ, R. LO CIGNO, and F. DRESSLER, "How Shadowing Hurts Vehicular Communications and How Dynamic Beaconing Can Help," in *32nd IEEE Conference on Computer Communications (INFOCOM 2013), Mini-Conference*. Turin, Italy: IEEE, April 2013, pp. 110–114.

[222] C. SOMMER, A. SCHMIDT, Y. CHEN, R. GERMAN, W. KOCH, and F. DRESSLER, "On the Feasibility of UMTS-based Traffic Information Systems," *Elsevier Ad Hoc Networks, Special Issue on Vehicular Networks*, vol. 8, no. 5, pp. 506–517, July 2010.

[223] C. SOMMER, A. SCHMIDT, R. GERMAN, W. KOCH, and F. DRESSLER, "Simulative Evaluation of a UMTS-based Car-to-Infrastructure Traffic Information System," in *3rd IEEE Workshop on Automotive Networking and Applications (AutoNet 2008)*. New Orleans, LA, USA: IEEE, December 2008.

[224] C. SOMMER, O. K. TONGUZ, and F. DRESSLER, "Adaptive Beaconing for Delay-Sensitive and Congestion-Aware Traffic Information Systems," in *2nd IEEE Vehicular Networking Conference (VNC 2010)*. Jersey City, NJ, USA: IEEE, December 2010, pp. 1–8.

[225] C. SOMMER, O. K. TONGUZ, and F. DRESSLER, "Traffic Information Systems: Efficient Message Dissemination via Adaptive Beaconing," *IEEE Communications Magazine*, vol. 49, no. 5, pp. 173–179, May 2011.

[226] C. SOMMER, Z. YAO, R. GERMAN, and F. DRESSLER, "Simulating the Influence of IVC on Road Traffic using Bidirectionally Coupled Simulators," in *IEEE Workshop on Mobile Networking for Vehicular Environments (MOVE 2008)*. Phoenix, AZ, USA: IEEE, April 2008, pp. 1–6.

[227] STATISTISCHES BUNDESAMT, "Verkehrsunfälle," Statistisches Bundesamt, Wiesbaden, Germany, Tech. Rep. Fachserie 8 Reihe 7, August 2015.

[228] H. STÜBING, M. BECHLER, D. HEUSSNER, T. MAY, I. RADUSCH, H. RECHNER, and P. VOGEL, "simTD: A Car-to-X System Architecture for Field Operational Tests," *IEEE Communications Magazine*, vol. 48, no. 5, pp. 148–154, May 2010.

[229] S. SUBRAMANIAN, M. WERNER, S. LIU, J. JOSE, R. LUPOAIE, and X. WU, "Congestion Control for Vehicular Safety: Synchronous and Asynchronous MAC Algorithms," in *9th ACM International Workshop on Vehicular Internetworking (VANET 2012)*. Low Wood Bay, UK: ACM, June 2012, pp. 63–72.

[230] L. SWEENEY, "k-Anonymity: A Model for Protecting Privacy," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 05, pp. 557–570, October 2002.

[231] V. TALIWAL, D. JIANG, H. MANGOLD, C. CHEN, and R. SENGUPTA, "Empirical Determination of Channel Characteristics for DSRC Vehicle-to-Vehicle Communication," in *1st ACM Workshop on Vehicular Ad Hoc Networks (VANET 2004)*. Philadelphia, PA, USA: ACM, October 2004, p. 88.

[232] A. TANG and A. YIP, "Collision Avoidance Timing Analysis of DSRC-Based Vehicles," *Accident Analysis and Prevention*, vol. 42, no. 1, pp. 182–195, January 2010.

[233] C. TINGVALL and N. HAWORTH, "Vision Zero - an ethical approach to safety and mobility," in *6th ITE International Conference Road Safety & Traffic Enforcement (Beyond 2000)*, Melbourne, Australia, September 1999, pp. 1–14.

[234] R. TOMKEWITSCH, "Dynamic Route Guidance and Interactive Transport Management with ALI-Scout," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 1, pp. 45–50, February 1991.

[235] O. TONGUZ, N. WISITPONGPHAN, F. BAI, P. MUDALIGE, and V. SADEKART, "Broadcasting in VANET," in *Mobile Networking for Vehicular Environments (MOVE 2007)*. Anchorage, AK, USA: IEEE, May 2007, pp. 7–12.

[236] O. K. TONGUZ and M. BOBAN, "Multiplayer games over Vehicular Ad Hoc Networks: A new application," *Elsevier Ad Hoc Networks*, vol. 8, no. 5, pp. 531–543, July 2010.

[237] O. K. TONGUZ, N. WISITPONGPHAN, and F. BAI, "DV-CAST: A Distributed Vehicular Broadcast Protocol for Vehicular Ad Hoc Networks," *IEEE Wireless Communications*, vol. 17, no. 2, pp. 47–57, April 2010.

[238] M. TORRENT-MORENO, J. MITTAG, P. SANTI, and H. HARTENSTEIN, "Vehicle-to-Vehicle Communication: Fair Transmit Power Control for Safety-Critical Information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3684–3703, September 2009.

[239] M. TREIBER, A. HENNECKE, and D. HELBING, "Congested Traffic States in Empirical Observations and Microscopic Simulations," *Physical Review E*, vol. 62, no. 2, pp. 1805–1824, August 2000.

[240] M. TREIBER, A. HENNECKE, and D. HELBING, "Microscopic Simulation of Congested Traffic," in *Traffic and Granular Flow '99*, D. HELBING, H. HERRMANN, M. SCHRECKENBERG, and D. WOLF, Eds. Heidelberg: Springer, 2000.

[241] M. TREIBER and A. KESTING, "Modeling Lane-Changing Decisions with MOBIL," in *Conference on Traffic and Granular Flow '07*. Paris, France: Springer, June 2007, pp. 211–221.

[242] S. UPPOOR and M. FIORE, "Large-scale Urban Vehicular Mobility for Networking Research," in *3rd IEEE Vehicular Networking Conference (VNC 2011)*. Amsterdam, Netherlands: IEEE, November 2011, pp. 62–69.

[243] O. URETEN and N. SERINKEN, "Wireless Security Through RF Fingerprinting," *Canadian Journal of Electrical and Computer Engineering*, vol. 32, no. 1, pp. 27–33, May 2007.

[244] A. VARGA, "The OMNeT++ Discrete Event Simulation System," in *European Simulation Multiconference (ESM 2001)*, Prague, Czech Republic, June 2001.

[245] N. VRATONJIC, K. HUGUENIN, V. BINDSCHAEDLER, and J.-P. HUBAUX, "How Others Compromise Your Location Privacy: The Case of Shared Public IPs at Hotspots," in *Proceedings of 13th International Symposium on Privacy Enhancing Technologies (PETS 2013)*, ser. LNCS 7981. Bloomington, IN, USA: Springer, July 2013, pp. 123–142.

[246] I. WAGNER and D. ECKHOFF, "Privacy Assessment in Vehicular Networks Using Simulation," in *Winter Simulation Conference (WSC '14)*, Savannah, GA, December 2014, pp. 3155 – 3166.

[247] R. E. WALPOLE, R. H. MYERS, S. L. MYERS, and K. YE, *Probability and Statistics for Engineers and Scientists*, 9th ed. Pearson, January 2012.

[248] S.-Y. WANG and C.-C. LIN, "NCTUns 5.0: A Network Simulator for IEEE 802.11(p) and 1609 Wireless Vehicular Network Researches," in *68th IEEE Vehicular Technology Conference (VTC2008-Fall)*, Calgary, Canada, September 2008, pp. 1–2.

[249] Y. WANG, A. AHMED, B. KRISHNAMACHARI, and K. PSOUNIS, "IEEE 802.11p Performance Evaluation and Protocol Enhancement," in *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*. Columbus, OH, USA: IEEE, September 2008, pp. 317–322.

[250] M. WELLENS, B. WESTPHAL, and P. MÄHÖNEN, "Performance Evaluation of IEEE 802.11-based WLANs in Vehicular Scenarios," in *65th IEEE Vehicular Technology Conference (VTC2007-Spring)*, Dublin, Ireland, April 2007, pp. 1167 –1171.

[251] J. WENGER, "Automotive Radar - Status and Perspectives," in *Compound Semiconductor Integrated Circuit Symposium, 2005 (CSIC'05)*, IEEE. IEEE, October 2005, pp. 21–24.

[252] A. WESTIN, *Privacy and Freedom*. Atheneum, 1967.

[253] C. WEWETZER, M. CALISKAN, K. MEIER, and A. LUEBKE, "Experimental Evaluation of UMTS and Wireless LAN for Inter-Vehicle Communication," in *7th International Conference on ITS Telecommunications (ITST 2007)*. Sophia Antipolis, France: IEEE, June 2007, pp. 1–6.

[254] W. WHYTE, A. WEIMERSKIRCH, V. KUMAR, and T. HEHN, "A Security Credential Management System for V2V Communications," in *5th IEEE Vehicular Networking Conference (VNC 2013)*. Boston, MA: IEEE, December 2013, pp. 1–8.

[255] B. WIEDERSHEIM, Z. MA, F. KARGL, and P. PAPADIMITRATOS, "Privacy in Inter-Vehicular Networks: Why simple pseudonym change is not enough," in *7th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2010)*. Kranjska Gora, Slovenia: IEEE, February 2010, pp. 176–183.

[256] L. WISCHHOF, A. EBNER, H. ROHLING, M. LOTT, and R. HALFMANN, "SOTIS - A Self-Organizing Traffic Information System," in *57th IEEE Vehicular Technology Conference (VTC2003-Spring)*, vol. 4. Jeju, South Korea: IEEE, April 2003, pp. 2442–2446.

[257] N. WISITPONGPHAN, F. BAI, P. MUDALIGE, V. SADEKAR, and O. TONGUZ, "Routing in Sparse Vehicular Ad Hoc Wireless Networks," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 25, no. 8, pp. 1538–1556, October 2007.

[258] N. WISITPONGPHAN, O. K. TONGUZ, J. S. PARIKH, P. MUDALIGE, F. BAI, and V. SADEKAR, "Broadcast Storm Mitigation Techniques in Vehicular Ad Hoc Networks," *IEEE Wireless Communications*, vol. 14, no. 6, pp. 84–94, December 2007.

[259] Q. XU, R. SEGUPTA, and D. JIANG, "Design and Analysis of Highway Safety Communication Protocol in 5.9 GHz Dedicated Short Range Communication Spectrum," in *57th IEEE Vehicular Technology Conference (VTC2003-Spring)*, vol. 4. Jeju, Korea: IEEE, April 2003, pp. 2451–2455.

[260] J. YOON, M. LIU, and B. NOBLE, "Random Waypoint Considered Harmful," in *22nd IEEE Conference on Computer Communications (INFOCOM 2003)*. San Francisco, CA, USA: IEEE, March 2003, pp. 1312–1321.

[261] W. ZIMDAHL, "Guidelines and some Developments for a new Modular Driver Information System," in *34th IEEE Vehicular Technology Conference (VTC1984)*, vol. 34. Pittsburgh, PA, USA: IEEE, May 1984, pp. 178–182.