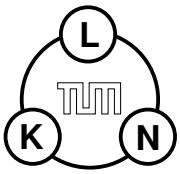


Technische Universität München
Lehrstuhl für Kommunikationsnetze

TECHNISCHE
UNIVERSITÄT
MÜNCHEN

Solutions for Scalable Communication and System Security in Vehicular Network Architectures

Dipl.-Ing. Univ. Claus Stephan Eichler

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr.-Ing. Josef S. Kindersberger
Prüfer der Dissertation:	1.	Univ.-Prof. Dr.-Ing. Jörg Eberspächer
	2.	Univ.-Prof. Dr.-Ing. Klaus Diepold

Die Dissertation wurde am 28.07.2008 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 01.04.2009 angenommen.

Solutions for Scalable Communication and System Security in Vehicular Network Architectures

Dipl.-Ing. Univ. Claus Stephan Eichler

Zusammenfassung

Kommunikationstechnologie hat in den letzten Jahren in viele Bereiche des täglichen Lebens Einzug gehalten. Mobiltelefone, kabellose Internetverbindungen und das so genannte Überallfernsehen sind nur einige Beispiele für die zunehmende Verbreitung moderner Kommunikationstechnik. Dieser Trend, vor allem aber die Verfügbarkeit von leistungsstarker und integrierter Kommunikationselektronik ermöglicht neue Anwendungsgebiete. Ein Beispiel dafür sind Fahrzeugnetze. In einem Fahrzeugnetz wird Kommunikationstechnologie zur Funkkommunikation zwischen Fahrzeugen verwendet, um aktuelle Kontextinformationen dezentral auszutauschen. Zusätzlich wird die Anbindung an Servicedienste realisiert. Das Konzept der Fahrzeugnetze verspricht sowohl die Sicherheit als auch den Komfort für Fahrzeuginsassen zu steigern. Derzeit existieren jedoch noch viele offene Fragen im Zusammenhang mit dieser Technologie. Insbesondere das dezentrale Ad-hoc Netz, welches einen zentralen Teil der Fahrzeugnetze ausmacht, stellt eine Vielzahl von Anforderungen an die Kommunikations- und Sicherheitstechnologien. Die in dieser Arbeit präsentierten Forschungsergebnisse liefern einen Beitrag sowohl zur Kommunikation als auch zur Informationssicherheit in Fahrzeugnetzen. Neue Verfahren und Konzepte zur Realisierung und Integration von Kommunikationstechnik und Informationssicherheit, sowie deren Verknüpfung, werden vorgestellt und bewertet. Für die Fahrzeug-zu-Fahrzeug Kommunikation werden drei unterschiedliche Kommunikationsstrategien vorgeschlagen und analysiert. Dazu gehören geographische Nachrichtenverteilgebiete, kontextbezogene Paketpriorisierung und Datenverteilung auf Anforderung. Des Weiteren wird die Festlegung von Vertrauen und ihre Absicherung evaluiert. Insbesondere wird die Leistungsfähigkeit einer Public-Key Infrastruktur im Kontext von Fahrzeugnetzen simuliert und bewertet. Diese Analyse der Vertrauensbasis dient als Basis für mehrere Sicherheitskonzepte, die für den Einsatz in Fahrzeugnetzen vorgeschlagen werden. Dabei werden zentralisierte und dezentralisierte Dienste berücksichtigt. Zum Schutz der Privatsphäre wird der Einsatz von Pseudonymen diskutiert und wichtige Konfigurationsparameter für den Pseudonymwechsel identifiziert. Die Ergebnisse zu Kommunikations- und Sicherheitstechnologien werden abschließend in einer Architekturlösung kombiniert. Das vorgeschlagene Architekturkonzept unterstützt sowohl dezentralisierte Fahrzeug-zu-Fahrzeug Kommunikation als auch zentra-

lisierte Dienstbereitstellung. Die Beiträge dieser Dissertation werden das Verständnis der komplexen Herausforderungen, die in Fahrzeugnetzen existieren, unterstützen. Außerdem können die vorgeschlagenen Lösungsbausteine zu einer einheitlichen Systemarchitektur verhelfen.

Abstract

Communication technology has found its way into many parts of our daily lives. Mobile phones, wireless Internet connections, and mobile television are just a few examples for the increasing prevalence of communication technology around us. This trend and the availability of powerful integrated communication devices leads to new application areas constantly, one prominent example being Vehicular Networks (VNs). In VNs communication technology is used to connect vehicles with each other to exchange up-to-date context information in a decentralized way as well as to provide access to service applications. While this concept promises to increase safety and comfort for future automobile passengers, still many open issues exist today. Especially the ad hoc network part of a VN, the so-called Vehicular Ad Hoc Network (VANET), has multiple requirements on the applied communication and security technologies. The work presented in this thesis tackles both of these challenges. New approaches to realize and integrate communication and security for a VN are introduced and evaluated. Three different communication strategies for the VANET setting are suggested and analyzed: geographic dissemination areas, context-based packet prioritization, and request-based data distribution. The trust establishment and its preservation for VNs is discussed. The performance of a Public Key Infrastructure (PKI) in the given setting is evaluated. This trust evaluation is used as a basis for several different security concepts suggested for centralized services as well as decentralized ad hoc applications. In addition, the use of pseudonyms to preserve privacy in VANET scenarios is evaluated and important configuration parameters are identified. The results on communication and security are combined in an architecture concept for VNs which can be used to provide both decentralized vehicle-to-vehicle (V2V) communication as well as centralized service provisioning by a service provider. The contributions of this thesis will help to understand the complex challenges of VNs and point out possible approaches to come to an unified system architecture.

Acknowledgments

Pursuing research work in the area of vehicular networks at the Institute of Communication Networks (LKN) has been a great experience and a challenging start into my professional career. I am very happy that I took the chance to write my thesis at the LKN, it has been a wonderful and prosperous time. The research results of almost five years have left their marks on this thesis. During that time I've worked together with many different people, which contributed to this thesis in form of discussions, reviews, and alternative views on the topic.

First of all I would like to thank my doctoral adviser Prof.-Dr. Jörg Eberspächer, for giving me the chance to write this thesis and supporting my ideas and results. I am very grateful that he gave me the opportunity to be a member of the research team at the LKN. His continuous support and trust in my work at the institute, as well as the encouraging mentoring of my research activities, motivated me a lot and helped to develop this thesis. In addition, I want to thank Prof.-Dr. Klaus Diepold for writing the second expertise for my thesis.

The LKN has a great research team and administrative staff. The friendly and cooperative atmosphere at the institute is outstanding and helped me to stay motivated even on the hard days of the long journey. In return I want to thank all my colleagues at the LKN. Specifically I would like to express my gratitude to Ingo Gruber, Christian Merkle, Bernd Müller-Rathgeber, Robert Nagel, Christian Schwingenschlögl, Robert Vilzmann, and Hans-Martin Zimmermann, for their support and many interesting discussions. In addition, I would like to thank my bachelor and master student Christoph Schroth.

Finally, I want to thank my parents and my wife Birgit. Their patience and understanding has helped me to go through with my research activities and finalize this thesis. Thank you very much for your continuous support.

Contents

List of Figures

List of Tables

1 Introduction

Wireless communication technology and its applications have been capturing our daily lives in the last couple of years. Starting out in the mid 1990's with mobile phones using the Global System for Mobile Communications (GSM) technology [EV99] more and more people were capable of communicating anywhere. This trend continued and with the introduction of Wireless Local Area Networks (WLANs), using the Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 access technology, it broadened to an "always on" philosophy – being connected anywhere at any time usually using a digital broadband access technology. The availability of new communication means strongly influenced related research activities and led to many new trends and technologies.

The possibility to exchange data without a cable connection, using a digital wireless transmission technology makes new network concepts feasible. Especially the idea of decentralized and fully uncoordinated network scenarios can be realized this way. In addition the wireless connectivity allows the nodes to move while being an active part of the network. This network scenario, also known as Mobile Ad Hoc Network (MANET), became a very popular research area during the last ten years. Many different protocols and services have been suggested over the years. One very specific application area of the MANET concept is for Inter-Vehicle Communication (IVC). In the literature the term Vehicular Ad Hoc Network (VANET) is applied for a MANET whose nodes are vehicles which use the communication technology to exchange, for example, traffic safety information. Using wireless communication to realize IVC makes many new application settings possible and promises higher safety and more comfort for the passengers. Traveling can be made increasingly comfortable and much safer using IVC. However, before IVC can be realized and is beneficial to vehicle passengers many open issues concerning communication, security, and architecture concepts need to be resolved.

In this thesis several contributions addressing IVC topics are presented. New approaches and solutions are introduced for all three major topics of IVC: Communication protocols and strategies, security and trust concepts, and architecture design. All contributions can be combined to realize a comprehensive architecture concept for IVC and telematics services.

1.1 Introduction to Vehicular Networking

As early as 1989 the first vehicle-to-infrastructure (V2I) communication system has been proposed in [TTIF89]. In their survey article on intelligent vehicles Collier and Weiland presented in 1994 the first ideas on Vehicular Networks (VNs) and their influence on traffic management and traveling in the future [CW94]. Many of their ideas are still challenging issues today. Cox suggested that personal communication will eventually be extended into the vehicle and provide new types of services to passengers [Cox95]. This brief view on the literature shows how long the topic of vehicular networking has fascinated the research community. Mainly the growing availability of wireless communication technologies boosted the research activities on VNs and helped to generate new ideas and advance existing visions. Starting out with the vision of V2I communication presented in [TTIF89, CW94] and the concept of vehicular telematics services [Zha02], which can provide a variety of services to passengers, the availability of new technologies advanced the vision to include ad hoc communication between vehicles [CKV01, EU02]. This vision brought up many open issues related to communication, security, and architecture concepts. Hence, several national and international research projects started to look into the issues of IVC [Fra04].

The availability of wireless communication technologies on the one hand and the need for new traffic safety solutions on the other hand were the most important factors to trigger research on VNs. In addition, the multiple possibilities for a commercial exploitation of a VN system helped to motivate the vehicle manufacturers to participate in and fund many research activities in the last years. The concept of vehicular networking promises to increase vehicle and traffic safety while still ensuring commercial interests of the Original Equipment Manufacturers (OEMs).

To realize a VN, vehicles need to be equipped with communication technology. The amount of up-to-date information available to a vehicle and its passengers can be significantly increased by having one or more communication devices integrated into the vehicle. Moreover, a large variety of services can be provided using these devices, for example, co-operative driver assistance, traffic control mechanisms, information services, and entertainment applications.

1.1.1 Examples for Vehicular Network Services and Inter-Vehicle Communication

Safety Services: To increase traffic safety several services have been discussed in the context of VNs. Intersection assistance, accident and general local danger warning, and severe weather warnings are some examples for safety services. They mainly rely on IVC and belong to the driver assistance services.

Traffic Control and Toll Services: Infrastructure like road signs can be equipped with communication technology as well. Hence, the signs can actively disseminate warning information to surrounding vehicles. In addition, vehicles can detect traffic conditions like congestions and provide this information to other vehicles. This information can be used to realize active navigation services. In addition, V2I communication can be used to handle and charge toll road fees.

Information and Entertainment Services: The communication devices can be used to provide all kinds of information services to the driver and the passengers of a vehicle. Typical examples are information on parking spaces, tourist attractions, the closest gas station, or the next dealership. Moreover, entertainment services will be offered in a VN. Thus, passengers will for example be able to buy and download music, reserve seats at the next movie theater, or book a hotel room.

Web-Services: An OEM providing VN technology for its vehicles will provide a fully integrated service concept. Therefore, the customer can configure vehicle settings, for instance radio stations, and services at the PC. The configuration will be updated at the next start of the vehicle's On-Board Unit (OBU). In this respect, better vehicle servicing can be offered. The system reminds the owner of the next service interval and can inform the dealership of the required parts. Additionally, software updates and support data like digital maps can be uploaded to the vehicle using a Web-service on the PC.

Autonomous and Cognitive Vehicles: In the future vehicles will have extended cognitive features and will be able to drive autonomously in a cooperative way. Vehicles will rely on a great variety of sensor and camera information to analyze their neighborhood, recognize the current traffic situation, and generate an appropriate behavior. This approach is studied in the German research project "KogniMobil". It strongly relies on communication between the cognitive automobiles to enable a cooperative behavior and use information from surrounding vehicles. For cognitive vehicles very specific communication needs exist, requiring highly adapted mechanisms which provide strong real-time capabilities and Quality of Service (QoS). An introduction to the wide field of cognitive automobiles and their communication needs can be found in [NEE07, SFK07, GAB+08]. The results on communication mechanisms and security presented in this thesis can be seen as preliminary results and starting point for the communication mechanisms needed for autonomous vehicles.

1.1.2 Advantages and Risks of Vehicular Network Services

The use of VNs provides many advantages for the customers as well as the vehicle manufacturers. The users benefit from the large variety of new service applications and a significant increase in vehicle and traffic safety. Traveling will become more comfortable and relaxed for vehicle passengers. The vehicle manufacturing process can be improved using the communication devices to identify the vehicles. Moreover, dealerships will be able to connect to the vehicle without a cable connection. Hence, the on-board diagnosis can be improved. In addition, new products and service subscriptions requiring the vehicle's connectivity can be sold by the OEM, creating a new business case.

Besides the manifold advantages also risks and difficulties exist in coherence with VNs. Especially the safety services need to be very reliable and trustworthy, since they influence the driver's behavior on traffic situations. The accountability for a system malfunction needs to be clear otherwise legal conflicts might arise. A big risk is the currently unclear customer acceptance. If not enough customers are willing to buy an IVC device or are even interested in the provided features, the whole concept will not succeed.

Besides these risks also several technical issues need to be resolved before VNs become reality. These open issues include for instance IVC schemes, security concepts, and system architecture design topics. Some of them are addressed in this thesis.

1.2 Main Contributions of the Thesis

This thesis contributes mainly in the areas of communication mechanisms, security and privacy strategies, and architecture design for VNs. More precisely the contribution is as follows:

Communication: New concepts for message diffusion in distributed wireless environments with mobile nodes are suggested. This includes especially the use of benefit functions to determine the utility of a given packet, as well as concepts to request information in VANET scenarios supported by gateway nodes. Additionally the limitations of the designated communication standard for VANETs, IEEE 802.11p, have been analyzed.

Security and Privacy: Analysis of Public Key Infrastructure (PKI) concepts and revocation mechanisms in distributed network environments like VANETs show the limitations of different approaches. Based on these results a trust setup for VNs is suggested. Additionally, contributions to the field of secure message exchange are made, using certificates and message reputation techniques. The Content Reputation System (CoRS) is suggested to increase content reliability in IVC systems. In the field of node privacy the use of pseudonyms in VANETs is analyzed.

System Architecture: Based on the results in the fields of communication and security a system architecture concept for VN environments is proposed. This concept includes the Internet support architecture as well as the Mobile Entities (MEs) and their interactions. The suggested architecture can be used to design respective systems in the future.

1.3 Organization of the Thesis

The following chapters of the thesis are organized in a similar fashion. Before the contributions to the respective area are presented, an overview on the respective related work is given. Each chapter closes with short individual conclusions.

Chap. 2 gives the motivation for the presented research work and explains the most important terms and definitions used throughout the thesis. Based on a short introduction of the typical VN scenarios and potential business cases, the building blocks of the scenarios are outlined. Besides the component description, the requirements for the communication architecture and the desired security functionalities are presented. Further, the existing research challenges in relation to VNs are outlined.

Chap. 3 outlines the concepts for efficient and scalable communication in VNs. In a first step an overview on the related work is given and the main difficulties of message distribution in VNs are presented. Based on these difficulties several approaches, for instance,

to increase scalability, to prioritize message delivery, and to limit network congestion are introduced. Their applicability and impact are backed by simulation results.

A selection of new security and privacy mechanisms applicable to VNs is discussed in Chap. 4. First, an overview on related work and the main difficulties of securing distributed VNs is presented. To introduce security in an architecture a trust basis is required. A possible solution to this demand is presented as a basis for all of the following security mechanisms outlined in Chap. 4. Using the trust architecture a secure platform architecture for telematics services is presented. Further, the message security for data exchange in a VANET is discussed and analyzed. In this context the concept of information reputation is used. Besides the security aspects of VNs, the necessity for privacy mechanism is established and the use of pseudonyms is analyzed.

Using the concepts introduced in Chap. 3 and Chap. 4 a full system architecture for VNs is outlined in Chap. 5. The architecture components of the support infrastructure and the MEs are introduced and explained. In addition, the interactions between the system components are described.

The thesis closes with a conclusion and a brief outlook on future research challenges in the area of VNs in Chap. 6.

Information on the simulation system and commonly used parameters, which have been applied for most research results presented in this work, are given in App. A.

2

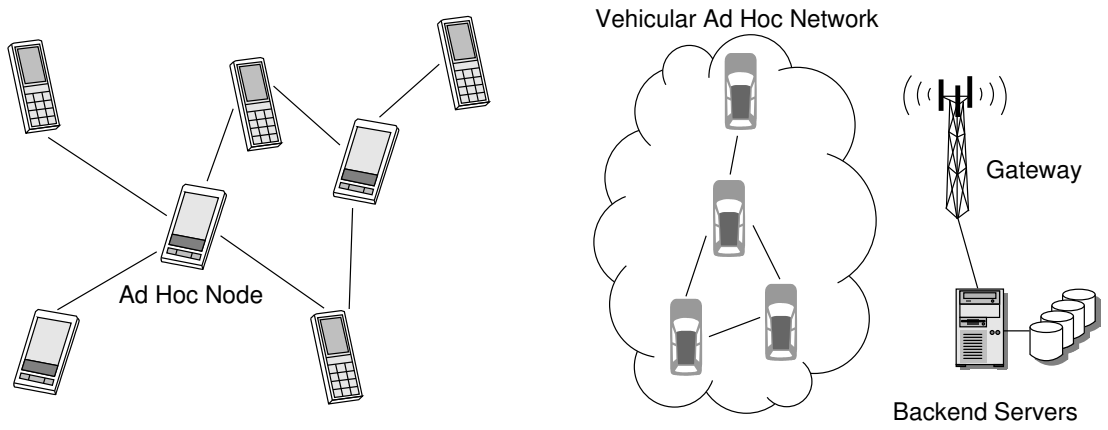
Vehicular Networks: Motivation, Scenarios, and Requirements

Over the last couple of years a new research area connected to wireless networks has emerged – mobile networks. Since wireless technology provides the means to transport data without having a cable connection, this implicates the possibility to communicate to mobile nodes. A prominent example for this evolution are the Global System for Mobile Communications (GSM) networks around the world. But not only operator driven communication advances in the mobile domain, also non-commercial wireless technologies such as Wireless Local Area Network (WLAN) allows users to connect to networks all over the world. In general the network nodes are stationary while communicating and move only when no network connection exists. Allowing the nodes to move while communicating requires handover mechanisms and a close knit access point infrastructure. On the other hand, wireless technologies like WLAN led to a new form of communication network called Mobile Ad Hoc Network (MANET) (see Fig. 2.1(a)). In a MANET generally all nodes are equal peers which can move. The network structure is fully decentralized, hence, no central coordination is used whatsoever.

A special realization of a MANET is a so-called Vehicular Ad Hoc Network (VANET), which basically is a MANET where all the nodes are vehicles. While in many research papers the term VANET is used predominantly, in this thesis also the term Vehicular Network (VN) is used. A VN consists of mobile nodes, usually vehicles, forming a VANET as well as gateways and a server infrastructure (see Fig. 2.1(b)).

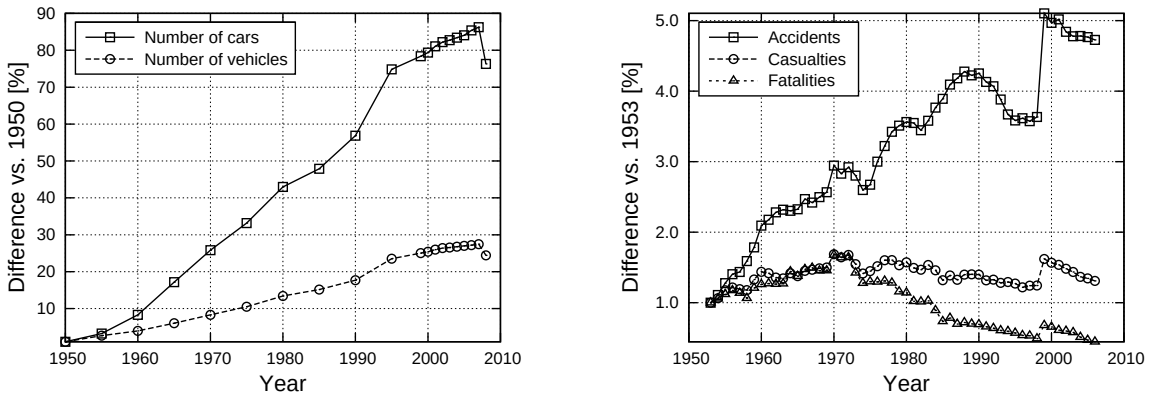
This chapter mainly motivates for research in the field of VNs. First, the terminology used in the thesis is explained. The most common scenario settings are introduced and explained. In addition, potential business models are outlined. In Sec. 2.5 the vision of this work is presented, highlighting the most important research issues connected to the field of VNs.

The statistics for road traffic in Europe and Germany prove that the number of vehicles has strongly increased over the last 50 years (see Fig. 2.2). Due to a rising number of vehicles both traffic density and number of accidents increased as well. However, with the introduction of safety systems like Anti Blocking System (ABS) or Electronic Stability



(a) Ad Hoc Network **(b)** Vehicular Network

Figure (2.1) Mobile network scenarios



(a) Number of vehicles **(b)** Number of accidents

Figure (2.2) Traffic statistics for Germany [Fed07, KB08]

Program (ESP) the number of road fatalities has decreased over the last decades. In a next step more active and especially interactive systems shall be introduced to reduce the number of accidents and the casualties even further.

The European Commission (EC) has set the goal to reduce the number of road fatalities by 50% until the year 2010 [Com03]. Thus, many different research projects have been funded in the context of the e-Safety Initiative [Eur08], prominent examples are Cooperative Vehicle-Infrastructure Systems (CVIS), Global System for Telematics (GST), PReVENTive and Active Safety Applications (PReVENT), and Secure Vehicular Communication (SeVeCom). Research projects at national level have looked into similar topics, projects like FleetNet or Network on Wheels (NoW) have to be named in this context. A new initiative is the pilot "Sichere Intelligente Mobilität – Testfeld Deutschland (SIM-TD)", which will provide a test environment for vehicular networking in Germany in the near future.

2.1 Related Work on Vehicular Networking

Since each of the following chapters has its own section on related work, only a more general overview is given at this point. The first concepts of intelligent vehicles using vehicle-to-infrastructure (V2I) have been outlined in [TTIF89] and [CW94]. The main goal of all these early concepts was to make traveling more comfortable and transportation systems more intelligent. These aims are still valid today.

The first larger European research projects related to Inter-Vehicle Communication (IVC) and telematics were launched in the late nineties. One of the first ones was CarTalk2000, which aimed at co-operative driver assistance systems based on ad hoc communication. In addition, the project looked at legal aspects and market introduction strategies using cost/benefit analyzes. The respective results have been presented in [RMM+02, MC04]. In parallel, a successful and widely-known research project called FleetNet was launched in Germany [Fle08, FEL01]. In the introduction papers on FleetNet the project goals and first results on position-based routing concepts are discussed [HBE+01a, FEL01]. FleetNet addressed the realization of a wireless ad hoc network used for IVC. The papers give insights in requirements for a potential communication hardware for IVC. The hardware needs to support bit rates exceeding 100 kbit/s in a fading channel environment and has to be capable of handling communication durations below 10 s. The hardware has to provide multihop communication capabilities. The project also addressed the issue of Internet integration of a VANET. An overview on the project's contributions has been outlined in [FWMH04]. Besides the multiple research contributions in the area of IVC, mainly communication and routing concepts, also a prototypical implementation has shown the performance of the approach providing safety and convenience applications to vehicle passengers.

One of the first in-vehicle system concepts has been discussed in [EU02]. The authors present their concept how to connect Intelligent Transportation System (ITS) services to the Internet using the Transmission Control Protocol (TCP)/Internet Protocol (IP) protocol suite and especially IP version 6. Within the paper the functional needs are described giving an overview on the early system concepts. The authors addressed the in-vehicle communication platform, the permanent Internet availability to the vehicle, wireless communication technology with fast handover capability as well as vertical handover between technologies, and scalability and flexibility of the system concept and its protocols as the main requirements, some of which are still valid today.

An overview on scenarios and realization issues of situation-based information systems in future vehicles is given in [KSB02]. The authors point out the capabilities of using Dedicated Short Range Communication (DSRC) to realize multihop communication between vehicles. Services like accident warning and infotainment services are discussed. The authors argue that ad hoc routing will play an important role for most applications. They reason the importance of geo-information like positions and present a position-based routing approach using Ad hoc On-Demand Distance Vector Routing (AODV) as a basis. The prototype of the implementation is discussed in detail, giving a first idea on how future system setups can look like.

Besides research in the direction of IVC other communication-based ITS services have been proposed. One example is the so-called Floating Car Data (FCD) service. Vehicles act

as mobile sensors and generate traffic information. This information is made available to other participants increasing the availability of up-to-date traffic information. FCD concepts have been discussed for example in [HLO99, STW02]. The FCD service can be realized with a cell-based communication system, however, it can clearly benefit from an IVC technology, making it to an enabler for both VN communication approaches in parallel.

Several of the early publications on IVC related concepts addressed highway scenarios [CKV01, ZMTV02]. The mobility of the vehicles is seen as beneficial for message distribution. This is analyzed in [CKV01].

The publication by Zarki et al. [ZMTV02] can be seen as the initial publication addressing security aspects in VNs. The authors have presented the Driver Ad Hoc Networking Infrastructure (DAHNI) which is a concept for a driver assistance system in future vehicles. A detailed discussion on security and privacy required for systems like DAHNI has been presented. Based on the system features and requirements, the authors identify digital signatures, the presence of a Public Key Infrastructure (PKI) or a similar trust environment, as well as exact and reliable time information as crucial security requirements. In addition, they suggested access control policies for secure information management. In [Her04] a motivation for security supporting automotive telematics services was given, arguing that road safety can only be increased with secure telematics. The importance of a security and privacy framework for telematics services was also pointed out in [HK02a], listing basic requirements like authentication of users and software, end-to-end (E2E) security, and intrusion detection. More specific ideas were given in [GM02], presenting a first outline of a trust setup for ITS using digital certificates.

The potential of VNs and IVC has been recognized by the EC in 2002 by establishing the e-Safety Working Group in strong cooperation with the automotive industry [Com03, Eur08]. In their report on Information and Communication Technology (ICT) the EC argues why VNs are an important building block for future European transportation systems. The demand for transport services grows continuously, mainly the road transport. The average number of 1.3 million accidents in Europe, causing roughly 40 thousand fatalities and 1.7 million injuries per year, come at the expense of about €160 billion (equiv. 2% of the European Gross Domestic Product (GDP)). To reduce these numbers by 50% until the year 2010, the EC recognizes ICT as an important building block. For this reason, the EC funds multiple European research projects, coordinated by the eSafety Working Group, to develop intelligent vehicle safety systems increasing passive safety, enabling active safety, and provide accident prevention.

Besides the technical aspects, comprehensively addressed by the European research projects, the economical aspects of VNs are very important. Without a valid business case the automobile industry will not invest in the technology on their own. In [Pra03] and [MML04] the economic side of IVC systems was discussed and evaluated. The author of [Pra03] argues that telematics services will increase comfort of vehicle passengers, therefore, the Original Equipment Manufacturer (OEM) integrating these capabilities in the vehicles will be able to earn money with the applications. Further, several existing business alliances are pointed out which support the industries' interests in the topic. An introduction into the economical background of IVC systems, addressing different players (customers, manufacturers, public authorities), was given in [MML04]. The authors point out that no OEM can pursue with

VNs on its own and primarily benefit from it. A cooperation between OEMs and public authorities is necessary to successfully deploy VNs. Two variants for the implementation phase exist: the "visible added value" concept versus a "regulative order". While a regulative order would lead to a large equipment rate within several years, the authors do not see this as a realistic setting. Hence, the manufacturers and the public authorities need to generate sufficient "visible added value" for the potential customers to reach the required 10% equipment rate as fast as possible.

A brief survey on IVC as one implementation part of ITS was presented in [LH04]. After reviewing some of the major research projects in the field, the authors pointed out potential applications, for instance, co-operative assistance systems and information systems. The main contribution of the paper is a brief overview on all aspects related to IVC: Physical layer communication technologies, Medium Access Control (MAC) protocols, routing, group communication, security issues, and simulation aspects. The paper provides a first introduction into the field.

A second overview on general technical challenges, potential solutions, and architectural concepts was presented by Lübke in [Lüb04]. The requirements on the functionalities of different layers and the security of a future system were pointed out. In addition, the author discussed the integration of the new communication technologies into the existing vehicle system stack and the vehicle Application Programming Interface (API). A more recent overview on requirements and research challenges in VANETs was presented in [ESE06]. Besides the need for efficient and information distribution, scalable network protocols, and information security, the authors pointed out the potential of future vehicle-to-vehicle (V2V) and V2I communication systems.

The direction of the current research activities in the field of VNs can best be seen by reviewing recent publications. Publications like [ACG+07, KRM07, FSTE07] give an up-to-date overview on the history of VN research activities as well as current and future research challenges. Some of the main open questions today are security and privacy solutions, MAC protocols for V2V communication, and system architecture concepts. Moreover, the scalability and flexibility of system concepts and protocols is one of the ultimate design goals in the field, which is still not fully solved.

2.2 Terminology and Terms

2.2.1 Networking and Communication

Backend: The term *Backend* refers to all architecture components located in the fixed network part of the system design. Hence, all servers and databases, mainly needed as a prerequisite for the whole system to run, compose the Backend. Typically, these components are connected via the Internet.

Benefit and Utility: In Chap. 3 the terms *benefit* and *utility* are frequently used. At first sight they seem to denote the same meaning. However, in the context of this thesis the term *benefit* is always used to describe the added value of a single piece of information or context. In contrast, the term *utility* is always used for the combination of several distinct benefit values belonging to the same message.

Mobile Entity (ME): Usually nodes in a VANET are vehicles. Nevertheless, also other devices equipped with the needed wireless communication interface, mobile phones or Personal Digital Assistants (PDAs), could be nodes in such a scenario. Thus, the term Mobile Entity is used as a synonym for all compatible devices in a VANET.

Telematics Platform: All components required to operate centralized services make up a Telematics platform. This includes the On-Board Unit (OBU) in the vehicle as well as the servers in the Backend.

Vehicular Ad Hoc Network (VANET): In this thesis the term is used to refer to decentralized and uncoordinated wireless networks mainly used for IVC. The VANET consists of several independent MEs and can operate without infrastructure support in most cases. Referring to [EEH+07] a VANET corresponds to the "pure ad hoc" network classification.

Vehicular Network (VN): In contrast to the term VANET the term VN is used to describe the full network scenario including the MEs, gateways, and all Backend components. This corresponds to an extended "ad hoc access" network classification [EEH+07], since the Backend servers are relevant for the scenario.

2.2.2 Security and Privacy

All terminology in the context of anonymity and identity management is used according to [PH07].

Privacy: Any kind of information knowledge limitation used to protect a user is a privacy function. In the context of this thesis the term is mainly used to describe mechanisms hiding identities or providing any kind of unlinkability to system external or uninvolved peers. Further it mainly addresses the technical aspects, however, these can be strongly connected to the non-technical aspects of privacy.

Reputation: The term *reputation* denotes a subjective opinion of somebody on something. Aggregating several opinions into a profile expresses the reputation. This can be used to increase reliability and trust in systems.

Security: Cryptographic mechanisms can be used to realize features like confidentiality, authenticity, integrity, and non-repudiation. The term *security* is the generic term for all of these features. In addition, it includes the technical method to generate and manage trust in the system. For most security realizations discussed in this thesis a trade off has been made between realizability, potential costs, and level of security. The higher the level of security the higher the costs. Hence, a trade off is needed providing a proportionate level of security for a reasonable investment.

Trust: The term *trust* can be used for both technical and non-technical contexts. While technical trust is realized with some kind of cryptographical mechanism and/or protocol and the question of trustability can be explicitly answered, this is not the case for non-technical contexts. In non-technical contexts social aspects become very

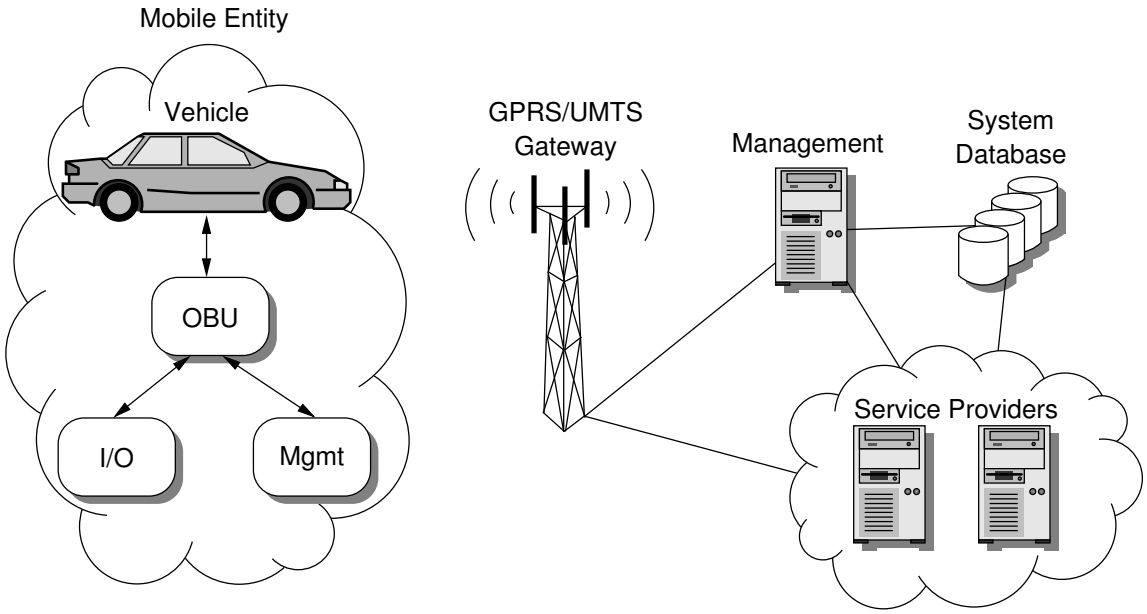


Figure (2.3) Infrastructure-based platform service scenario

relevant. In this thesis the term *trust* is exclusively used within the technical context. A detailed description of trust and its handling in a technical setting is introduced in Sec. 4.2.1.

2.3 Scenarios for Vehicular Networks

Simply equipping vehicles with communication technology is not enough. Applications and scenarios have to be designed using the communication possibilities. But what are typical scenarios for VNs? Even though the corresponding research area is still young, many different scenarios have been proposed in the literature. In the following section a selection of services and use cases is introduced. Further, potential business models are discussed and connected to the use cases.

Two different scenarios using communication technology in VNs are plausible: Infrastructure-based platform services and decentralized ad hoc information services. These two scenarios are completely different, but they can be based on the same requirements. Security is an important feature in both system settings. Further, scalability and performance is crucial in both scenarios. In addition, the system architecture needed for either setting can be complemented to fulfill the requirements of the other scenario.

Infrastructure-based Platform Services: Based on a system architecture with components in the support network infrastructure (also referred to as Backend), the platform services run on a defined architecture in the vehicle using an OBU. Platform services provide individually configured service applications to the end user in the vehicle. Usually these services are provided by one or more service providers in the Backend and

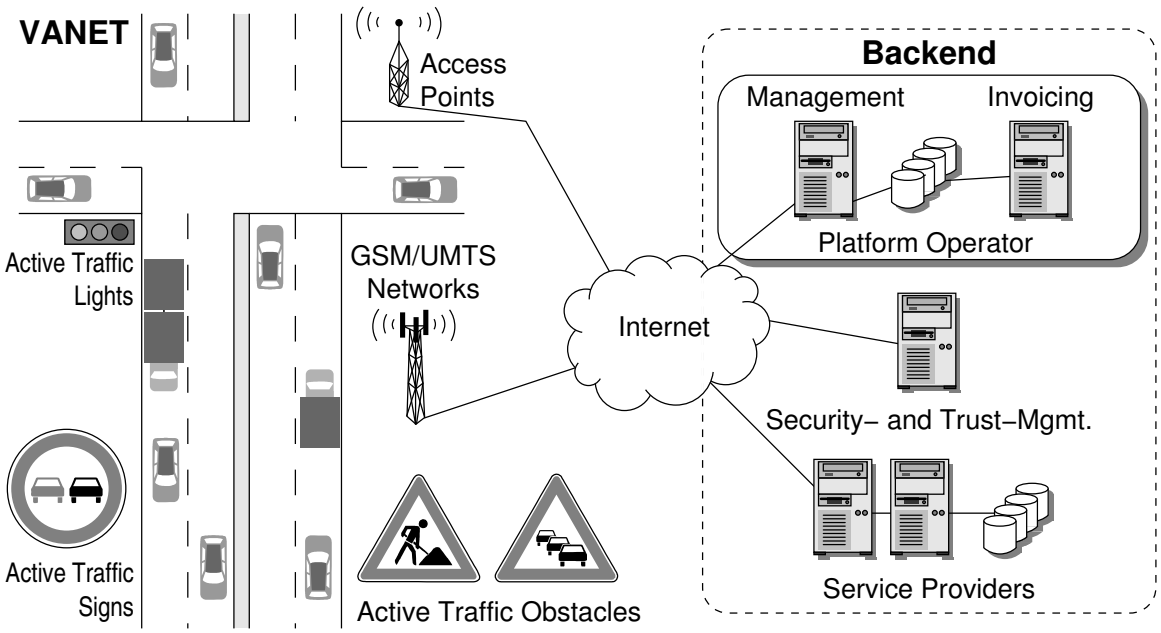


Figure (2.4) Vehicular network reference scenario

are delivered through cell-based communication networks such as GSM or Universal Mobile Telecommunications Standard (UMTS). A typical network scenario for platform services is shown in Fig. 2.3.

Decentralized Ad Hoc Information Services: The architecture concept for ad hoc information services is decentralized, therefore, no centralized servers are required. The service applications run on the ME and try to interact with surrounding MEs on a spontaneous basis. The communication capabilities are provided by a wireless technology similar to the widely used WLAN. The data handling and processing is done by the MEs directly. Central servers can support the scenario, but they are in no way required to operate the decentralized services.

2.3.1 A Vehicular Network Reference Scenario

Both the existing work in the field and the vision of the thesis help to design a typical VN scenario. This reference scenario can be seen as the basis for the research contributions discussed in the following chapters.

Scenario Organization

The setting of the reference scenario is depicted in Fig. 2.4. It can be divided into two parts, the mobile network part (VANET) and the infrastructure part (Backend). The two network parts are connected through wireless communication technologies. The GSM, General Packet Radio Service (GPRS), and UMTS networks are mainly used to provide connectivity for E2E communication links. These are mainly needed for all platform service applications. Access

points, using the ad hoc wireless communication technology of the VANET, are used to provide incidental Backend access in addition to the cell-based networks.

Vehicular Ad Hoc Network Scenario

The VANET scenario mainly consists of MEs such as vehicles. All MEs are equipped with a common wireless communication technology, this could be an interface of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 standard family (for example IEEE 802.11a/b/g/p). All communication is spontaneous and event-triggered, hence, usually no E2E communication relations will be used.

In addition to the MEs, fixed communication entities are part of the scenario. Examples are wireless equipped traffic signs, traffic lights, and obstacles like road construction. Further, automatic traffic control systems can use wireless technology to distribute the current control status to passing vehicles. On the left-hand side in Fig. 2.4 such a VANET setup is shown.

Backend Scenario

The second network part of a VN is the Backend network. The Backend contains the support infrastructure for the VANET part. Servers needed for user and subscription management are located in the Backend. The same is true for the base security functions. The Backend components can be controlled by one stakeholder, nevertheless, multiple stakeholders are more realistic in a real-life scenario. This leads to multiple business connections between stakeholders, which are solely relevant in the Backend. In an optimal setting the users do not have to bother with multiple stakeholders. The open telematics market (refer to Sec. 2.3.3 and Fig. 2.5) helps to operate the system with a single business connection between the user and *one* of the multiple operators.

2.3.2 Services and Use Cases

The use cases for VNs are manifold. Depending on the underlying network scenario, different services can be provided using a VN. Generally, service applications in VNs can be grouped in different categories. Applications for safety, information, and entertainment can be differentiated.

Safety applications are all services which actively or passively help to ensure a maximum level of comfort and safety for vehicle passengers. Examples besides others are collision warning, intersection assistance, and adaptive cruise control. All safety applications that rely on V2V communication have very strict requirements especially on delay and packet loss. In addition, they need to be absolutely trustworthy and intrusion resistant.

The group of information services contains a large variety of services. Their main purpose is to increase the level of up-to-date information mainly for the driver of a vehicle. These services range from mere informational services to warning applications. Examples for information provided by these types of services are road status, weather, parking space information, as well as navigation.

Entertainment applications include all other applications not related to the other two categories. This ranges from simple Internet access over tourist information services to online gaming applications.

Depending on the services to be used in a vehicle, one or both network parts have to be integrated in the vehicle. All service applications which are personalized or require a subscription need the platform service architecture. Most information and safety services can be run with the ad hoc service architecture only.

2.3.3 Potential Business Models

Besides the service applications potentially running on a VN, important considerations are the business cases supporting the economic base of such a system. No vehicle manufacturer has an advantage by simply integrating communication technology into a vehicle. In fact, the business case supporting the integration of such a system in future vehicles is most important when taking the industry's position. The integration of communication technology into vehicles will increase the cost of the product. But it is not yet clear if the customers are willing to pay for this additional feature. This is especially questionable for the first years when the saturation of "equipped vehicles" is low, therefore, the added value for the customer is small.

Hence, the OEMs are somewhat caught up in a dilemma. They have to provide the communication technology in their vehicles, however, none of them has the supporting infrastructure to manage users and services running on such a platform. Moreover, it is not the key competence of an OEM to provide and operate such an infrastructure. Therefore, to be able to provide communication-based services in future vehicles, cooperations are needed between vehicle manufacturers, the communication providers, as well as potential service providers.

In a first step the market has to be opened from both sides: The service providers need easy access to the platform. Both the potential and existing customers wish for the freedom of service provider selection (see Fig. 2.5). This goal has been targeted by integrated projects on European level such as 3rd Generation Telematics (3GT) and GST [Juh03, VVM+04]. An open telematics market is seen as the crucial building block for a successful commercial exploitation of telematics and V2V services. The main goal is to have *one* system where each customer has *one* contract only, hence, all charges are summarized on a single invoice. This goal was targeted by the European project GST [Glo04].

By realizing an open market, a very crucial requirement for the commercially successful exploitation of VN technology is met. In a next step services have to be deployed. We can differentiate between commercial and complimentary services.

Providing improved and advanced servicing capabilities is a first business model for the OEMs. A service technician can analyze technical problems of the vehicle using the wireless communication interface instead of the On-Board Diagnostics (OBD) connector. In addition the wireless interface can already be beneficial during the vehicle production process. The vehicle can be identified and configured over the air.

A promising business model is the possibility of selling service applications to the customers of vehicular network technology. A variety of so-called telematics services can be offered to the customers. One example is the Virtual City Portal (VCP) [BKK+03], which is

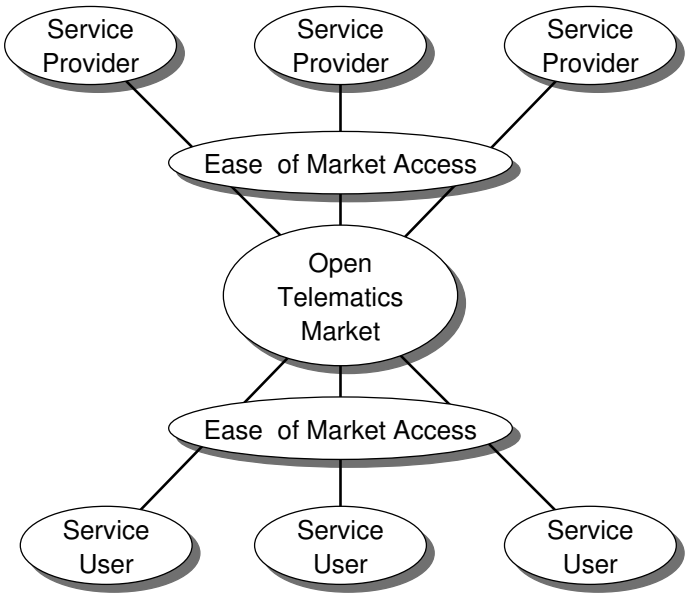


Figure (2.5) The open telematics market [VVM+04]

an information service platform providing contents on tourist attractions, available parking space, and restaurants and hotels to in-vehicle passengers. In this area of the value chain the open market is again very beneficial. Besides offering services on its own, the platform operators can invite external Service Providers (SPs) to sell services to the customers of the platform.

In addition to regular services also road management can be handled by an open telematics platform. One example would be the collection of toll, for example highway usage fees. Using a special service application the vehicle is registered when entering a toll road. When leaving the toll road the application will calculate the traveled distance and the respective traveling costs, which will then be charged to the customers invoice.

It is much harder to define business models for the decentralized communication scenario, however, even for this more or less uncontrolled setting a commercial application can be found. The system can be used to distribute virtual ads, for example, a gas station could distribute current specials using V2V communication.

2.4 Overview on System Components in Vehicular Networks

2.4.1 Components of the Backend Architecture

The Backend architecture consists of three main components. The component organization may be done according to Fig. 2.4, by contrast, other setups providing the same functionalities are equally valid. Moreover, in a real-life implementation not one single operator in the Backend architecture needs to manage all nodes. Practically, several operators will coexist,

similar to today's mobile phone providers. Still, the general setup of these architectures will follow the design shown in Fig. 2.4.

Platform Operator: A very important component in the Backend architecture is the platform operator. It is the main management component. All users are registered and administrated at the operator. In addition the platform operator is responsible for all financial transactions. It will provide an invoicing capability to all service providers and issue the user's invoices. Since it can be expected that many different OBUs will be installed in the market, many different software versions need to exist in parallel. Hence, the platform operator will hold the necessary information on OBUs in the field and provide it to the service providers. Further, it will be able to provide a remote software update feature to its subscribers, keeping the platform up-to-date. The operator holds strong links to the two other main components of the Backend architecture, since it interacts with both security management and the service providers.

Service Provider: The functionalities for the users are provided by the service provider(s). Each service provider can have a whole portfolio of services. Since several operators can coexist a service provider can potentially provide its services to more than one operator, thus, extending its constituency. During the service provisioning phase the provider will strongly cooperate with the operator to provide the correct version of its service for the subscribing clients. The regular service interaction can be done without involving the operator, mainly due to the third Backend architecture component, the security management.

Security and Trust Management: The contractual relations between the system's stakeholders need to be converted into a technical representation which can be handled without human interaction. This is done by the security and trust management components. The trust environment generates a so-called circle-of-trust including all stakeholders and users. It enables seamless interactions between the components. In addition the security is used to provide functionalities like confidentiality of user data or entity authentication. The security and trust management is a crucial component of the system, especially since financial data is handled.

2.4.2 Components of the Mobile Entities

A ME will practically be a vehicle in most cases, thus, the components are very vehicle specific. In other cases the setup of the ME, for example, a mobile phone, needs to be similar to provide the same functionalities.

On-Board Unit: The core component of a ME is the OBU. Basically an OBU is a special computer system installed in the ME. The OBU is the main hard- and software component running all platform- and V2V-related service applications. It requires a defined system setup, especially for the telematics platform services part, to be compatible with the systems in the Backend.

Wireless Interface(s): In order to be able to communicate to surrounding MEs or Backend components the ME requires wireless communication capabilities. To be fully compli-

ant to the VN scenario shown in Fig. 2.4, a ME needs both, a cell-based and an ad hoc wireless communication interface.

Sensors and Actuators: Especially for many IVC applications the sensors of a vehicle are important data sources. The sensors are devices like Global Positioning System (GPS), the speedometer, and the external thermometer. They provide the data basis for many V2V information exchange applications, usually relying on the data of several sensors in combination. The OBU can use the actuators to influence the driving behavior of the vehicle, for example it could initiate an emergency braking due to a collision warning.

2.4.3 Supporting Components in Vehicular Networks

Especially during the early deployment phase of VNs, when only few vehicles are equipped with communication technology, supporting components play an important role. They provide the first functionalities to the early users and help to justify the potential costs for the technology.

Roadside Infrastructure: Any type of roadside infrastructure can be equipped with communication technology and participate in the VN. Hence, road signs, traffic lights, and traffic management systems can use IVC to provide additional services to the early adopters. The OBU can provide the current road sign limitations to the driver as well as the switching frequency of a traffic light the vehicle is approaching. Generally, roadside infrastructure is providing information to the vehicles which can be used within the vehicle to make traveling more comfortable or ease route planning.

Gateway Nodes: In addition to the Backend access through a cellular network, gateway nodes can be used to provide Internet connectivity to vehicles only equipped with V2V communication technology. For certain protocols the integration of gateways is important to provide an occasional Backend access to all nodes. Moreover, the gateways are an important component to realize a fast and widespread information dissemination.

Many other components can be integrated into a VN, however, most of them are very application specific. Therefore, they are not discussed here in detail. Nevertheless, it is very important to keep the possibility of integrating fixed infrastructure components in mind.

2.5 Vision and Research Challenges for this Thesis

The research work presented in this thesis has been guided by a common vision. This vision defined the research challenges and the goals for the thesis.

2.5.1 Vision for Vehicular Networks

Future Vehicular Networks shall be capable to provide both centralized and decentralized services to the vehicle passengers with the following attributes:

Open platform: The architecture for a VN has to be open, which means protocols and interfaces are public. Hence, the realization of new services is not limited to the platform operator. Potentially any service provider can make a new service available to subscribers.

High efficiency and performance: It is very crucial that all services provided by a VN are efficient, thus, they use only a limited amount of resources. Only services providing a maximum added value while using a minimum amount of resources show a high performance and will be successful. Further, the system shall be extensible, hence, multiple applications have to be able to share scarce resources like bandwidth.

Scalability: All protocols and services are operable for low as well as very high node densities without significant performance reductions, meaning they are scalable. Especially the use of the wireless channel needs to be handled with care, to remain operational even if the node density is high (10 or more neighbors).

Reasonable security: All services as well as the platform system itself operate with the necessary security mechanisms and are reasonably well protected, depending on the susceptibility to attacks. Therefore, the degree of security and its complexity and cost need to be balanced.

The security in VNs has to be decentrally usable, on the other hand its organizational structure has to be centralized, in order to allow the operator to control the security of the system. It must be possible to add or remove a subscriber to the security system at any given time. The security integration into the system should be transparent, revisable, and usable for any service running on the platform.

2.5.2 Open and Addressed Research Challenges

The vision presented in Sec. 2.5.1 leads to many different research challenges. They include aspects of the system architecture, means of communication, strategies for data dissemination, and security and privacy integration [LH04, Lüb04].

In the field of system architecture the research challenges include general architecture management, the overall architecture composition, security integration for platform as well as communication security. Moreover, the system integration into existing vehicle architectures is an open issue.

Several different communication means, which could be applied to VNs, exist today. Examples are the conventional IEEE 802.11 WLAN or the newly standardized IEEE 802.11p V2V communication standard. But even though communication technologies exist, many issues and demands in relation to VNs are still not fully resolved. These include scalability related to the bandwidth use in dense network scenarios and time critical message exchange, for example, needed for collision warning applications.

Data dissemination is a new research area which emerged from research areas like MANETs or Wireless Sensor Networks (WSNs). The main goal is to distribute information to all or a group of nodes in the network. This shall be done as efficiently and fast as possible. Especially due to the mobility of the nodes and the limited network capacity caused by

the shared medium, the distribution of data in any type of mobile network is still a great challenge.

Last but not least the realization of security as well as privacy under the constraints described in Sec. 2.5.1 is a big challenge. Even though efficient security mechanisms exist, using them in any kind of distributed environment such as a VN brings up questions regarding their scalability, efficiency, and data overhead. In addition, the distributed setting can cause limitations in terms of security due to the missing constant connectivity to the security support infrastructure. Privacy protocols need to be adapted to the decentralized scenario, taking into account the mobility of nodes in particular.

The integration of all these features into a holistic architecture providing communication for both decentralized and centralized services with appropriate security is the biggest challenge. After all it needs to be adaptable for changes and future requirements.

Obviously, not all open issues can be addressed in this thesis. A selection of aspects is addressed and new approaches are suggested. The areas of research covered in the following are communication and data dissemination in Chap. 3, security and privacy in Chap. 4, and architecture issues in Chap. 5.

2.6 Overall System Requirements for an Efficient, Scalable, and Secure Vehicular Network

A system concept like the one presented in Sec. 2.3.1 and Sec. 2.4 is designed quickly on a theoretical basis. But the practical realization is bound to a variety of requirements. They are mainly based on the related work presented in Sec. 2.1.

2.6.1 Requirements for Communication

In the reference scenario two different types of communication can be differentiated: Session-based communication with dedicated end points and uncoordinated broadcast communication between MEs. These two variants need to be integrated into the system architecture and have to be available for the service applications. This has the main implication that different communication technologies need to coexist and have to be handled in parallel with the same system stack. This infrastructure independent access capability is a crucial prerequisite for the realization of VNs [KVS02].

While the cellular-based communication technologies can practically be used without modifications, this is not true for the ad hoc IVC technologies. Since they make use of a shared medium, a decentralized MAC mechanism will be used, limiting the capacity. Especially in decentralized and uncoordinated ad hoc communication environments the capacity is limited [GK00]. However, the required capacity increases with the number of services and nodes participating in the system. Hence, new communication strategies are required which optimize the usage of the available capacity and meet the high demand in dense scenarios. Moreover, during the deployment phase of VNs the node density can be expected to be low (less than 10% of equipped vehicles). Therefore, the communication mechanisms need to be able to cope with both dense and sparse node densities, providing equal service quality.

An important issue is the parallel usage of different types of services using IVC. The services will have different requirements on packet delay, reliability, and data rate. Therefore, the communication stack needs to be capable of handling different types of characteristics and parameter sets in parallel. This is complicated further due to the distributed system organization. Many nodes can be involved in the distribution process. Thus, uncoordinated distribution mechanisms are required that can be integrated in the split system stack, can cope with the communication requirements and limitations, and scale with increasing number of nodes.

One of the most important requirements is reliability. VNs will only be successful if the services work reliably, which is closely connected to the reliability of the underlying communication technology.

2.6.2 Requirements for Security and Privacy

In addition to the communication requirements also requirements for security and privacy exist. The main goal of using security is to exclude unwanted MEs or even attackers and provide a certain degree of system reliability. Moreover, security is an important component to generate user confidence in the system.

Since the security mechanisms can be used to differentiate valid from invalid peers, they are very important control features for the platform operator. Hence, a fully integrated security framework is required for a VN, giving the operator the control over participation and subscription at least for the centralized services. In this context the protection and secure handling of financial data is very important and an absolute requirement for the trustworthy operation of the system.

Especially wireless communication technologies using a shared medium, possibly in an unlicensed frequency band, are vulnerable to manipulation, eavesdropping, and other attacks. Therefore, communication security, providing message and sender authenticity as well as content integrity, is a very crucial requirement. However, when integrating security features the limited system resources, for instance, the channel capacity, have to be kept in mind. Thus, a very adapted security realization is needed which provides reasonable security with low overhead. Security mechanisms cause overhead in processing capabilities and data consumption. Both should be minimal to keep the system scalable.

The security functionalities are required for many different aspects in the system. Hence, the security implementation needs to be an integral building block of the system. In addition, the hard- and software architecture itself needs to be secured, thus, functions like secure storage and tamper resistance are required to ensure the security even for terminals deployed in the field. To keep the security integration simple and verifiable, it is reasonable to use one system wide security concept with a unified trust architecture. All protocols and components requiring security functionalities make use of this single security framework. Therefore, the security integration needs to be versatile and cover all security aspects of the architecture. After all, the security integration needs to be able to provide privacy features. It is crucial that automatic tracking of users can be prevented with respective privacy features. Unlinkability of MEs and events as well as general anonymity are important features needed in a VN.

3

Efficient and Scalable Communication Mechanisms for VANETs

Using Vehicular Networks (VNs) requires communication means that are adapted to the specific characteristics of these environments. As shown in Sec. 2.6 above, many different requirements have to be taken into account when designing communication mechanisms for VNs and Vehicular Ad Hoc Networks (VANETs) especially. The contributions related to communication strategies in VANET environments are presented in the following chapter. They have been designed based on the reference scenario and its requirements introduced in Chap. 2. While security is an important building block for VNs and is closely connected to communication it will not be in the focus of this chapter. The mere communication mechanisms and possible improvement strategies are presented, but keeping the security requirements in mind.

First of all the main purpose of communication should be in the center of attention. The wireless communication is required to exchange or distribute data between network nodes. So far the VN scenario does not differ from any other scenario. However, the communication in VNs is subject to specific requirements, which call for special communication capabilities. Therefore, characteristics like timeliness, reliability, scalability, and efficiency move in the focus of designing communication mechanisms. Thus, these design goals are important properties targeted in research projects. New strategies are developed to reach improvements in these areas.

Three main approaches have been studied in the course of the research work for this thesis to improve communication mechanisms mainly for the VANET part of VNs. The first step is to adapt the simple flooding mechanisms by using the messages' data content, to limit broadcast ranges or combine equal or similar messages to one single message. A second mechanism is to use prioritization to grant guarantees to specific messages, while disregarding other messages not as important or urgent. The third strategy is to distribute content only when it is requested, which can be beneficial for specific data like status information in a VANET.

The chapter is organized as follows. In Sec. 3.1 the related work in the field is presented and discussed. The main challenges of wireless communication in the context of VNs are

introduced in Sec. 3.2. In Sec. 3.3 the improvement possibilities due to limited dissemination areas and content aggregation are presented. The new Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11p Wireless Access in Vehicular Environments (WAVE) standard promises to solve scalability and throughput issues for VANETs. An evaluation of the standard is presented in Sec. 3.4. The prioritization of messages based on the information relevance can increase the global network utility of vehicle-to-vehicle (V2V) communication. The system concept and an extensive evaluation by simulation is discussed in Sec. 3.5. In contrast to broadcast-based distribution of information also the request-based information distribution promises improvements in terms of scalability, efficiency, and delay. The Mobile Data Request Protocol (MDRP) is introduced and evaluated in Sec. 3.6. The chapter closes with conclusions in Sec. 3.7.

3.1 Overview on Existing Communication Technologies and Protocols

Before the different contributions of the thesis in relation to communication concepts for Inter-Vehicle Communication (IVC) are presented, the most important related work is discussed. Research on communication technologies for IVC has become very popular in the last years, hence, a large variety of publications has been published.

The research on VANET communication has been developed from the research activities on Mobile Ad Hoc Networks (MANETs). That is why many publications on MANET research, for example routing protocols [Bou04, LL99, BMJ+98, GKL04, GE05] or performance aspects and scalability [XKG02, ZR06, HL02, SMSR02], are also relevant for the more specific VANET scenario. Taking the MANET results as a starting point for VANET research, specific topics evolved.

Concepts and Protocols for Information Distribution

Probably the most important feature of IVC is the distribution of information to surrounding Mobile Entities (MEs). Many different approaches and strategies have been suggested in the last years, some of them similar to the contributions in this thesis. The main goals of all distribution mechanisms are scalability, low delays, reliability, and efficiency. Hence, only the desired MEs shall be reached with low packet loss and high speed even in highly populated scenarios.

A first strategy to change conventional routing schemes to a more adapted mechanism for vehicular scenarios was the suggestion to use positioning information obtained by a Global Positioning System (GPS) receiver in networking protocols [NI97]. This allows the realization of geographical services as well as the definition of sending areas. Using this geographical approach a geocasting mechanism has been suggested in [KV99]. The protocol provides a location-based message forwarding scheme and makes use of geocasting areas. This approach reduces the number of messages compared to conventional multicast. An even improved protocol (GeoTORA) has been presented in [KV00], which combines single-cast routing mechanisms with the geocast flooding approach. This hybrid dissemination

mechanism provides a high destination accuracy with few messages and is a very important starting point for following VANET geo-distribution approaches.

The first communication protocols in MANETs as well as VANET scenarios used routing protocols to identify peers and communication routes. An overview and comparison of several MANET routing protocols has been presented in [XKG02]. Besides reviewing the protocols the paper also discusses scalability issues and introduces MANET routing in general. Security issues and a model solution for the MANET routing protocol Ad hoc On-Demand Distance Vector Routing (AODV) have been discussed in [EDS+04]. The MANET routing schemes have soon been extended to transport mechanisms for data between vehicles. In [FMH+02] a comparison of different routing approaches used in vehicular environments has been presented. The simulation results proved that regular MANET routing protocols perform poorly in the vehicular context, while position-based routing has great advantages and a much higher performance. Besides demonstrating the need for new mobility models, which are more suitable for the specific V2V scenarios, the paper identifies parameters to optimize future message dissemination protocols. These include the radio-range, the beaconing frequency, and the general distribution strategy.

One of the first adapted message distribution protocols for IVC has been presented in [FWK+03]. The Contention-Based Forwarding (CBF) is a position-based forwarding scheme using a new distributed channel contention scheme, which reduces the load on the channel. Only the best suited node wins the contention and suppresses the surrounding nodes' forwarding action with its packet sending action. The protocol reduces packet duplications and collisions, hence, increasing the scalability and throughput of the protocol.

A similar approach has been followed for the Smart Broadcast protocol suggested in [FZZ06]. This position-based broadcast protocol has specifically been designed for highway scenarios and uses a contention mechanism similar to the Ready-to-Send (RTS)/Clear-to-Send (CTS) concept known from the IEEE 802.11 standard. The waiting period in the contention process is selected relative to the node density, thus, regulating the collision probability accordingly.

Within the FleetNet project a distributed traffic information distribution scheme using a Universal Mobile Telecommunications Standard (UMTS)-based IVC approach (SOTIS) has been suggested [WER+03]. Each vehicle collects information and provides it to neighboring vehicles. Using the continuously growing knowledge base, related contents are aggregated to reduce packet sizes. In order to keep the message distribution scalable for large scenarios, an information degradation relative to the distance to the event is introduced. Hence, with increasing distance, less information on a distinct event reaches the vehicles.

Analog to SOTIS is the TrafficView concept introduced in [NDLI04]. Similar to the Floating Car Data (FCD) approach the vehicles gather information, for example on the road status. This information is provided to other vehicles using IVC. Within the paper different information diffusion mechanisms using data aggregation and compression are discussed. The main goal is to send as much information as possible in one data packet. Three different aggregation mechanisms are discussed: ratio-based, cost-based, and time-based aggregation, all of them being semantic aggregation schemes. Overall aggregation appears to be a very beneficial approach to increase scalability and performance of data distribution mechanisms in V2V scenarios.

The aggregation of information for an effective distribution has first been discussed in the context of sensor networks, where the channel and energy limitations are much more firm than in VANETs. Several approaches for MANET and sensor network scenarios have been suggested in [SWP03, MR04, PSP03, SOC04]. An important feature for data aggregation is security. In the papers several security concepts have been introduced, which could also be applied in VANETs, especially the secure aggregation using Merkle hash trees presented in [PSP03]. The content pieces are hashed. The hash values represent the leafs of the tree and are combined and hashed in pairs to compute a root hash value. The root value is signed and can be used to authenticate any leaf.

An important issue in V2V networks is the spacial data propagation. In [WFR04] an analytical evaluation of information distribution has been presented. The authors verified their analytical model via simulations and identified parameters which influence the data propagation more than others. These are vehicle density as well as average and relative vehicle speeds. Although simple assumptions for one-way traffic have been used, the evaluation is still an important result to keep in mind for information distribution mechanisms in VANETs.

A mobility-centric approach for data dissemination has been suggested in [WFGH04]. The concept uses the node mobility for the benefit of the distribution mechanism in combination with trajectory-based and position-based forwarding mechanisms. This helps to distribute messages with lower delays to the designated areas.

An approach to optimize safety message distribution in the data link layer has been presented in [XMKS04]. In the paper distribution requirements for different message classes are discussed and used to optimize the protocol approach. The authors suggest a Medium Access Control (MAC) layer extension to reduce the probability of reception failure. The main methodology is carrier sensing and the distinct repetition of the message sending process. The approach presented is one of the first trying to optimize message distribution in Dedicated Short Range Communication (DSRC) systems using a link layer extension.

A system design combining network layer and application layer mechanisms to improve data forwarding in VANET scenarios has been suggested in [TMFH06]. Two approaches, packet-centric and information-centric data forwarding, are discussed and jointly used with strategies like aggregation and data modification in the suggested system concept. Especially the information-centric forwarding is beneficial for the protocol scalability, while aggregation reduces network load. The results presented in the paper strongly support the idea to use hybrid system concepts with optimization in several Open Systems Interconnection (OSI) layers and a cross-layer approach. The concept has been further evaluated and compared with other concepts in [TM07b], additionally two mechanisms using power control and contention strategies are suggested to realize a traffic shaping by prioritizing relevant messages. Moreover, an analysis of DSRC wireless technology potentially used in future V2V networks is presented in [TM07b].

Besides the broadcast-based data dissemination schemes a second information distribution approach is useful for VANETs, namely distribution on request using a content-aware data request protocol. So far this approach has not yet been evaluated in the context of V2V networks. However, in Sec. 3.6 a respective protocol concept is suggested. A data request protocol is practically based on a routing protocol. Very important is the use of a

position-based routing approach, for example Greedy Perimeter Stateless Routing (GPSR). The Zone Routing Protocol (ZRP) is a hybrid routing protocol using routing zones [Haa01]. Its concept or similar approaches can be used to realize a data request protocol for VANETs. The idea to make networking protocols content-based has first been introduced in [CW01]. Nodes are no longer addressed by unique network addresses, however, so-called receiver-predicates, identifiers for the data, are used. A content-based routing protocol following the concepts of [CW01] has been suggested in [CRW04]. The protocol uses broadcast trees based on the predicates of content-based networking to forward requests and replies. The routing protocol does exchange routing information just like most known routing protocols, replacing network addresses with predicates bound to content.

Communication Technology Evaluations

Several wireless communication technologies are candidates for DSRC between vehicles. Starting from the Wireless Local Area Network (WLAN) standard IEEE 802.11 [LAN99] the development has led to the IEEE 802.11p variant in combination with the WAVE mechanisms [Com06, IEE05, Tas06]. All of these technologies have been analyzed and evaluated.

A first simulative evaluation of the WLAN standard has been presented in [CWKS97]. In [HL02] the scalability of the synchronization mechanism used for the ad hoc communication mode of the WLAN technology has been evaluated. It was shown that the original algorithm does not scale without complications, however, newer variants of the standard are practically not affected by this problem, they have other scalability issues such as capacity problems [GK00].

Several optional extensions have been suggested for the popular IEEE 802.11 standard family. An important one is the IEEE 802.11e Quality of Service (QoS)-aware MAC layer extension [Sta05]. A short introduction into the extension as well as an analytical and simulative evaluation has first been discussed in [ZC03]. Assuming saturated queues the analytical model can be used to calculate the throughput of the QoS extension. A very similar evaluation approach has been presented in [TP06], which additionally supports the estimation of the service-delay distribution. Since the Enhanced Distributed Channel Access (EDCA) mechanism of the QoS extension can be adapted by changing several parameters for the contention calculation, an optimized parametrization can be found. Such an optimization, maximizing the throughput, has been presented in [BV06]. The WLAN QoS extension is part of the new WAVE multi-channel concept, hence, its performance has a significant influence on the designated V2V communication technology. A first evaluation of the QoS extension in a VANET scenario has been presented in [TMJH04]. The priority-based channel access is evaluated using a network simulator with two different channel models. The results of this evaluation support the use of priority-based access methods, especially for a saturated medium.

Besides the research activities on the conventional WLAN communication technologies and their application to IVC other concepts have been suggested in the literature. In [MLS05] a multi-channel communication concept is suggested which combines scheduled time division access and ad hoc contention based access. This concept allows an integrated use of distributed V2V and coordinated vehicle-to-infrastructure (V2I) communication, however,

a central coordination for the scheduled access is required. In addition, the multi-channel benefits are not used since different channels can not be used simultaneously. An overview on DSRC for IVC is presented in [JTM+06], discussing communication technology as well as connected message handling protocols. The paper discusses the concepts of the IEEE 802.11p communication standard and identifies three important research issues: congestion control, broadcast enhancement, and concurrent multi-channel operation.

Integration and Usage of Gateways in Vehicular Networks

Future VNs will not solely rely on IVC, however, communication with infrastructure components and gateways will be used to provide additional services. That brings up the question how gateway nodes will be integrated into the network and especially how they are advertised to the MEs. The integration and use of gateway nodes in the context of MANETs has been discussed in [XB02], suggesting a combined active and passive gateway discovery mechanism in combination with a routing protocol. A gateway organization system enabling the parallel use of several gateway nodes in a MANET has been introduced in [MB05]. The Internet connectivity is realized by access points connected to the gateway infrastructure, however, no discovery mechanism is used since a regular MANET routing protocol shall be used.

Within the FleetNet project the Internet access from vehicles has also been realized with gateways [BFW03]. In [BWSF03] a service discovery protocol is introduced which can be used to identify gateway nodes in a VANET. The protocol helps to detect and select the most suitable gateway node. It uses a proactive notification mechanism by distributing gateway notification mechanisms. But the paper does not discuss the performance of such an approach.

Fairness, Prioritization, and Context-based Dissemination Concepts

Since it can be expected that future IVC systems will use a communication technology relying on a shared medium, fairness of channel access is an important issue. In addition, high density scenarios cause a saturated communication system, therefore, new concepts like packet prioritization and context-based message handling promise to maximize the benefit of V2V communication and disseminate important messages first.

To increase fairness and improve the medium access a distributed scheduling scheme called Distributed Wireless Ordering Protocol (DWOP) has been suggested in [KLS+02]. Packets shall access the medium in an orderly fashion comparable to an ideal reference scheduler such as First In, First Out (FIFO). The approach provides a certain degree of QoS and fairness. In order to realize the contention triggered by the packet arrival time of all nodes within radio range, status information needs to be exchanged. The packet arrival time information is piggybacked on the RTS messages, which prevents the protocol use for broadcast communication.

The already mentioned CBF protocol is also an approach to introduce higher quality for message dissemination into an IVC system while reducing the load [FWK+03]. Due to the contention mechanism, which selects a short contention period for nodes far away of the previous sender, the number of newly informed nodes can be maximized. In addition,

retransmissions of nodes close to the previous sender are suppressed, reducing the packet collisions on the channel.

A protocol which handles local throughput overload situations with a fairness criterion has been introduced in [TMSH05]. Based on power control, the protocol restricts packet transmission ranges to reduce the interference range and limit the area throughput to a specified optimum. However, the contention for the channel has not been considered in this approach. A more detailed discussion on power control and contention as optimization parameters for IVC has been presented in [TM07b].

Besides technical mechanisms also information-based mechanisms can be used to increase quality and fairness. The idea of situation adaptive content dissemination in VANETs has first been elaborately discussed in [Kos05]. Exploiting the context of the vehicle and that of the incoming messages in combination with a rule-base can be used to set the message priority. Hence, more important messages have a higher probability to be distributed. This concept has been further elaborated in [Str07] as well as in this thesis in Sec. 3.5.

3.2 Challenges of Message Distribution and Promising Improvement Strategies

Information distribution in decentralized wireless networks, such as VANETs, is a challenging task. The existing wireless communication devices like IEEE 802.11 WLAN have been designed for different use cases, hence, they are not optimized for information distribution with message broadcasts. Several drawbacks of wireless communication in general and properties of the VN scenario make information distribution, taking the requirements of Sec. 2.6.1 into account, a difficult task. But several promising strategies exist how these challenges can be tackled.

3.2.1 Challenges and Drawbacks of Wireless Communication for Information Dissemination

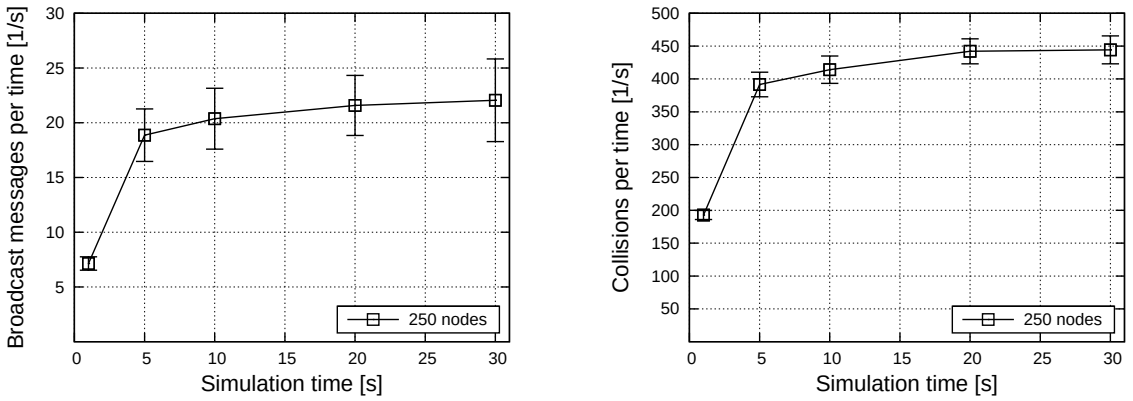
The main limitation of wireless communication is its throughput capacity. In a wireless ad hoc network this throughput (R) is limited to a maximum upper bound [GK00]. In a network with randomly placed nodes (number of nodes (N_n)) this upper bound amounts to

$$\mathcal{O}\left(\frac{R}{\sqrt{N_\mathrm{n}}}\right). \tag{3.1}$$

In a more general network setting using a global scheduling scheme with a randomized traffic pattern this upper bound falls to

$$\mathcal{O}\left(\frac{R}{\sqrt{N_\mathrm{n}\log N_\mathrm{n}}}\right). \tag{3.2}$$

This finding, presented in [GK00], helps to understand the scalability issue of VANETs. The more nodes a network contains the lower is the data throughput. Hence, following the



(a) Broadcast tries per time and node

(b) Detected collisions per time and node on the shared medium

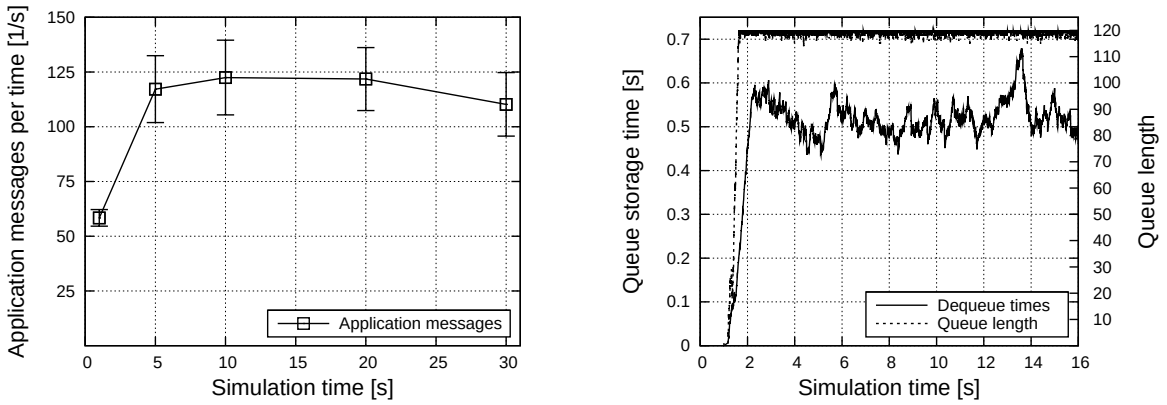
Figure (3.1) Broadcast storm influence on the packet throughput

upper bound of Eqn. (3.1), large networks would have an infinitesimal small throughput. In [LBC+01] Li et al. present that the feasibility of large decentralized wireless networks depends on the data traffic's degree of locality. The authors show that an ideal forwarding chain can achieve 1/4 of the throughput that a single-hop transmission can achieve. Since the IEEE 802.11 MAC is not ideal, it achieves a throughput of only 1/7 in the Network Simulator 2 (NS2) [LBC+01]. Nevertheless, despite the pessimistic upper bound of Eqn. (3.1) large scale ad hoc networks are feasible if the majority of its data traffic can be kept local.

Besides the throughput limitation, a second factor is especially important for the throughput and the successful packet distribution: The number of packet collisions. Due to the decentral network organization in ad hoc network scenarios no global scheduling exists. Therefore, packet collisions occur, the more traffic the more collisions (see Fig. 3.1). Packet collisions are an even bigger limitation than the mere limited throughput capacity, since a collision results in a lost packet which blocked the channel and potentially needs to be retransmitted. The retransmission can only be done if the collision is detected, which is difficult in a setting with single transceiver interfaces. See [KT75, TK75, Rom86, PCS07] for mechanisms of collision detection in wireless networks.

In order to disseminate information, flooding is used in many cases, since it is a simple mechanism to potentially reach all network nodes. However, non-restrictive flooding in a wireless network causes a so-called Broadcast Storm [NTCS99]. To better understand the concepts of limited throughput capacity and packet collisions and their relation to the Broadcast Storm problem, a simple example scenario was simulated. 250 nodes were placed randomly on a simulation area of 1000 m $\times$ 1000 m. A node at the center of the simulation area transmitted a packet with 520 B at $t_{\text{sim}} = 1$ s. Each receiving node forwarded the message unconditionally. This setting caused a Broadcast Storm within seconds. The simulation results are shown in Fig. 3.1 and Fig. 3.2.

The number of broadcasts sent per node, relative to the passed simulation time, are shown in Fig. 3.1(a). The number increases strongly in the beginning and saturates at about



(a) Received messages per time and node

(b) Queue length and dequeue delays

Figure (3.2) Broadcast storm influence on queues, dequeue delay, and reception of data packets

10 s into the simulation. The same is true for the number of collisions that are detected per node, shown in Fig. 3.1(b).

Important for information dissemination is the number of received messages, rather than the number of sending attempts. The result for the received messages is shown in Fig. 3.2(a). The result is similar to the sending events and the collisions, nevertheless, the number of successful receptions decreases after 20 s simulation time. Moreover, about 1300 to 5800 packets are sent every second (e.g. 250 nodes $\times$ 20 Msg/Node), however, only about 70 to 125 packets are received. Thus, the majority of packets is lost due to collisions or queue overflows. The message queue status of a random node in the scenario is depicted in Fig. 3.2(b). The queue of 120 packets is fully filled within one second after the first packet transmission.

This simple simulation example shows the drastic effect of unconditional message forwarding in an ad hoc network environment. It proves that limited throughput capacity, packet collisions, and Broadcast Storm are strong limitations for decentralized wireless networks and have to be tackled in order to realize large wireless VN scenarios in the future.

Besides these three major limitations, several other challenges have to be solved to realize efficient and scalable wireless communication for VANETs. Examples are viable communication destination selection, decentralized network and communication protocol organization, meeting of real-time requirements for specific data, and simultaneous service provisioning of multiple applications using the same interface.

In [Per01, pp. 353] the existing client-server model, known from the Internet, is questioned for decentralized ad hoc environments. This is also a legitimate question for VANET scenarios. Hence, the selection of a communication peer in decentralized scenarios is an important challenge to be solved. New addressing schemes have to be designed, specifically adapted to the decentralized and frequently changing ad hoc environment.

Due to the decentralized nature of VANETs many existing concepts from fixed and centralized networks can not be used. Therefore, the organization of communication

mechanisms and information distribution strategies has to be redefined for this specific scenario. This is associated with the addressing issue mentioned above.

Especially safety services have strict real-time requirements. The packet delay should not exceed 100 ms [MLS05], to provide up-to-date position data for collision warning services. However, if these critical messages have to be exchanged with the same network interface over the same limited channel like infotainment messages, a traffic differentiation is required. The challenge is to find a viable and efficient differentiation mechanism that is adaptable to any packet type and can operate in the decentralized setting.

3.2.2 Possible Strategies to Improve Message Distribution

Before presenting the research contributions for efficient and scalable communication mechanisms the most promising strategies to improve the information exchange between MEs are discussed. The strategies aim at tackling the main challenges presented in Sec. 3.2.1.

Message aggregation and a limitation of the distribution area are strategies to counter the throughput limitation. The aggregation of similar or equal message contents helps to reduce the number of messages, thus, channel capacity is saved. A similar effect is achieved by using dissemination areas. Depending on the information, the broadcast range of a message can be adapted, hence, only MEs in the vicinity will receive and forward a certain message. This reduces the number of messages in the network.

Despite the danger of a Broadcast Storm, message broadcasting will remain an important way to distribute information in decentralized networks. But no unconditional flooding will be used. Therefore, message forwarding rules have to be applied to decide if a received message has to be forwarded and which delay is tolerable. This simple strategy to tackle flooding problems is a preliminary stage towards message prioritization. Message prioritization of any kind (based on size, age, content) is a promising strategy to reduce network flooding. In addition it can help to fulfill strict delay requirements and real-time behavior.

The Broadcast Storm issue can also be addressed by introducing request based information distribution. The strategy is to request information as soon as it is required. This is primarily useful for data that is provided on a regular basis, for example status information. In a network scenario with gateway nodes, like it is the case in a VN, this request-based information diffusion promises to be very efficient and scalable.

Most of the mentioned strategies are analyzed and tested in the course of this thesis. The results of the approaches and the respective network simulations are presented in the remainder of the chapter.

3.3 Improving Information Distribution with Dissemination Areas and Data Aggregation

Limiting message dissemination to a certain area or combining similar or equal content to one single message are promising approaches that can be used to increase the scalability of V2V communication systems by reducing the network load. They can be used in combination

with any underlying MAC and with arbitrary physical communication technologies, which makes them a valuable extension to concepts integrated into lower layers of the system stack.

3.3.1 Limited Flooding with Defined Dissemination Areas

The first analyzed strategy to improve scalability by reducing the number of messages is the use of dissemination areas. The idea is as simple as effective and follows the pseudo code in Alg. 3.1. Only previously unknown messages are processed at all. If the ME is still located inside the dissemination area of the information it will forward the message. Messages not fitting these rules are deleted.

```
1  initialization;
2  while message forwarding active do
3  |  receive message;
4  |  if message is new then
5  |  |  process content;
6  |  |  if position inside dissemination area then
7  |  |  |  forward message;
8  |  |  else
9  |  |  |  delete message;
10 |  |  end
11 |  else
12 |  |  delete message;
13 |  end
14 end
```

Algorithm (3.1) Processing information messages with forwarding areas

This geographically confined information diffusion is similar to the geo-based routing schemes, for example ZRP [Haa01] or GPSR [KK00], where information is targeted at a specified geographical area. In the diffusion case, the information is supposed to be contained in the specified area. Fig. 3.3 shows a simulation result for the use of dissemination areas. Five different messages were distributed in an area of Munich's city center in the depicted simulation example. Each message had its own diffusion area. Fig. 3.3 shows that the messages mainly stay in the designated areas on the city map.

Usually the mechanism of containing messages to certain dissemination areas is not used alone, moreover, a combination of mechanisms is applied. This is also described in [AEK+06]. Most likely an area limited message will have a lifetime, thus, after the lifetime has expired the message will not be disseminated even in the designated diffusion area anymore.

3.3.2 Reducing Messages by Aggregation of Message Content

Due to the multitude of services in a future VANET, many different messages will be transmitted simultaneously. This is one reason for the congestion of the wireless channel. One problem specific to messages generated by vehicles, for example, a traffic congestion

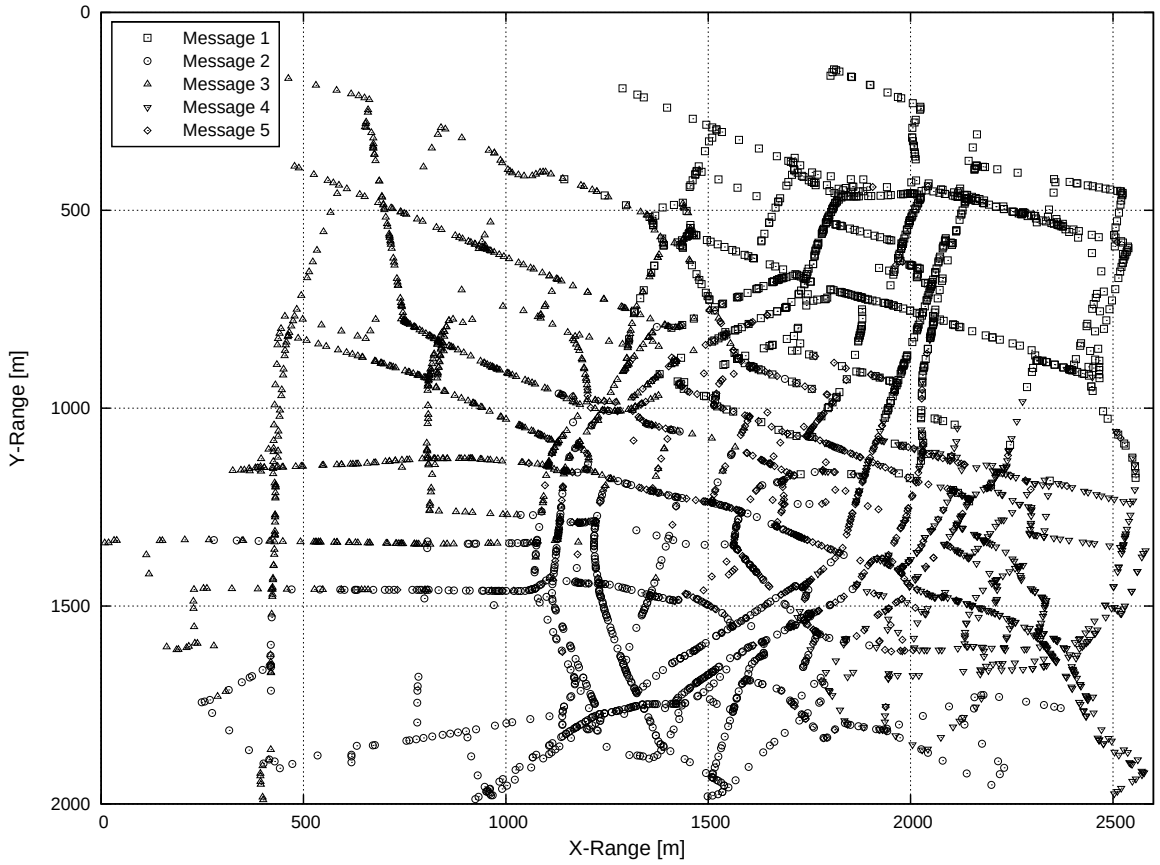


Figure (3.3) Dissemination areas for five traffic incidents [AEK+06]

warning, is the fact that not one but many different vehicles will detect the same event almost at the same time. Hence, they send practically identical warning message to surrounding MEs. Even if a Broadcast Storm protection is used, these different messages with identical content would cause a fully congested network. Therefore, an additional mechanism is required to reduce the network load – data aggregation is one possible approach.

Data aggregation characterizes the combination of information pieces from several different messages into one single message. In place of several messages only one aggregation message will be forwarded. This strategy has several advantages. Besides the reduced channel usage the aggregation helps to apply a weighing of information. More accurate information could get a bigger influence on the aggregate than less accurate information. In addition, aggregation messages carry additional information, for example, the number of single messages contained in the aggregate message. This can be used to rate the reliability of the information, as long as MEs can be identified individually. Further, information aggregation enables the dissemination of event areas, like a bad weather area. Whereas a single ME can only detect one event position at a given point in time, several MEs can combine their information on an event, thus, a whole cloud of event positions is generated. In case of an event area this is a very much desired feature, while the position of an event point could become uncertain. Both the advantages and disadvantages of the data aggregation

strategy are addressed in the concept introduced in the following. The presented results have previously been published in [EMS06].

The Data Aggregation Concept

In a dense network scenario many MEs will detect the same event. Thus, multiple messages are distributed over the VANET, containing practically identical content. The idea to combine these messages leads to the aggregation concept. However, combining messages to an aggregate message can be done in several different ways. In addition, not all types of events are equally suitable for aggregation.

Many different kinds of events exist in a VN environment, for example bad weather warnings, road status information, and local danger warnings, to name a few. Events have different characteristics, they have different lifetimes, may change over time, or have different confidence levels. The aggregation mechanism needs to be closely connected to the type of data, therefore, a classification of events is helpful to design an aggregation system. A more general event classification is adequate to realize data aggregation. Hence, the following event classification is suggested.

1. Static events: Any event that remains static at the same position over time, for example a road construction site.
    a) Infrastructure-based events: Any event that has been send by an infrastructure component, for example a road sign.
    b) Vehicle-based events: Any event that is originated from a vehicle, for example an accident warning.
2. Dynamic events: Any event that changes its position over time, for example a bad weather event.

While aggregation helps to reduce the number of messages, it can however lead to a reduction of information at the same time. Suppose several vehicles detect the end of a traffic jam and send out a warning message containing the position of the event. A potential aggregator has two options: calculate an “average” position or include all positions into one message. While the first option leads to the loss of the individually detected event positions and makes the information slightly more uncertain, the second option leads to a much larger message. A trade-off has to be made for each event class, deciding if the average value is sufficient or not.

A ME can use aggregation as soon as it holds at least two messages describing the same event. In the simulation model each message contains a node ID, an event ID, a time stamp, and a position. These parameters help to differentiate messages and events. In the simulator the differentiation of events can be achieved by simple means, due to the global knowledge. On the other hand, in a real-life scenario this is not the case, since no global knowledge exists. Thus, mechanisms have to be found to match messages to identical events. The matching of events in VNs has not been inquired in this thesis.

Before messages are aggregated their information has to be weighted depending for example on the timeliness and variability of the information. This is crucial for highly dynamic events especially. The newer a piece of information the higher is its influence on

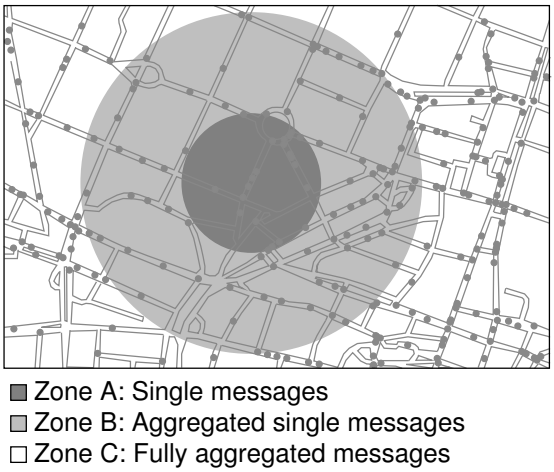


Figure (3.4) Three different data aggregation areas [EMS06]

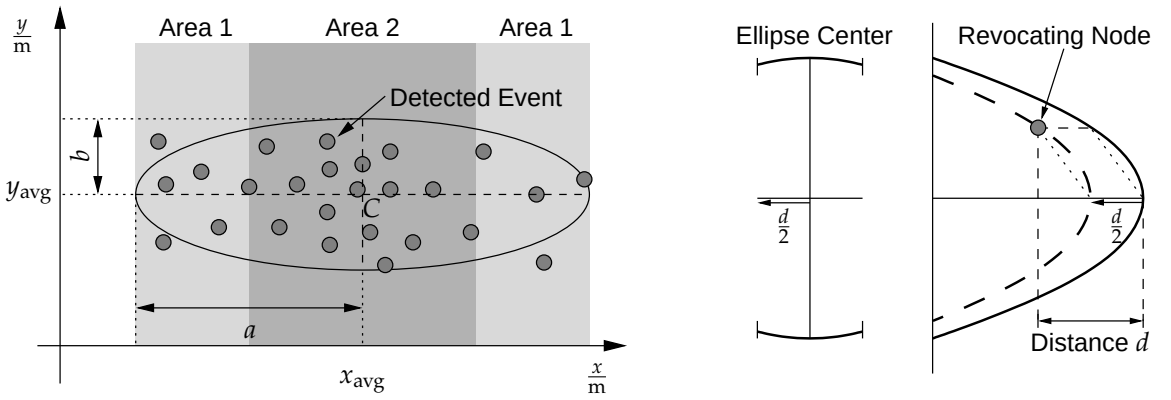
the aggregate message. This ensures that old and therefore most likely invalid information can not dominate the information of the aggregate message.

Aggregation is not reasonable in all areas around an event. Hence, different steps in the information distribution take place in defined regions around an event (see Fig. 3.4). Three aggregation areas are suggested. Individual, not aggregated messages are only distributed in the close vicinity of an event (zone A). Vehicles in this inner circle handle individual messages and generate aggregate messages. In zone B only aggregate messages are processed further. In this zone aggregation of aggregate messages can also be used to further reduce the number of messages. Finally, in zone C aggregate messages are simply disseminated as long as they are not too old for the corresponding event category. The use of the aggregation areas helps to reduce the number of aggregators for individual messages. This is desired to reduce the aggregation of identical messages, which can distort the information significantly. The distribution of messages in zones can be realized by using geographical zone information as well as a simple hop counter.

Another important mechanism of the aggregation concept is the use of revocation messages, which is mainly useful for dynamic event areas. If a vehicle enters a potential hazard area, but can not verify the event, it can generate a revocation message to correct the area information of the corresponding event. This revocation information is included in the event aggregate message like a normal information message. Nevertheless, as soon as the number of revocation messages dominates the aggregate, it is no longer distributed.

Modeling of Hazard Areas using Ellipses

Ellipses can be used to efficiently aggregate position information for hazard areas [EMS06]. Using the individually detected event positions an average position $(x_{\text{avg}}, y_{\text{avg}})$ of the event area is calculated (see Eqn. (3.3) for calculating x_{avg}). This position is used as the center (C) of the hazard area (see Fig. 3.5(a)). In a second step the dimensions of the area are determined. This is done by calculating the distances of all individual positions to the center of the area.



(a) Hazard event area modeled with an ellipse

(b) Revocation-based ellipse adaptations

Figure (3.5) Event area modeling using ellipses [EMS06]

The maximum distances in the direction of the x- and y-axis are used as the semi-minor and semi-major axis a and b (see Eqn. (3.4) and (3.5)).

Due to the use of revocation messages the parameters of the ellipse have to be altered. To modify the ellipse while keeping it as large as possible, therefore, minimizing the error, the ellipse is divided into two areas (see Fig. 3.5(a)). All events positioned in $x < x_{\text{avg}} - a/2$ or $x > x_{\text{avg}} + a/2$ are located in area 1, all other events are located in area 2. In case the revocation event is located in area 1 the semi-major, otherwise the semi-minor axis is altered. The position of the revocation event is assumed to be located on the circumference of the new ellipse. The new axes are calculated using Eqn. (3.6). The respective axis is reduced by half the distance d of the vehicle's position to the circumference of the old ellipse. To maintain the symmetry the center of the new ellipse has to be moved by $d/2$ accordingly (see Fig. 3.5(b)).

$$x_{\text{avg}} = \frac{1}{N_{\text{n}}} \sum_{i=1}^{N_{\text{n}}} x_i \tag{3.3}$$

$$a = \max_{i=1}^{N_{\text{n}}} |x_{\text{avg}} - x_i| \tag{3.4}$$

$$b = \max_{i=1}^{N_{\text{n}}} |y_{\text{avg}} - y_i| \tag{3.5}$$

$$1 = \frac{x^2}{a^2} + \frac{y^2}{b^2} \tag{3.6}$$

Simulation Model and Simulation Parameters

In contrast to all other results of the thesis the aggregation simulation model has been integrated into NS2 [Inf08]. Since the work has been carried out in cooperation with BMW, their proprietary traffic simulator CARISMA has been used to model the vehicle's mobility. CARISMA uses a microscopic traffic model similar to [Kra98] and places vehicles on a digital city map provided by the Environmental Systems Research Institute (ESRI). The coupling

of the simulators has been done according to the concept introduced in [EOSK05]. This selection of simulators has been chosen since it offered a realistic mobility even on a large simulation area, while in contrast having a not so sophisticated radio wave propagation model. The focus of the simulations was to prove the general benefits of aggregation in a VN, rather than the evaluation of aggregation with a specific DSRC technology.

The aggregation model was implemented in the application layer in form of an ad hoc agent. The agent can detect events as well as generate event messages, aggregates, and revocation events. The model uses a simple store-and-forward message dissemination mechanism. Incoming messages are evaluated and stored, while only the new messages are forwarded right away. In case an event message from the memory is still valid but has not been received for a certain time period, the stored message is broadcast again. If the agent has collected several messages related to one event an aggregate message is generated. Depending on the distance to the event position a wait-time is followed by the agent before broadcasting the aggregate. The farther the ME is away from the event position the longer is this wait-time. This ensures that other messages related to the event shall also be received and included into the aggregate message.

The simulations were done using a simulation area of 8 km^2 and 220 vehicles. The communication range was set to 400 m and a simulation duration of 1000 s was used. A hazard area was randomly placed within the simulation area. The shape of the hazard was chosen to be an ellipse to be able to better evaluate the functionality of the aggregation mechanisms. The vehicles check periodically if an event is detected. As soon as the hazard is detected a warning message is generated. Both individual and aggregated warning messages had a size of 200 B. The lifetime of aggregate messages was set to 200 s.

To evaluate the aggregation mechanism two different scenarios have been simulated: One with a static hazard area and another using a moving hazard area. To verify the detection result and evaluate the reliability, three intersecting planes have been calculated and evaluated (see Fig. 3.6):

- Not detected rain area,
- falsely detected rain area,
- correctly detected rain area.

The evaluation of the three areas has been done at the end of the aggregate's lifetime.

Simulation Results for Static and Dynamic Events

The argument for aggregating messages is to reduce the number of messages. In Fig. 3.7 the overall number of messages during the simulation runs for static and dynamic events is shown. The plot shows that the number of messages can be reduced significantly for both event types. For static events the largest reduction is about 48%, in case no revocation is used. If the revocation mechanism is activated the reduction of messages is about 26%. Nevertheless, the information of the aggregate messages is more reliable, as Fig. 3.8 shows. For dynamic events the degree of reduction is equivalent, about 50% for static hazard areas and about 22% for dynamic events can be achieved (see Fig. 3.7). The magnitude of the reduction effect strongly relates to the size of the aggregation zones introduced in Fig. 3.4.

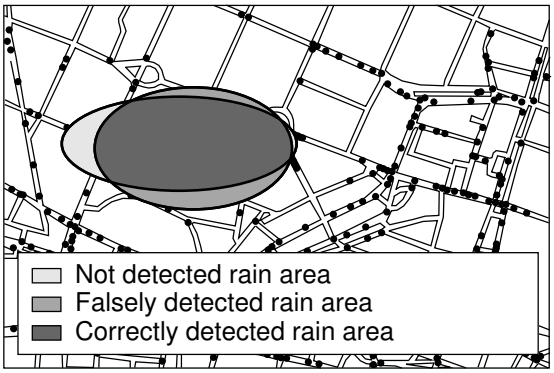


Figure (3.6) Example for a rain area and detection area overlap [EMS06]

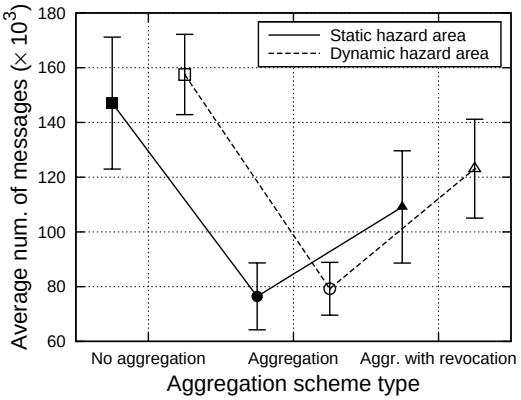
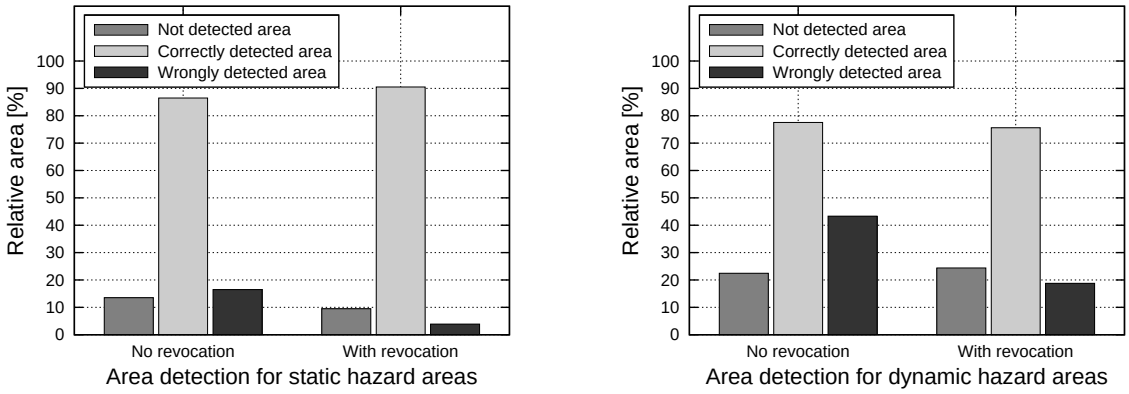


Figure (3.7) Number of messages required for static and dynamic hazard areas [EMS06]

The detection quality for hazard areas using message aggregation is shown in Fig. 3.8. It is measured by comparing the size and location of the true hazard area with the detected area.

Fig. 3.8(a) shows the result for a static hazard area. Using the aggregation mechanism without revocation results in a positively detected area of 86.5% at the end of the aggregate's lifetime. Evaluating the same value after only 100 s, which corresponds to half the aggregate's lifetime, the positively detected area amounts to 71.1%. Due to the longer detection period more individual events can be detected, resuling in a more accurate event detection. The falsely detected hazard area is about 16.5% of the actual hazard area. Using revocation messages this result can be improved significantly, since only 3.9% of the area are false positive. In this case the correctly detected hazard area increases to 90.5%.

The detection of dynamic hazard areas is more difficult. Hence, the results shown in Fig. 3.8(b) are inferior compared to the static hazards. The aggregation without revocation manages to detect 77.5% of the real hazard area. About 22.5% of the hazard area are not detected and the large amount of 43.3% is falsely detected. However, by using the revocation mechanism the falsely detected area can be reduced more than 50% down to a value of 18.8%. The weighing of messages for dynamic event aggregates needs to be carefully adjusted to



(a) Static hazard areas

(b) Dynamic hazard areas

Figure (3.8) Event detection quality for hazard areas [EMS06]

the characteristics of the hazard changes to reduce the detection error, while still reducing the number of messages.

Overall the results prove that aggregation is a valid strategy to reduce the number of messages needed to disseminate hazard events in a VN. Therefore, it is a mechanism which improves dissemination scalability and efficiency. Moreover, the message aggregation mechanism can be used to disseminate information on hazard areas. A new approach using ellipses for area modeling has been suggested and evaluated. This works especially well for static events, however, the reduction effect still holds true for dynamic events. Most of these results have previously been published in [EMS06].

3.4 Evaluation of Message Distribution using IEEE 802.11p/WAVE

After the introduction of conditional flooding and aggregation mechanisms in Sec. 3.3 the use of prioritization will be evaluated and used in the next two sections. The new communication standard for vehicular communication, the IEEE 802.11p or WAVE standard, which has been defined over the last years by the IEEE standards committee, also uses prioritization [Tas06]. The IEEE 802.11p standard is based on the well-known IEEE 802.11 WLAN standard [LAN99]. However, in contrast to WLAN it uses a multi-channel concept, which is adapted to the vehicular scenario. In addition, it accounts for the priority of messages using Access Classes (ACs), which use different MAC parameters. This is practically identical to the EDCA amendment of WLAN [Sta05]. The use of the EDCA mechanism shall enable the distribution of highly time critical safety messages in dense scenarios, parallel to the distribution of best-effort type messages.

Since the WAVE standard is specially designed for VANET communication it should fulfill most if not all of the requirements for VANETs. The following analytical and simulation-based evaluations of the standard help to assess its capabilities and limitations. The presented results have previously been published in [Eic07b].

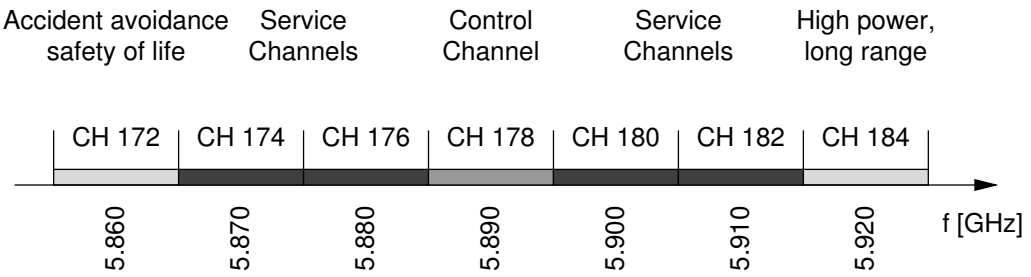


Figure (3.9) Physical channels for IEEE 802.11p/WAVE

This section on the IEEE 802.11p standard evaluation is organized as follows. In Sec. 3.4.1 the main definitions of the standard are introduced. An analytical evaluation of collision probability and throughput is presented in Sec. 3.4.2. In Sec. 3.4.3 the results for several simulation scenarios are discussed. Finally, in Sec. 3.4.4 the WAVE communication standard is assessed based on the simulation results.

3.4.1 Inter-Vehicle Communication Using WAVE

Before analyzing the IEEE 802.11p and WAVE standard family the most important features are introduced. These can be found in [Com06, Tas06]. The physical layer of IEEE 802.11p uses seven channels of 10 MHz bandwidth each. The bandwidth for these channels is allocated in the upper 5 GHz range as depicted in Fig. 3.9. Since the design shall allow both single- and multi-transceiver units, the different channels can not be used simultaneously, however, each ME alternates between the Control Channel (CCH) and one of the Service Channels (SCHs) or the safety channel (CH 172). In order to meet the strong delay requirements, for example, of collision warning messages, a channel period (t_p) containing one CCH and one SCH interval must not last more than $t_\mathrm{p} = 100$ ms. For the evaluation of the standard the CCH and SCH equally share t_p, thus, every interval time (t_i) of $t_\mathrm{i} = 50$ ms the channels are changed.

The MAC layer in WAVE is equivalent to the IEEE 802.11e EDCA QoS extension [Sta05]. This allows the categorization of messages into different ACs, having different channel access priorities. Four different ACs are defined, where AC0 has the lowest and AC3 the highest priority. Within the MAC layer a packet queue exists for each AC. During the packet selection process for the next transmission, all four ACs contend internally. The selected packet then contends for the channel externally using its selected contention parameters. This system concept is shown in Fig. 3.10. The contention parameters used for the CCH are shown in Tab. 3.1. To calculate `CWmin` and `CWmax` the values `aCWmin` $= 15$ and `aCWmax` $= 1023$ have to be used.

Depending on the AC, first the Arbitration Inter-Frame Space (AIFS) is selected. The AIFS period (t_a) is defined as $t_\mathrm{a} = \texttt{AIFS} \cdot t_\mathrm{s}$, where the slot time ($t_\mathrm{s}$) is set to $t_\mathrm{s} = 16$ µs. Next, the size of the Contention Window (CW) is determined, randomly selecting a value between 0 and `CWmin` for the first transmission attempt, hence, the contention period (t_c)

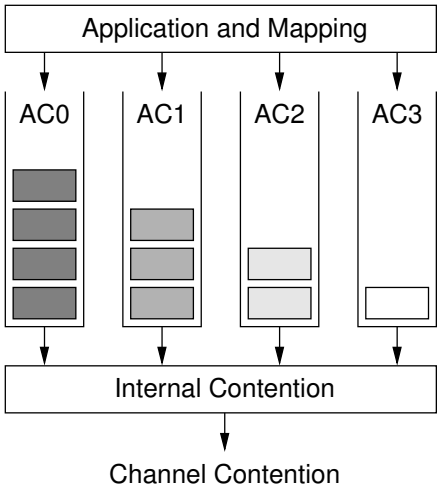


Figure (3.10) MAC queues of WAVE with four access classes [Com06]

AC	CWmin	CWmax	AIFS	t_w
0	aCWmin	aCWmax	9	264 µs
1	$\frac{\text{aCWmin}+1}{2} - 1$	aCWmin	6	152 µs
2	$\frac{\text{aCWmin}+1}{4} - 1$	$\frac{\text{aCWmin}+1}{2} - 1$	3	72 µs
3	$\frac{\text{aCWmin}+1}{4} - 1$	$\frac{\text{aCWmin}+1}{2} - 1$	2	56 µs

Table (3.1) EDCA parameters for the CCH used in the WAVE system

equals to $t_c = \texttt{CW} \cdot t_s$. In case of a collision, the transmission will be retried, using an increased CW of $2 \cdot (\texttt{CW} + 1) - 1$. This increasing will be continued until both CWmax and the maximum number of retries (7) is reached.

The described contention mechanism is similar to the ones known from conventional WLAN and the EDCA extension, however, WAVE uses specific parameters for its EDCA extension. The contention process leads to the wait times shown in Fig. 3.11. Each AC has to wait at least its AIFS slots, plus additional slots determined by the selected CW value, leading to the average wait time (t_w) in Tab. 3.1.

3.4.2 Analytical Evaluation of Throughput and Collision Probabilities

Knowing the characteristics and theoretical capabilities of the WAVE technology, the question arises on how effective this concept is in reality. Using an example setup of WAVE the feasible throughput and the probability for a collision-free channel access has been elaborated. In this scenario a packet size of 500 B was assumed. This is a reasonable average packet size, including both data and security information. Taking into account the constraints of the contention process (see Sec. 3.4.1), the throughput for the specified packet size during one interval time (t_i) of $t_i = 50$ ms can be calculated for each AC. This is done with Eqn. (3.7). The calculated upper bound of the throughput for a CCH using a data rate of 6 Mbit/s and

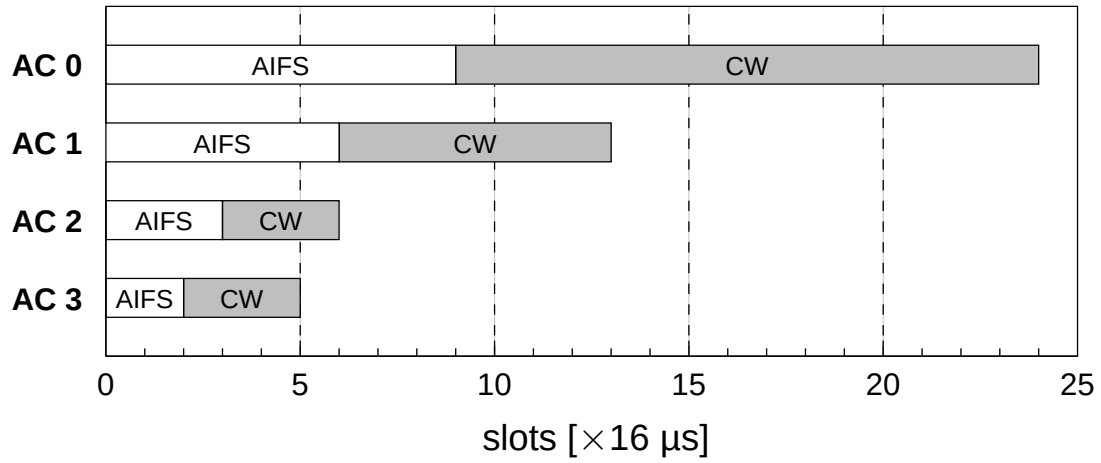


Figure (3.11) Wait-times for different access categories in the CCH caused by contention

the specified packet size is shown in Fig. 3.12(a). This upper bound can be reached as long as *only one* sender is using the channel. Collisions will reduce the throughput as soon as more senders access the channel simultaneously.

$$N_\mathrm{p} = \frac{t_\mathrm{i}}{t_\mathrm{a} + t_\mathrm{c} + t_\mathrm{d}} \tag{3.7}$$

$$t_\mathrm{d} = t_\mathrm{head} + t_\mathrm{data} = 96\ \mu\mathrm{s} + 667\ \mu\mathrm{s} \tag{3.8}$$

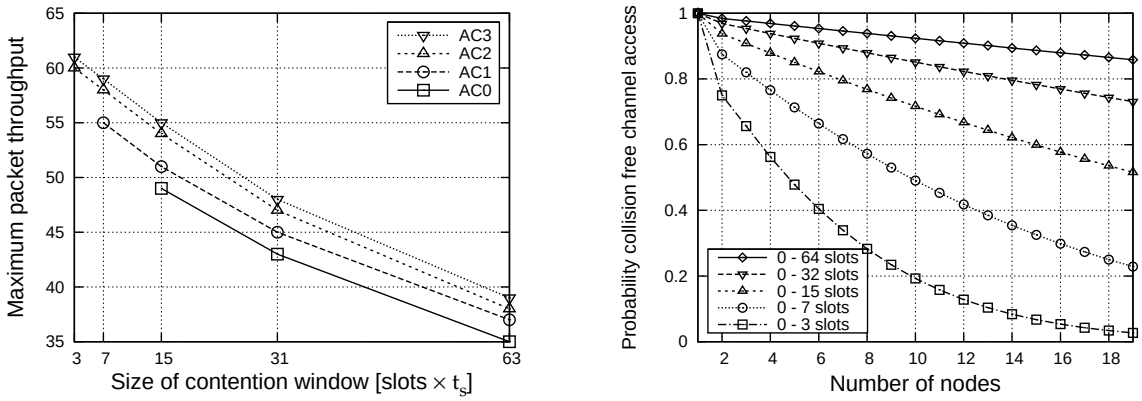
$$n_\mathrm{t}\left(N_\mathrm{n}\right) = {n_\mathrm{cs}}^{N_\mathrm{n}} \tag{3.9}$$

$$p_\mathrm{coll}\left(N_\mathrm{n}, n_\mathrm{cs}\right) = \frac{N_\mathrm{n}}{n_\mathrm{t}} \cdot \sum_{i=1}^{n_\mathrm{cs}-1} \left(n_\mathrm{cs} - i\right)^{N_\mathrm{n}-1} \tag{3.10}$$

The probability of a channel access collision depends on the available number of contention window slots (n_cs). The number of slots is specified for each AC used in the WAVE standard. In a scenario with N_n sending nodes a collision will occur if at least two nodes select the same t_c and no other node selects a shorter t_c. The number of combinations of N_n nodes each selecting a t_c is given by Eqn. (3.9). The collision probability (p_coll) can be calculated with Eqn. (3.10). It depends on the parameters N_n and n_cs.

Using Eqn. 3.10 the probability for a collision free channel access can be calculated. In Fig. 3.12(b) this has been done for different CW sizes and an increasing value of N_n. The plots in Fig. 3.12(b) clearly show the limitations for AC3, using a CW size of $n_\mathrm{cs} \in 0 \ldots 3$. As soon as several nodes contend using AC3, a collision becomes very likely, reducing the successful data throughput. For example, six nodes sending one AC3 message each, have a probability for a collision free channel access of just 40%.

To reduce the collision probability the CW size has to be increased, however, this leads to a slightly reduced throughput maximum as shown in Fig. 3.12(a). The second option to reduce the collision probability is to shape the traffic and reduce the number of high priority packets using AC3. However, the traffic shaping mechanism would have to consider the



(a) Maximum packet throughput for different access categories without collisions

(b) Collision free channel access probability for different CW sizes depending on simultaneously acting nodes

Figure (3.12) Influence of CW size on channel access and throughput

current number of neighbors to be effective, which is not necessarily an available information in a decentralized network.

The wait-times given in Fig. 3.11 clarify that message in AC3 and AC2 win the contention process against the lower priority categories (AC1 and AC0). The reason for that are the much lower contention periods and especially due to the long AIFS times for AC0 and AC1. However, AC3 and AC2 contend against each other, as well as AC1 and AC0.

3.4.3 Evaluation of IEEE 802.11p/WAVE by Simulation

To be able to further elaborate on the characteristics and constraints of the WAVE standard the system concept has been implemented in the simulation environment OMNeT++, to conduct a variety of simulations.

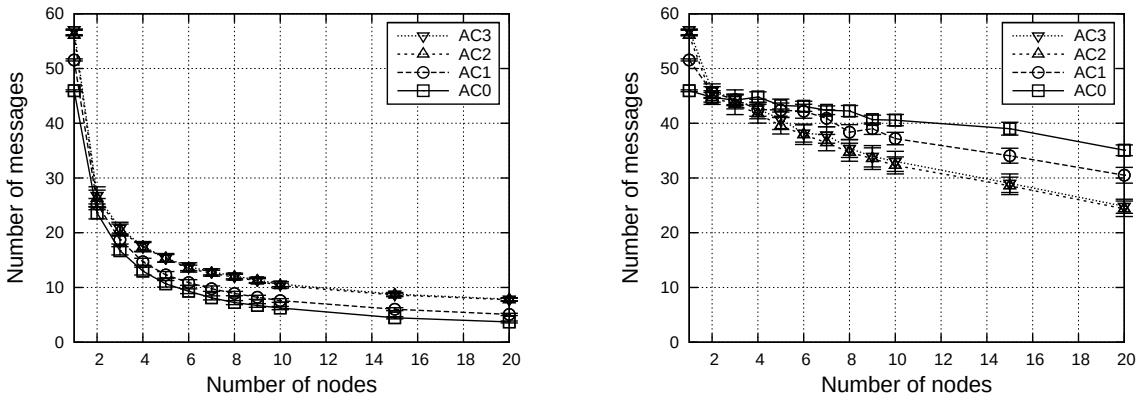
Simulation Environment and Parameters

For the simulations the OMNeT++ simulation environment has been used [Var01, VH07]. A detailed description of the base models used for simulation in this thesis is given in App. A. For the simulations presented in the following, the existing WLAN model has been replaced with a respective WAVE model. Compared to the analytical evaluation in Sec. 3.4.2 the same parameters were applied for the simulations. Hence, the simulated CCH had a data rate of 6 Mbit/s. The radio-range was set to 250 m and the interference of other transmissions were regarded up to a distance of about 4 km. The interval time (t_i) of the CCH was set to $t_i = 50$ ms. The WAVE simulation model detects packet collisions if the radio tries to transmit just at the same time as a new packet is being received, leading to a packet retransmission.

Besides a model validation scenario, two different VANET scenarios have been analyzed. Both scenarios were based on the Manhattan Grid Mobility (MGM) model with the parameters specified in Sec. A.3.2 [pp. 168]. The average node speed of 60 km/h was applied. Each

	Low Data Traffic		High Data Traffic	
	inter-arrival	packets/s	inter-arrival	packets/s
AC3	100 ms	10.0	80 ms	12.5
AC2	100 ms	10.0	60 ms	16.7
AC1	90 ms	11.1	30 ms	33.3
AC0	45 ms	22.2	20 ms	50.0

Table (3.2) Traffic load parameters used for the WAVE evaluation



(a) Number of sending attempts per node and AC

(b) Number of successfully received messages

Figure (3.13) Sending and receiving of messages using WAVE

simulation run used a simulation time of 900 s. The main difference of the two scenarios was the amount of data traffic, a low data traffic and a high data traffic scenario were simulated. The mean parameters for the exponentially distributed inter-arrival times of the generated messages are given in Tab. 3.2. Besides the different traffic values several different node densities were applied. A value of $N_\mathrm{n} = 100$ is equivalent to an average of 1.9 neighbors, $N_\mathrm{n} = 200$ is equivalent to 3.8 neighbors and so forth.

Model Validation and Performance Evaluation

In a first step the WAVE simulation model was validated using plausibility checks and the analytical results presented in Sec. 3.4.2. On a simulation area of 150 m $\times$ 150 m one receiving node and between one and twenty sending nodes were randomly distributed. The sending nodes were initialized with 60 messages of the same AC. Simulation duration was one CCH interval time (t_i).

In Fig. 3.13(a) the number of sending attempts per node and AC are plotted. The values for one sending node conform to the values given in the analytical evaluation (Fig. 3.12(a)). Since all nodes share the channel, the attempts exponentially decrease with increasing number of nodes. Moreover, for low priority ACs fewer sending attempts occur, mainly due

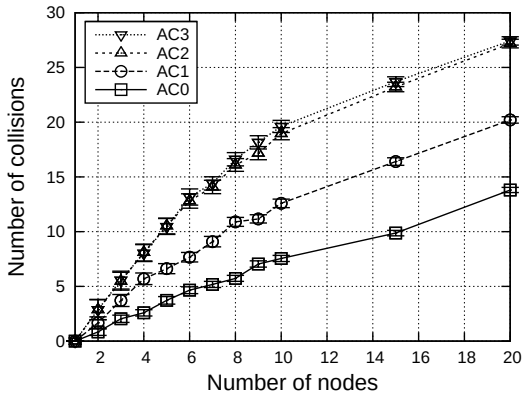


Figure (3.14) Detected packet collisions at the receiver

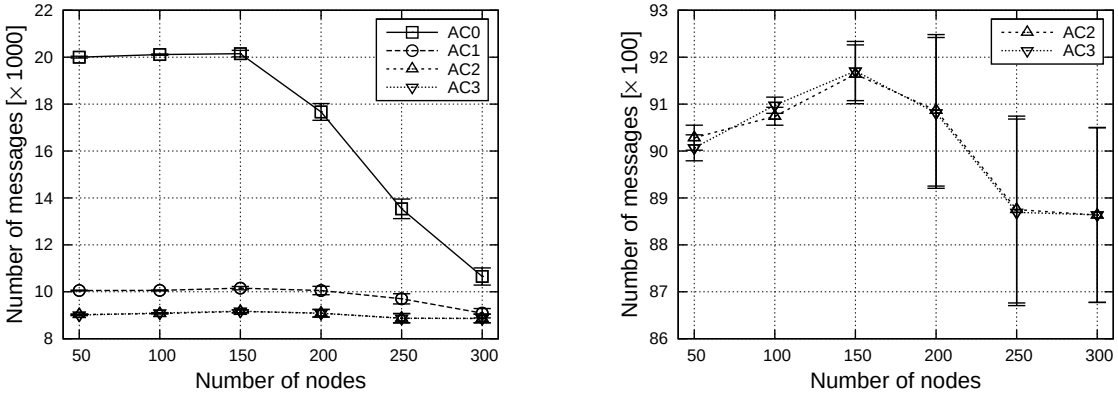
to the longer contention times. More important than the sending attempts are the number of received packets. The evaluation results at the receiving node are plotted in Fig. 3.13(b).

The plots show that for all categories the number of received messages is linearly decreasing. This can be explained by the increasing number of collisions on the channel (see Fig. 3.14). A somewhat surprising result is that the lowest priority category AC0 is not as much affected by the collisions as the high priority categories. This effect can be explained with the significantly smaller CW used for AC3. Due to the small CW more collisions occur for AC3 compared to AC0 (see Fig. 3.14). This result demonstrates that the WAVE standard can not cope with many high priority messages in a dense scenario. Hence, the traffic shaping algorithms, mapping messages to ACs, should incorporate the number of current neighbor nodes in the mapping decision. Alternatively the number of received messages per AC over the last channel interval could be regarded to improve the mapping. That way the high collision probability can be reduced.

Simulation Results for Vehicular Network Scenarios

The packet generation parameters specified in Tab. 3.2 are valid for the application layer, however, they do not necessarily correspond to the number of actually broadcast packets on the channel. The channel capacity and the interactions between the ACs limit the throughput. This is represented in the *send* and *received* results for both the low and the high traffic scenario given in Fig. 3.15, Fig. 3.16, and Fig. 3.17.

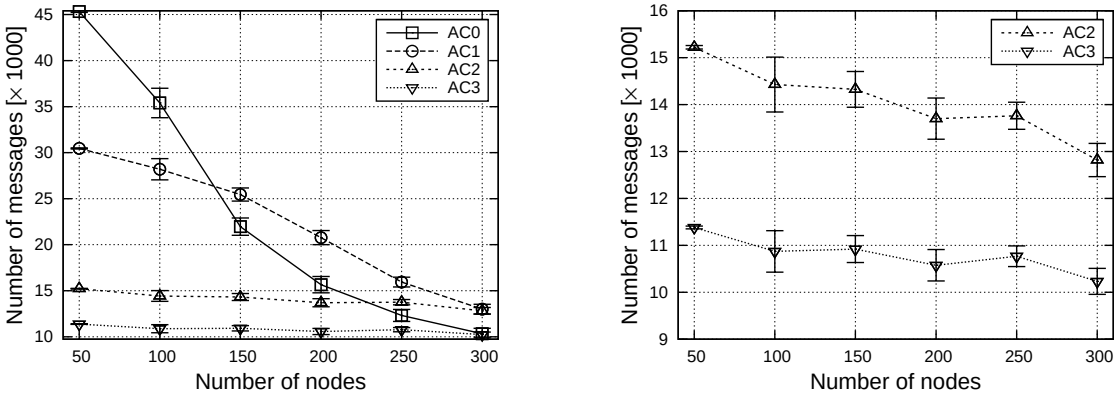
For the low traffic scenario the average number of sent broadcasts per node and its distribution to the four ACs is plotted in Fig. 3.15(a), and for AC2/AC3 in greater detail in Fig. 3.15(b). With an increasing number of nodes the general network load increases, leading to a stronger indirect interaction between the ACs. As soon as the maximum throughput capacity is reached and the traffic load further increases, the packets of the lowest category AC0 get fewer channel accesses. This leads to the decline in sent broadcasts for AC0. As the traffic increases further, AC1 is the next category to lose access shares on the channel (see Fig. 3.15(a) starting at 200 nodes).



(a) Average number of sent messages

(b) Detailed sent statistic for AC2 and AC3

Figure (3.15) Sending statistic for the low traffic scenario



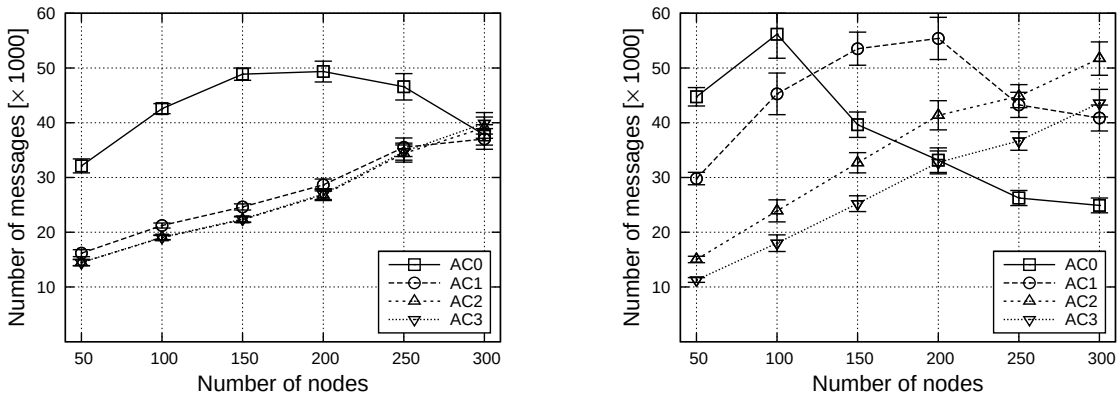
(a) Average number of sent messages

(b) Detailed sent statistic for AC2 and AC3

Figure (3.16) Sending statistic for the high traffic scenario

Looking at the sending behavior of the high traffic scenario the same characteristics can be found. However, due to the much higher load the decline effects of the ACs AC0 and AC1 are much more severe. This is shown in Fig. 3.16(a). The number of sent packets of AC0 practically declines exponentially, while the number of sent messages of AC1 still declines in a linear fashion. The same decline effect happens to AC2 and AC3 in the high traffic scenario, however, the decline is marginal compared to the other ACs (see Fig. 3.16(b)). While the number of messages in AC2 decreases by 16% compared to the average of 15000 sent messages, the number of message in AC3 still decreases by 9% compared to the average of 11250 sent messages.

At the receiver side the decrease in sending attempts for specific ACs causes a decrease in the number of received packets of the equivalent ACs. This effect is observed for both the low and the high traffic scenario (see Fig. 3.17). The first AC to be affected is AC0, which can be clearly seen in Fig. 3.17(a). The decline in the reception of the low priority ACs can



(a) Low traffic scenario

(b) High traffic scenario

Figure (3.17) Average number of received packets

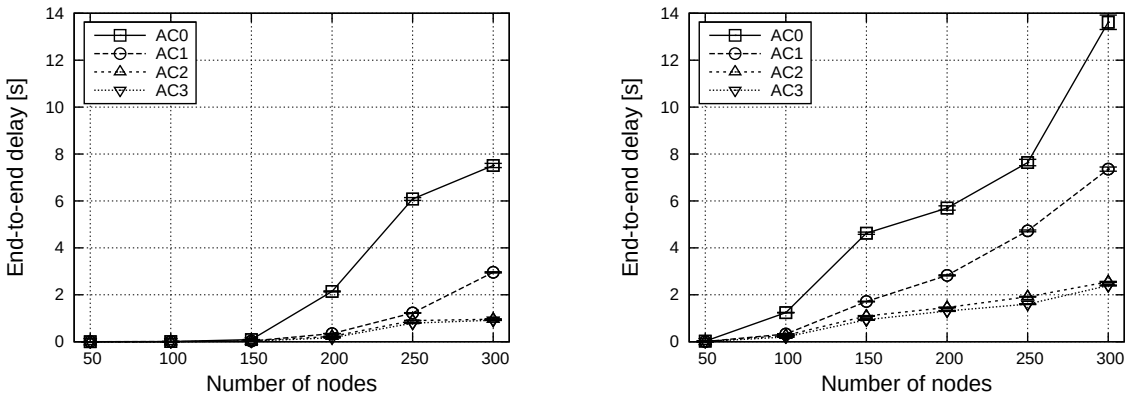
be observed even more severe in the high traffic scenario. In Fig. 3.17(b) the decline for AC0 and AC1 can be seen very clearly. The increase of N_n leads to an increasing traffic load, therefore, even the higher priority ACs start to be influenced by the limited channel capacity.

Since the different ACs were primarily introduced to provide a prioritization mechanism for time critical messages, an important parameter to evaluate is the end-to-end (E2E) delay. The average E2E delay is plotted in Fig. 3.18 for both scenario settings. The results show that especially the low priority data packets suffer from an exponential increase of the E2E delay with increasing node density. However, the E2E delay for AC3 also increases. The average value of 1 s for 300 nodes in Fig. 3.18(a) is relatively high, considering that collision warning messages should have a maximum delay of 100 ms to be able to provide a reliable safety service [MLS05]. The E2E delay for the high data traffic scenario is about twice as long as in the low data traffic scenario, which can be seen in Fig. 3.18(b). Therefore, especially in high load scenarios the distribution of high priority messages is challenging. WAVE can prioritize groups of messages over others, yet a time critical distribution with guarantees is not feasible in this way.

3.4.4 Assessment of the WAVE Communication Technology

The defined parameter set for the EDCA used in WAVE is capable of prioritizing messages, however, with an increasing number of nodes, sending AC3 especially, the collision probability increases significantly. Since collisions are detected *after* a transmission if at all, a high collision probability results in many dead times; times where the channel is blocked but no useful data is exchanged. Due to the continuous switching between CCH and SCH, which also use different packet queues, the collisions have an even worse impact. Messages for the CCH queue up further during the SCH intervals, resulting in longer queues and a higher end-to-end delay.

Especially in dense scenarios or in case of filled MAC-queues, the WAVE technology can not ensure time critical message dissemination (for example collision warning messages). One improvement possibility is to use different EDCA parameters, which would mitigate



(a) Low traffic scenario

(b) High traffic scenario

Figure (3.18) Average end-to-end packet delay for different access classes

the high collision probability. In addition, a continuous priority re-evaluation mechanism, similar to the concept presented in Sec. 3.5, could be integrated into the WAVE technology to improve its performance. This would continuously reduce the number of high priority messages due to aging, hence, prevent long message queues. Moreover, the assigned spectrum should be used more efficiently. Multi-transceiver communication devices could ensure a continuous usage of the safety channel, therefore, doubling its capacity compared to WAVE. Alternatively a Time Division Multiple Access (TDMA)-based communication could be used at intersections to provide exclusive channel access to distinct stations, thus, preventing packet collisions entirely. However, this setup would require a centralized scheduler at the intersection. For all possible solutions an assessment for performance gain, complexity, and costs has to be made to find the optimal solution.

3.5 Message Distribution with Prioritization: Using Content-Utility as Priority Index

The IEEE 802.11p/WAVE standard uses a simple prioritization capability based on four ACs to tackle the scalability issue in VANETs. This concept can be successful in specific network settings, however, it can not deal with a highly overloaded system adequately. Moreover, the standard does not suggest how to assign messages to the different ACs.

One possibility to realize this assignment is to use context information, for example gained from the messages' content. A message distribution system using prioritization based on the message content's utility to other nodes is introduced and evaluated in the following section. This concept promises to be a valid approach to increase scalability and quality of message dissemination in VNs.

3.5.1 The General Concept of Prioritization-based Message Distribution

Many different types of messages are simultaneously distributed in a VANET. This is mainly due to the fact that multiple services run in parallel. If all messages are distributed in the network without any limitations whatsoever, this will lead to a Broadcast Storm or at least to a major congestion. Therefore, mechanisms to limit the distribution area have to be used.

While the distribution areas presented in Sec. 3.3.1 only use the position information of messages to limit the message distribution, other criteria, for example, message age or time of day, can be used additionally. Furthermore, using several criteria simultaneously provides the possibility to calculate a priority value for each message. Based on this priority value the distribution decision is made. For example, only messages with a priority value above a given threshold will be sent. In addition, all packets in the sending queue can be ordered using the messages' priority.

Before the prioritization-based message distribution concept is introduced, the general features of the system are presented. In addition, its main advantages are discussed. The message prioritization concept using content-utility as priority index has previously been published in [ESKS06] and [KAE+06].

Altruism of the concept: Nodes can only control which messages they send to neighboring nodes instead of which messages they receive. Hence, each node tries to be as altruistic as possible by sending the message with the highest utility of all available messages. This will most likely be the message providing the highest degree of entropy to neighboring nodes. To be able to do so, the utility of all available messages is calculated frequently. Different context categories are evaluated to determine a message's utility value.

Sophisticated information differentiation: The concept uses context categories to differentiate messages. Specific context categories can be selected for each message type, providing the best content assessment possible. This ensures that different services can run on the same platform and share one wireless channel. The differentiation mechanism is capable of selecting the most important messages for distribution, given the fact that the utility value can be computed objectively.

Decentralized scheduling: The prioritization concept works fully decentralized. Hence, no coordinated or managed scheduling is required. Each node acts individually and uses a contention mechanism when accessing the shared channel.

Adaptability and flexibility: The system concept is adaptable to new service classes. Due to its flexible differentiation strategy based on context categories, it is easily adaptable to changing requirements. The concept works in both sparse and dense network scenarios, thus, it is applicable for initial VANET deployments as well as for future, high-density scenarios. Further, the differentiation mechanism can be integrated with many other VANET message dissemination concepts or works as a comprehensive substitute.

3.5.2 Using Information Benefit and Utility Maximization for Message Prioritization

The utility of up-to-date traffic safety information can be quantified with many different measures. Some examples are reduced number of accidents and injuries, reduced traffic jams and shorter average travel times, and reduced environmental pollution. However, these criteria are not suitable for the utility calculation in the vehicle when actually processing new messages. Therefore, different criteria need to be specified to calculate a message's utility.

Since different types of messages are distributed simultaneously over the same platform, the utility quantification needs to be very flexible. Hence, message context categories are used to quantify the utility.

Context Categories for the Benefit Calculation

The use of context categories provides a flexible and adaptable quantification strategy. It can easily be changed or complemented if new message categories are added to the system or existing message types need to be quantified differently. Three main context categories have been identified as most relevant to the utility quantification process.

Information context: The most important context is the information context (c_i). All message types have a specific information content. Depending on the message category, its content has a specific c_i which requires a customized benefit quantification. Message contents like *distance to event*, *accuracy*, *degree of timeliness*, or the message context like *time of day* or *current driving route* have to be handled individually for each message category. Everything that directly corresponds to the information of a specific message category belongs to the respective c_i.

Message context: The second context category is the message context (c_m). In contrast to the information context (c_i) the message context (c_m) is not different for various message categories, however, c_m can still be weighted differently for the utility quantification. Examples for c_m are *message age*, *time since last reception*, and *time since last broadcast*.

Vehicle context: The third context category, the vehicle context (c_v), includes all vehicle specific context information connected to a message. Similar to c_m the vehicle context is used analogically for different message categories, only the weighing factors differ. The vehicle context is described by values like *driving direction*, *number of neighbors*, *vehicle speed*, and *route flexibility*.

Other context categories are possible besides these three main categories, however, most context information can be allocated to one of these three categories. Thus, a flexible utility quantification can be set up using c_i, c_m, and c_v as context categories.

Benefit Calculation for Context Category Information

Having an allocation of information to different context categories enables its utility quantification. The calculation process of a message's utility value consists of several steps. The

first step is the calculation of benefit values for the individual information pieces or context categories. A benefit function ($\mathcal{B}$) is used to determine these individual benefit values. The benefit functions used to determine the benefit values have specific properties.

Property 3.1 *The benefit functions $\mathcal{B}$ are strictly monotonic.*

$$\left|\frac{\partial \mathcal{B}_j(c_j,t)}{\partial c_j}\right| > 0 \quad \forall c_j \geq 0 \tag{3.11}$$

Property 3.2 *All values of $\mathcal{B}$ are non-negative and belong to the closed interval $[0,1]$.*

$$\mathcal{B}_j(x) \in [0,1] \quad \forall\, x \in [0,\infty) \tag{3.12}$$

A benefit function ($\mathcal{B}$) usually is concave or sigmoid and a continuously differentiable function for values greater or equal to zero [FC05]. A typical representative of a basic benefit function is given by Eqn. (3.13), specifying a concave and sigmoid benefit function.

$$\mathcal{B}(x) = 1 \pm \frac{x}{\sqrt{1+x^2}} \tag{3.13}$$

However, this base function can not be applied directly. It needs to be adapted to fit the required properties of the context and message category. Four parameters are used to adapt the base function given by Eqn. (3.13), to move the curve and define the adequate decline rate. This leads to the two benefit functions given by Eqn. (3.14) and Eqn. (3.15).

$$\mathcal{B}(x) = \frac{1 + a - \frac{\left(\frac{x-b}{c}\right)}{\sqrt{1+\left(\frac{x-b}{c}\right)^2}}}{d} \qquad \forall\, x \in [0,\infty) \tag{3.14}$$

$$\mathcal{B}(x) = \frac{1 + a + \frac{\left(\frac{x-b}{c}\right)}{\sqrt{1+\left(\frac{x-b}{c}\right)^2}}}{d} \qquad \forall\, x \in [0,\infty) \tag{3.15}$$

The parameters a, b, c, and d are used to adapt the characteristics of $\mathcal{B}$. Each parameter has its specific task:

- $a|a \geq 0$ is used to move the curve up and down on the y-axis.
- $b|b \geq 0$ is used to move the curve left and right on the x-axis.
- $c|c \geq 0$ is used to adapt the rate of decline.
- $d|d \in [1,2]$ is used to keep the benefit in the interval $[0,1]$.

Based on the equations Eqn. (3.14) and Eqn. (3.15) as well as the four adaptation parameters, specific benefit functions with arbitrary characteristics can be defined. Examples for resulting benefit functions are introduced and shown in Sec. 3.5.3.

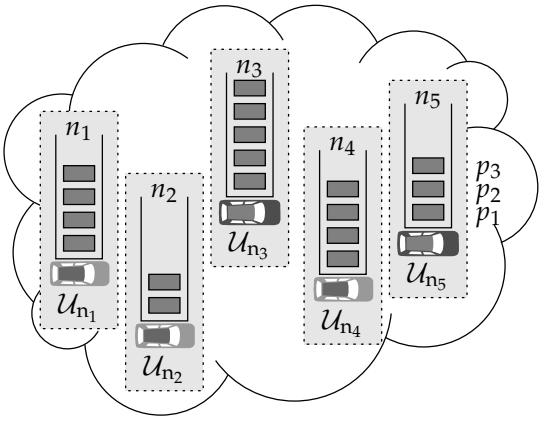


Figure (3.19) Snapshot of an idealized scenario

Combining Multiple Benefit Values to the Message Utility Value

After calculating the individual message context benefit values, an overall message utility value needs to be calculated. Hence, a utility function ($\mathcal{U}$) is required. The calculation process with its individual steps is shown in Eqn. (3.16). Based on the packet (p_i) the context categories are evaluated and the benefit values calculated. These benefit values $\mathcal{B}_j$ are assessed with individual weighing factors (w_j) to determine the overall message utility $\mathcal{U}$.

$$p_i \rightarrow \left\{ \begin{array}{lll} c_\mathrm{i} & \Rightarrow \mathcal{B}_j(c_\mathrm{i}) & \cdot w_j \\ c_\mathrm{v} & \Rightarrow \mathcal{B}_j(c_\mathrm{v}) & \cdot w_j \\ c_\mathrm{m} & \Rightarrow \mathcal{B}_j(c_\mathrm{m}) & \cdot w_j \end{array} \right\} \Rightarrow \mathcal{U}_i(p_i, c, t) \tag{3.16}$$

$$\mathcal{U}_i(p_i, c, t) = \frac{1}{\sum_{j=1}^{N_\mathrm{c}} w_j} \cdot \sum_{j=1}^{N_\mathrm{c}} w_j \cdot \mathcal{B}_j(c_\mathrm{i}, c_\mathrm{v}, c_\mathrm{m}) \tag{3.17}$$

The general utility function ($\mathcal{U}$) given in Eqn. (3.17) uses the packet and the related context information to determine the message's utility value. The benefit values of all context categories are assessed and summed up, using the number of context categories (N_c). The resulting value is divided by the sum of all context category weights, in order to obtain a utility value $\mathcal{U} \in [0, 1]$.

To be able to compare and analyze simulation results for scenarios using message utility as priority-index, the global utility ($\mathcal{U}_\mathrm{g}$) is an important parameter. With Eqn. (3.18) the value for $\mathcal{U}_\mathrm{g}$ can be determined. This is done by summing the utility values for all messages N_m, received by all nodes N_n, integrated over the whole simulation time (t_sim).

$$\mathcal{U}_\mathrm{g} = \int_{t=0}^{t_\mathrm{sim}} \sum_{n=1}^{N_\mathrm{n}} \sum_{i=1}^{N_\mathrm{m}} \mathcal{U}_i(p_i, c, t) dt \tag{3.18}$$

To better understand the characteristics of the utility-based message prioritization concept, an idealized and simplified scenario is discussed as an example. This scenario is

made up of N_n MEs within mutual communication range. For this example let $N_\mathrm{n} = 5$, as shown in Fig. 3.19. At a randomly given point in time t, the scenario is evaluated. At time t, all MEs have at least one packet to transmit. At the same time all MEs have an individual interest in certain pieces of information. The information is transported in form of wireless packets p, which each have an individual utility value quantified with a utility function ($\mathcal{U}$) specified in Eqn. (3.17). The overall goal is to globally maximize the utility provided to all nodes.

Imagine a scheduler with global knowledge of all packets existing in the network and their utility [ES07]. This scheduler could reach the optimized global network utility by performing the algorithm specified in Alg. 3.2. This process helps to select the packet p providing the highest utility to the network at time t.

```
1  while packets p exist for broadcast do
2  |  forall nodes n_i, i = 1 ... N_n do
3  |  |  forall packets p_j, j = 1 ... N_p do
4  |  |  |  Calculate U_j(p_j, c, t)
5  |  |  end
6  |  |  Define packet queue order using U_j(p_j, c, t);
7  |  |  Select packet p_j| max (U_j(p_j, c, t));
8  |  end
9  |  Select packet p_{i,j}| max (U_{i,j}(p_{i,j}, c, t))    ∀ i ∈ [1, N_n], ∀ j ∈ [1, N_p];
10 |  Send packet p_{i,j};
11 end
```

Algorithm (3.2) Message handling of a global scheduler to optimize the network utility

At first the packets at all nodes need to be ordered according to their utility value. Each node should broadcast the packet providing the maximum utility to the receiving nodes. This overall network utility ($\mathcal{U}_\mathrm{n}$) specified in Eqn. (3.19) is calculated as the sum of each nodes' gained utility value if packet $p_{i,j}$ is sent, minus the utility for the sending node. The normalized network utility ($\overline{\mathcal{U}_\mathrm{n}}$) given in Eqn. (3.20) normalizes $\mathcal{U}_\mathrm{n}$ to the number of receiving nodes.

$$\mathcal{U}_\mathrm{n}(i,j,t) = \left(\sum_{m=1}^{N_\mathrm{n}} \mathcal{U}_m(p_{i,j},t)\right) - \mathcal{U}_{i,j}(p_{i,j},t) \tag{3.19}$$

$$\overline{\mathcal{U}_\mathrm{n}}(i,j,t) = \frac{1}{N_\mathrm{n}-1}\left[\left(\sum_{m=1}^{N_\mathrm{n}} \mathcal{U}_m(p_{i,j},t)\right) - \mathcal{U}_{i,j}(p_{i,j},t)\right] \tag{3.20}$$

$$\max\left\{\overline{\mathcal{U}_\mathrm{n}}(i,j,t)\right\} = \max_{i=1\ldots N_\mathrm{n}}\left\{\max_{j=1\ldots N_\mathrm{p}}\left[\mathcal{U}_\mathrm{n}(i,j,t)\right]\right\} \tag{3.21}$$

However, the internal packet order is not sufficient to reach the overall maximum of the aggregated utility values. Therefore, a second step needs to be taken by the global scheduler. The five packets, one from each node in the scenario, that contend for the channel access, will most likely have different utility values. Hence, after the internal ranking also a global ranking of packets needs to be established in the second step. This ensures that the node

n_i, owning the packet $p_{i,j}$ with the overall highest utility value, will send first. Hence, the second step in the optimization results in the maximum normalized global utility specified in Eqn. (3.21).

This idealized scenario using a global scheduler, previously published in [ES07], is not feasible in a decentralized VANET environment, since no global control exists. Therefore, the practical realization of this scheduling concept needs to be adapted to the requirements of the scenario. Each node will have to individually calculate and estimate the utility each packet will potentially provide to receiving nodes. The limited knowledge of the neighbors' interests prevents an optimal packet ordering. Nevertheless, this adaptation can almost match the optimal utility calculation.

The adaptation required for the second step is more severe. No global contention mechanism is used to handle channel access according to the actual utility values. Hence, a decentralized contention mechanism needs to be adapted to best fit the required needs. Moreover, nodes will not all be in mutual communication range in a real-life scenario. Therefore, interference will further reduce the throughput and decrease the achievable global aggregate utility value.

The two main adaptations required for the real-life VANET scenario have been taken into account in the system concept specified in Sec. 3.5.3. Existing wireless communication systems like WLAN or even the IEEE 802.11p WAVE vehicular communication system can be modified to fit the needs for the suggested utility-based information distribution.

3.5.3 System Concept for Utility-based Information Distribution

The utility calculation introduced above needs to be integrated into a system architecture to use it for information distribution. To optimally use the utility values in a communication architecture an OSI layered system approach with cross-layer elements is proposed. The suggested system design is depicted in Fig. 3.20, which has been previously suggested in [ESKS06] in similar fashion.

The application layer manages all message related tasks: Generating, broadcasting, receiving, and storing messages. It is embedded into the vehicle's system architecture, therefore, it can access necessary resources to generate messages or utilize the information of incoming messages. In addition, it can generate the context information, required for the utility value calculation. The utility values are needed both in the application and the link layer, hence, a Cross Layer Communication (CLC) is required to provide the functionality in both layers. The utility value is used for three tasks:

1. Remove messages with a utility value below a given threshold,
2. provide a utility-based queue management,
3. implement a utility-based medium access strategy.

The utility-based queue management changes the de- and enqueue behavior of the link layer. Therefore, the frequently used drop-tail queue overflow handling is replaced with an utility-oriented queue discard (see Fig. 3.21(a)). In addition, the dequeuing of packets, commonly done with the FIFO mechanism, is also replaced with an utility-based dequeuing

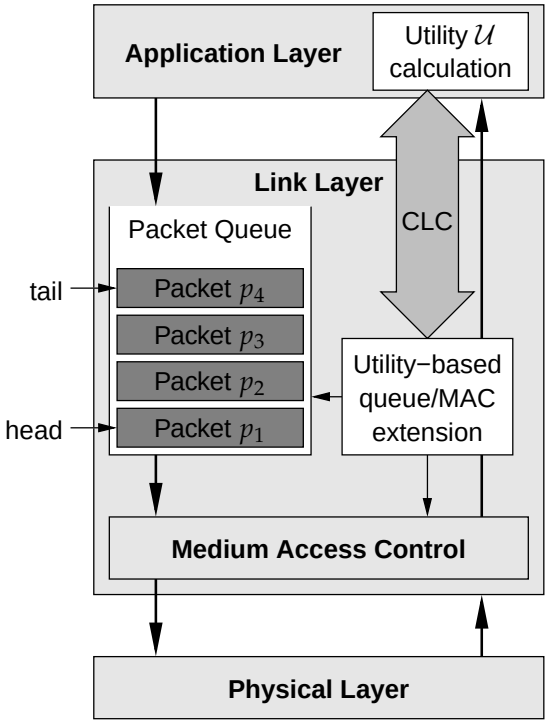


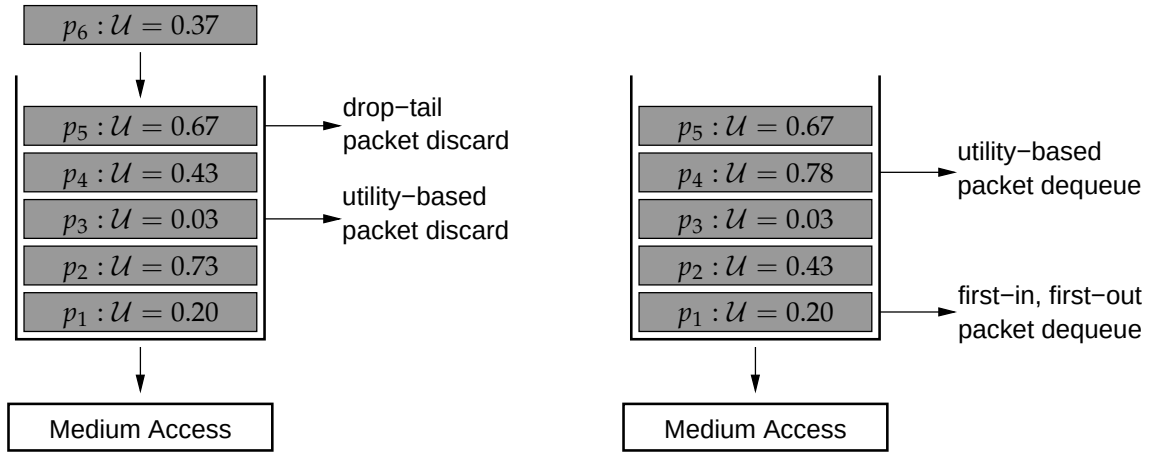
Figure (3.20) Cross-layer architecture for utility-based message prioritization in VANETs

policy (see Fig. 3.21(b)). The packet with the lowest utility is discarded, while the packet with the highest utility is selected to be send next. These new queue management mechanisms can only be optimally used if the utility values of the enqueued messages are re-evaluated constantly. Hence, the extension module needs to handle the message not only during the de- and enqueue procedures but also in between.

Each message needs to run through the MAC procedure of the communication device after passing the link layer message queue. A decentralized channel contention is used in most wireless devices today. This helps to realize an almost collision-free channel access without any central scheduling unit. The contention mechanism used in the IEEE 802.11 standard family in case of the Distributed Coordination Function (DCF) is shown in Fig. 3.22. The two MEs contend for the shared medium. After each transmission the medium has to remain idle for a general waiting period, the so-called Distributed Coordination Function Inter-frame Space (DIFS). Then the node individual back off, the so-called CW, is waited. If the medium is still idle after the CW has been completed the node will start its transmission. In case another node starts transmitting while the CW has not completed, the remaining CW back off slots will be transferred into the next contention period (see Fig. 3.22).

This contention mechanism was used as a starting point for the contention in the cross-layer architecture. Two main levers exist to influence the behavior of the contention: vary the defer period or change the CW calculation.

Usually the defer period is a fixed number of slots in form of the DIFS, however, to increase the channel access probability for packets with a high utility value, a changing



(a) Enqueue mechanisms **(b)** Dequeue mechanisms

Figure (3.21) Utility-based de- and enqueue mechanisms in comparison

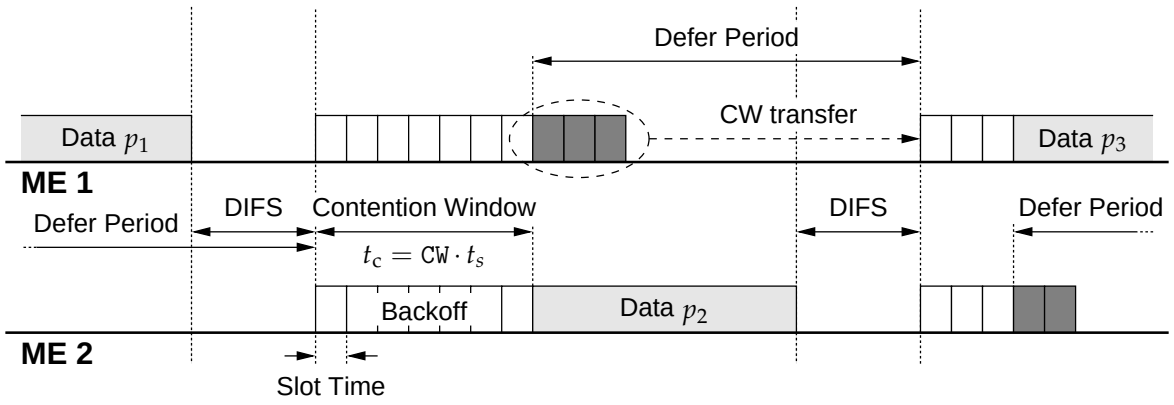


Figure (3.22) Example for channel contention mechanism using the Distributed Coordination Function for two Mobile Entities

defer period can be used. Therefore, the DIFS period needs to be calculated depending on the utility value of the respective message. The higher the utility value the smaller is the respective CW, thus, the message will have a higher channel access probability. This idea is similar to the different AIFS periods used in WAVE (see Fig. 3.11) or the IEEE 802.11e QoS standard [Sta05].

The number of backoff slots in a contention period is selected from the closed interval $\mathtt{CW} \in [0, \mathtt{aCWmin}]$, where $\mathtt{aCWmin} = 7$. This fixed CW size is replaced with a variable CW size in the cross-layer architecture for utility-based message dissemination. Therefore, the size of the CW is calculated individually for each message, depending on its utility value. Moreover, a smaller CW interval is used for messages with a high utility value, whereas messages with a low utility value have a larger CW interval. This leads to a higher channel access probability for messages with higher utility values.

An existing alternative to the cross-layer architecture is the IEEE 802.11e standard specified in [Sta05]. In contrast to the cross-layer architecture the WLAN QoS extension uses several message queues, identical to the queue concept of the WAVE standard (see Sec. 3.4.1 and Fig. 3.10). The queues could be used for different utility value ranges, therefore, leading to a prioritization on a queue basis. However, this can not reach the same performance as the suggested cross-layer architecture as published in [KAE+06, SSEE06a, SSEE06b].

3.5.4 System Simulation and Evaluation of Results

The cross-layer architecture introduced in Sec. 3.5.3 has been integrated into the OMNeT++ simulator and evaluated in different scenario settings. As underlying simulation basis the models presented in App. A have been used.

Simulation Model and Parameters

The application layer of the simulation model generates new messages and handles incoming messages. It differentiates between five message types, having different utility intervals and an individual utility calculation regarding four context parameters. Two variants for the benefit calculation exist in the model, a simple linear calculation (see Eqn. (3.22)), and a more realistic calculation based on Eqn. (3.14) and Eqn. (3.15).

$$\mathcal{U}_\mathrm{s} = \frac{2 \cdot \frac{\widehat{d_\mathrm{e}} - d_\mathrm{e}}{\widehat{d_\mathrm{e}}} + 2 \cdot \frac{\widehat{a_\mathrm{m}} - a_\mathrm{m}}{\widehat{a_\mathrm{m}}} + \frac{N_\mathrm{mc} - c_\mathrm{msg}}{N_\mathrm{mc}} + \frac{t_\mathrm{r}}{\widehat{t_\mathrm{r}}}}{6} \tag{3.22}$$

The message types and the corresponding parameters for the benefit functions are listed in Tab. 3.3. The following five message categories were used:

- Collision warning message: m_coll,
- Local danger warning message: m_danger,
- Road status message: m_road,
- Information message: m_info, and
- Weather message: m_weather.

A selection of benefit value curves is shown in Fig. 3.23. The curves have been generated using the parameters given in Tab. 3.3. All benefit values corresponding to one message are combined to calculate the message's overall utility value. An example for two benefit values is depicted in Fig. 3.23(d).

The contention mechanism was adapted compared to the contention mechanism used for conventional WLAN [LAN99]. Hence, the CW interval is defined as $[0, \texttt{aCWmax}]$. For the simple scenario the interval size was computed with Eqn. (3.24), while for the realistic scenario Eqn. (3.23) was applied. As described in Sec. 3.5.3 [pp. 57], the conventional WLAN uses a CW size of $[0, 7]$ for broadcast messages.

Scenario	Realistic					Simple
	m_{coll}	m_{danger}	m_{road}	m_{info}	m_{weather}	
Utility interval	$[0, 1.0]$	$[0, 0.9]$	$[0, 0.85]$	$[0, 0.8]$	$[0, 0.8]$	$[0, 1.0]$
Message age, Eqn. (3.14)	$a = 0$ $b = 0$ $c = 1$ $d = 1$	$a = 0$ $b = 600$ $c = 100$ $d = 2$	$a = 0$ $b = 1500$ $c = 400$ $d = 2$	$a = 0$ $b = 2200$ $c = 800$ $d = 2$	$a = 0$ $b = 2500$ $c = 1000$ $d = 2$	$\widehat{a_{\text{m}}} = 50$ s
Distance to event, Eqn. (3.14)	$a = 0$ $b = 200$ $c = 10$ $d = 2$	$a = 0$ $b = 400$ $c = 80$ $d = 2$	$a = 0$ $b = 1000$ $c = 400$ $d = 2$	$a = 0$ $b = 2000$ $c = 1000$ $d = 2$	$a = 0$ $b = 2500$ $c = 2000$ $d = 2$	$\widehat{d_{\text{e}}} = 300$ m
Last reception, Eqn. (3.15)	$a = 0$ $b = 0$ $c = 1$ $d = 2$	$a = 0$ $b = 5$ $c = 1$ $d = 2$	$a = 0$ $b = 7$ $c = 1$ $d = 2$	$a = 0$ $b = 10$ $c = 1$ $d = 2$	$a = 0$ $b = 10$ $c = 1$ $d = 2$	$\widehat{t_{\text{r}}} = 50$ s
Timeliness benefit	$\mathcal{B}_{\text{t}}(t) = \frac{1}{t^2} \quad \forall\, t \in [0, \infty)$					–
Inter-arrival times	5 s	5 s	5 s	6 s	6 s	–

Table (3.3) Benefit function parameters used for the system evaluation

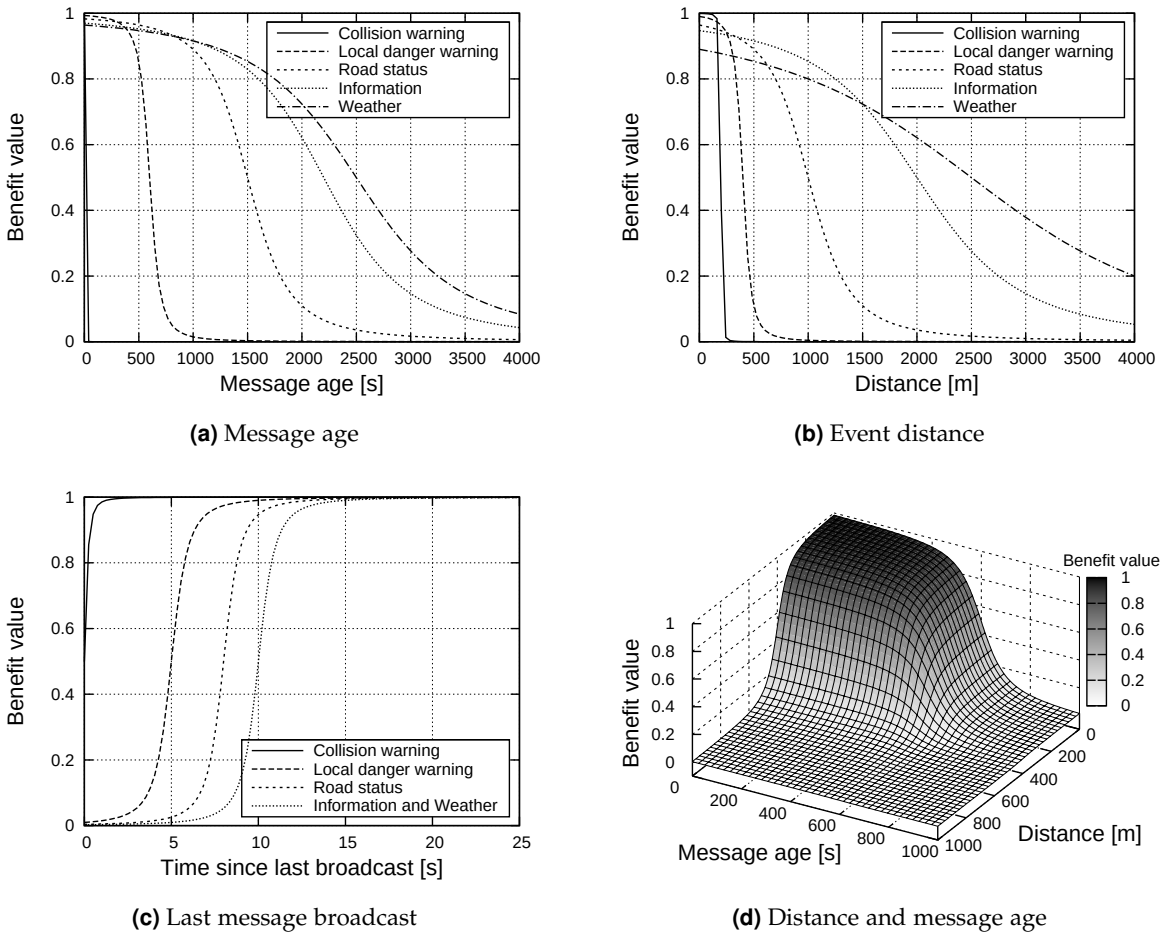
$$\texttt{aCWmax}(\mathcal{U}) = -100 \cdot \frac{\frac{\mathcal{U}-1}{5}}{\sqrt{1 + \left(\frac{\mathcal{U}-1}{5}\right)^2}} + 31 \qquad \texttt{aCWmax} \in [31, 121] \tag{3.23}$$

$$\texttt{aCWmax}(\mathcal{U}) = (1 - \mathcal{U}) \cdot 992 + 31 \qquad \texttt{aCWmax} \in [31, 1023] \tag{3.24}$$

Since the utility-based message prioritization is primarily designed for high-load scenarios the message queue in the link layer of a ME will be filled. In the simulation model a queue length of 80 packets was used. A trick has to be used to be able to simulate high-load settings in scenarios with several hundred nodes. The throughput available on the shared channel was set to a very low value of 0.1 Mbit/s. This is absolutely unrealistic for a real-life scenario, however, it reduced the number of simulation events significantly. This trick helped to keep the simulation environment scalable and simulate settings in reasonable time, while analyzing the system's performance at the same time. Due to the reduced throughput capacity the inter-arrival times needed to be altered from realistic settings to longer periods, resulting in the values given in Tab. 3.3.

Five message categories were used simultaneously to generate enough traffic to reach the throughput capacity limit and to be able to analyze different priority settings in parallel. A number of $N_{\text{n}} = 300$ nodes was used for the simulations. Each node continuously generated messages with a message size of 500 B and an exponentially distributed inter-arrival time. The average inter-arrival times are specified in Tab. 3.3.

Besides these fixed parameters, several parameters were varied to analyze the system's performance. In addition to the variation of the CW calculation and the different utility functions, four different scenarios were analyzed.



(a) Message age

(b) Event distance

(c) Last message broadcast

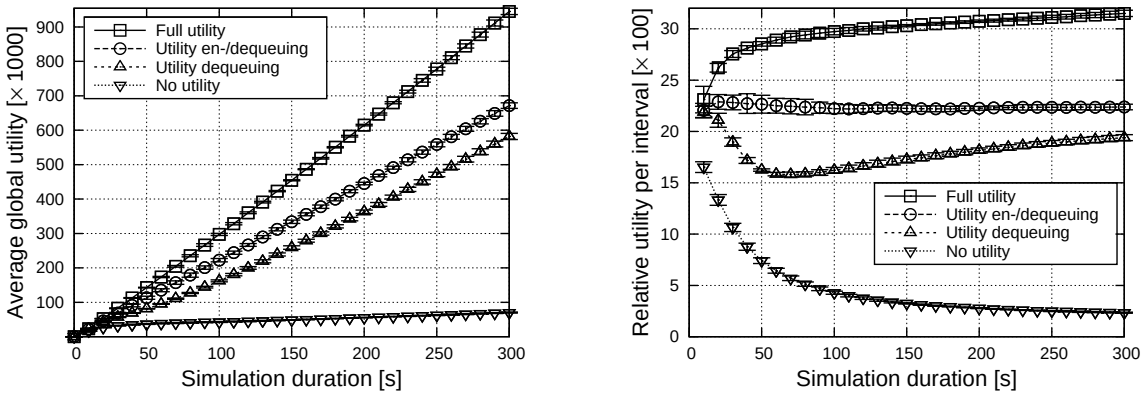
(d) Distance and message age

Figure (3.23) Benefit value curves for different context parameters (realistic scenario)

1. All utility functionality in the link layer deactivated (No utility),
2. Utility system with changed dequeue mechanism (Utility dequeuing),
3. Utility system with changed de- and enqueue mechanisms (Utility en-/dequeuing),
4. Utility system with changed de- and enqueue mechanisms and continuous utility re-evaluation (Full utility).

Simulation Results for Utility-based Message Dissemination in VANET Scenarios

Based on the specified simulation parameters many different scenarios have been simulated with the OMNeT++ simulator, applying the supporting models described in App. A. The evaluation results show the strengths and weaknesses of the approach and help to configure the concept for a real-life implementation. The presented results are equivalent to the results previously published in [ESKS06, KAE+06, SSEE06b, ES07]. In addition, they show the comparison between strictly linear utility functions (referred to as *simple scenario*) and



(a) Overall utility accumulation

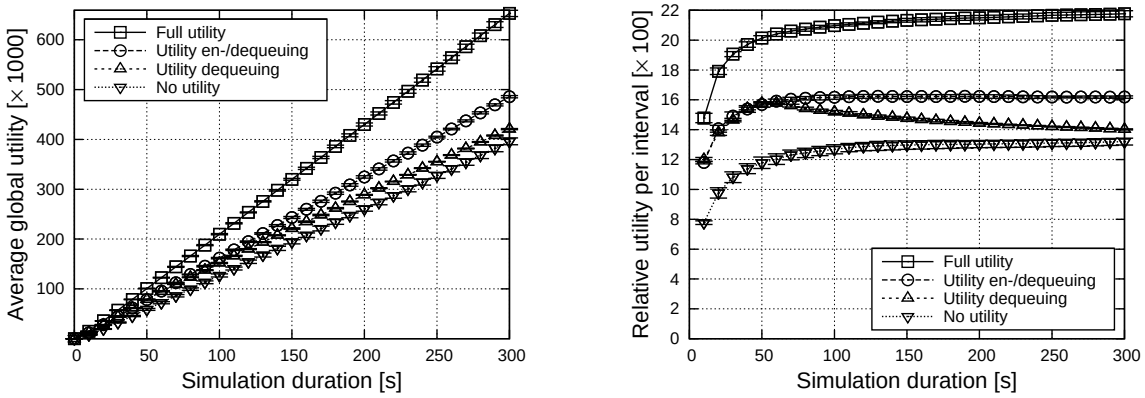
(b) Relative accumulated utility per interval

Figure (3.24) Globally accumulated network utility in the simple scenario

both concave and sigmoid functions (referred to as *realistic scenario*). A comparison of the utility-based message dissemination scheme and the IEEE 802.11e QoS enhancement has been done and presented in [KAE+06, SSEE06a, SSEE06b].

The most important result is the proof that the approach works as expected. To show this, the globally accumulated network utility is determined. All received messages are evaluated and their utility value is added to the global network utility value. Once every 5 s the global network utility is evaluated. The results for the different scenario settings, using the simple utility functions, is shown in Fig. 3.24. In Fig. 3.24(a) the accumulated values are shown, while in Fig. 3.24(b) the gained utility per interval is depicted. Both figures show the improvement that the dissemination approach causes. Using a network utility evaluation will increase the overall utility in the network, optimizing the utility of the respective services to the users. On average during the first 30 s to 50 s the message queues in the MEs need to fill up. Hence, the maximum effect can be detected only after this setup time. This effect can be seen in Fig. 3.24(b) for the *Utility dequeuing*-plot. The gained utility per interval decreases during the first 60 s, to constantly increase afterwords. Especially the *Full utility* setting shows great improvements, which is to be expected. In this setting all possible improvements are used: utility-based de-/enqueuing and regular utility re-evaluation.

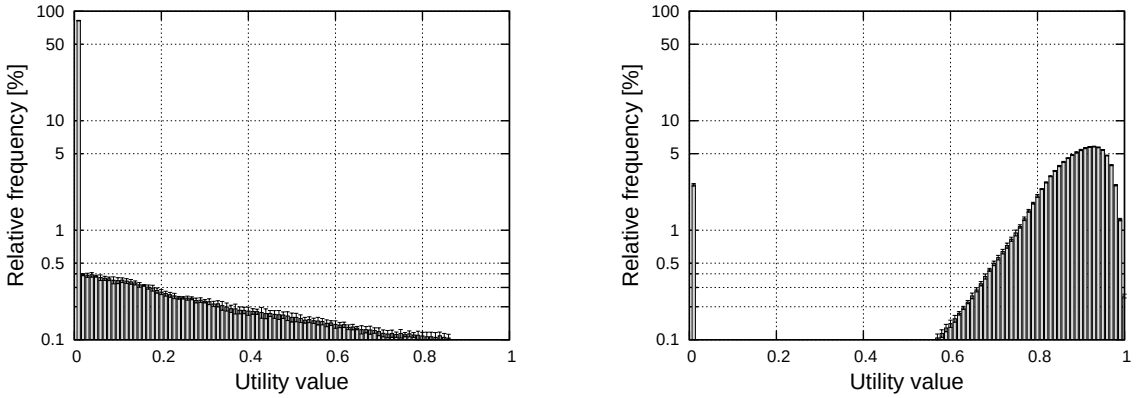
An equivalent result for the realistic scenario settings is depicted in Fig. 3.25. Like in the simple scenario the improvement can be seen. Nevertheless, a change due to the different utility functions can be recognized. The results in Fig. 3.25(b) show that in the *Utility dequeuing* setting a decrease in the gained utility per interval is occurring. The reason for this is the fact that the full message queue prevents new messages with a high utility value from entering the queue. Hence, the accumulated utility per queue decreases until new messages can be added again. For all other settings an increase can be seen. The graphs in Fig. 3.25(b) depict that the increase saturates after some time, hence, the system practically reaches a steady state. The important fact is that the utility-based message dissemination shows a much higher steady state than the conventional system.



(a) Overall utility accumulation

(b) Relative accumulated utility per interval

Figure (3.25) Globally accumulated network utility in the realistic scenario

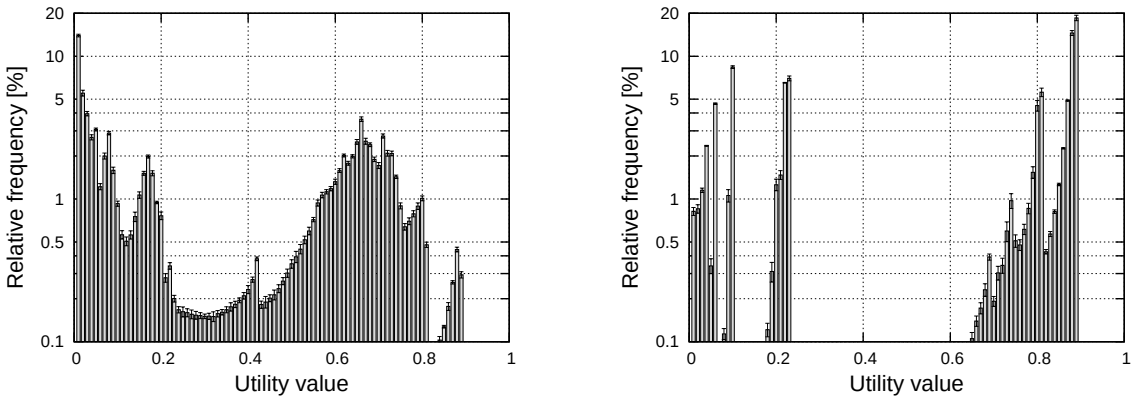


(a) No utility evaluation

(b) Full utility-based dissemination

Figure (3.26) Utility value occurrences in the simple scenario

After knowing that the approach is valid and increases the global network utility significantly, several system tests have been done to further evaluate the approach. In addition, a detailed system evaluation helps to understand its behavior and can be used as a basis for the configuration of a real-life system. The selection of utility-functions influences the occurrence of utility values. In Fig. 3.26 the occurrence of different utility value intervals is shown. Each interval has the size of 0.01, hence, 100 intervals are shown. In the scenario without utility evaluation the intervals at the lower end dominate (see Fig. 3.26(a)), while in the scenario with fully activated utility-based dissemination the intervals at the upper end dominate (see Fig. 3.26(b)). Without differentiation a packet stays in the message queue for quite some time, since it is added at the end of the queue. During the waiting period the utility value normally decreases. Therefore, the low utility values dominate the occurrences. A reciprocal behavior is seen for the utility-based diffusion scheme. The messages having



(a) No utility evaluation

(b) Full utility-based dissemination

Figure (3.27) Utility value occurrences in the realistic scenario

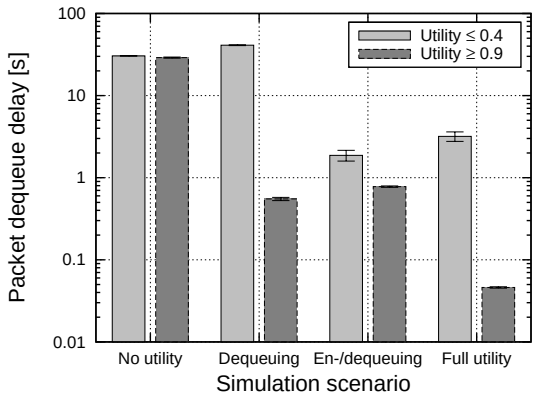


Figure (3.28) Packet dequeuing delay comparison for realistic scenario settings

a high utility value are preferably distributed, leading to a high occurrence of high utility values.

The occurrence evaluation for the realistic scenario is depicted in Fig. 3.27. The results are equivalent to the simple scenario results, however, the distributions are not nearly as smooth. The configuration of the utility functions strongly influences the value distribution. In addition the scenario specific parameters (number of nodes, node speed, message queue length) also influence the distribution in an indirect fashion. Only messages of type m_{coll} can have benefit values above 0.9. The very strong utility value decline for m_{coll} depending on the message age (see Fig. 3.23(a)) causes all messages of this type that waited longer than 50 ms on average to be deleted. This is a strict but realistic setting for the respective message type. Moreover, the results show that the utility functions need to be configured very carefully. Further, the use of linear utility functions results in a very smooth value distribution in comparison to the concave and sigmoid functions. The linear utility functions are not capable of separating different message types as consequently as the concave and sigmoid functions in terms of utility value calculation. This is mainly due to their minor

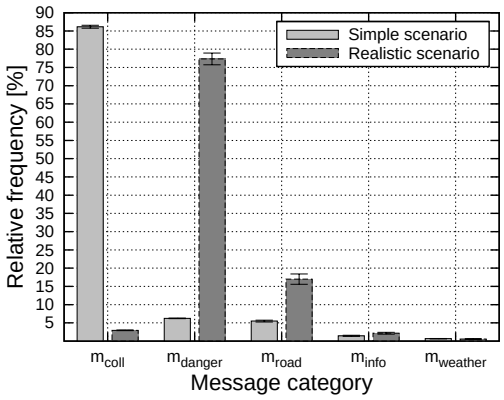


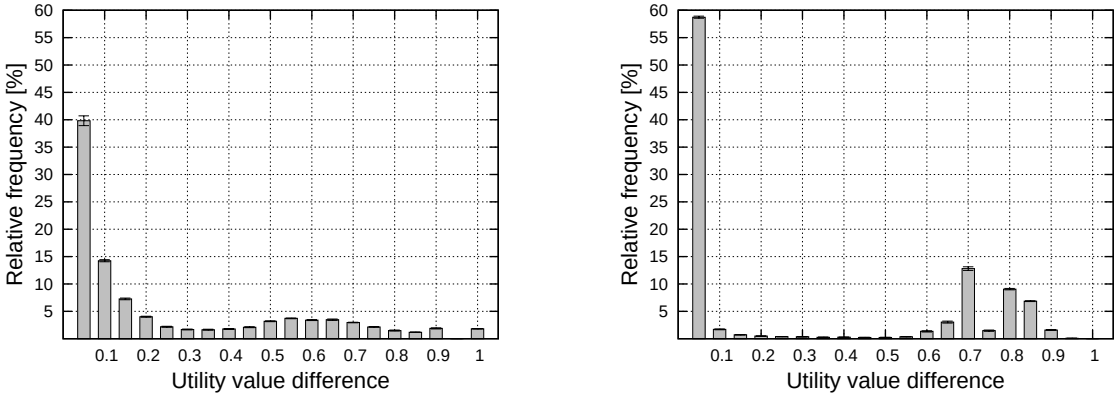
Figure (3.29) Message category distribution for utility-based message dissemination

decline compared to the exponential decline used for the non-linear functions. About 30% of all values range below a utility value of 0.3 (see Fig. 3.27(b)). This is caused by the fast declining sigmoid benefit functions. Hence, many packets have a low utility value. A more careful benefit function configuration will mitigate this behavior.

For the performance of an IVC system the packet delay is an important factor. The packet dequeuing delay for the realistic settings is shown in Fig. 3.28. The results show that without differentiation the different benefit values are handled equally. Using the differentiation leads to a delay difference between high benefit values and low benefit values. Further, the average overall delay can be reduced significantly. A packet in the conventional scenario suffers from an average delay of up to 30 s. In the scenarios using full utility-based message differentiation this is reduced to an average delay between 50 ms and 3 s. Therefore, even messages requiring a low distribution delay can be handled sufficiently in the high load case using the utility-based approach.

The different message categories are generated with almost the same frequency (see Tab. 3.3). The utility-based dissemination scheme acts as a filter for the messages. The distribution at the receiver side is shown in Fig. 3.29 for both scenario types. While the simple scenario clearly favors message category m_{coll} the realistic scenario predominantly sends message category m_{danger}. This different behavior comes from the different utility functions and shows how the filter characteristics can be influenced. Both settings are not very well suited for a realistic scenario and need to be adapted, the filter should not prefer one message category so drastically. Nevertheless, the result shows the possibilities of the utility-based message dissemination concept and its strong influence on the dissemination process.

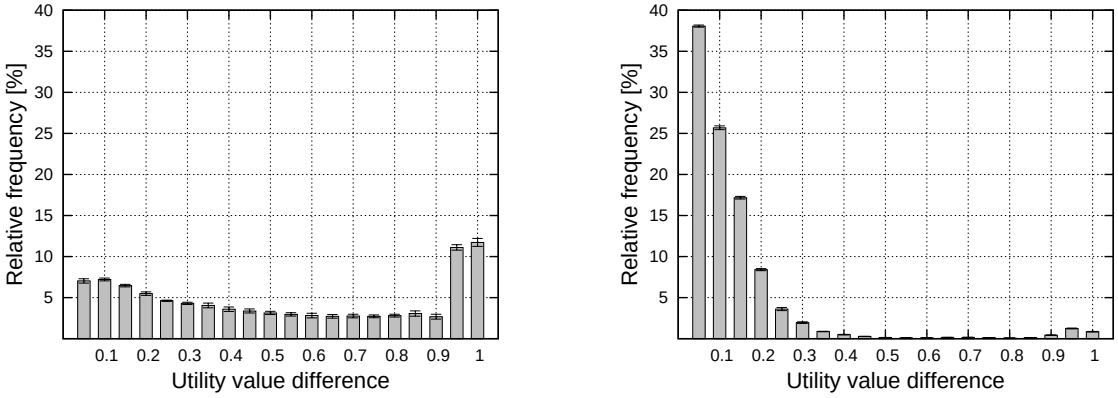
In the current approach, the sending ME determines the utility value based on its own context information. This most likely correlates with the context information of surrounding MEs. However, a slight difference in the context information will always exist, leading to a difference in the calculated utility value. This difference has been collected and evaluated during the simulations. A histogram divided into 20 intervals having a width of 0.05 per interval was generated for the realistic scenario settings. Fig. 3.30 shows the results for a scenario without benefit and a scenario with full utility-based message dissemination.



(a) No utility evaluation

(b) Full utility-based dissemination

Figure (3.30) Occurrences of utility differences between sender and receiver in the realistic scenario



(a) No utility evaluation

(b) Full utility-based dissemination

Figure (3.31) Occurrences of utility differences between sender and receiver in the simple scenario

Without differentiation the utility value differences are distributed over all intervals with the peak of 40% at the smallest possible difference of ≤ 0.05 (see Fig. 3.30(a)). Therefore, the assumption that context information differs in most cases only marginal between neighboring nodes is verified. Activating the utility-based message differentiation increases the peak of the first interval by almost 50% to a value of roughly 60%. However, some of the intervals for larger utility value differences also show a rather strong increase. This can mainly be explained with the strict filtering of message duplicates. If a message is received multiple times, the duplicates add almost no utility, hence, a relatively large utility difference between sender and receiver can occur. Another reason for this effect is the node individual utility calculation, which can lead to significant differences in the utility value, depending on the benefit function configurations.

The equivalent results for the simple utility function scenarios are depicted in Fig. 3.31. Without differentiation the utility value differences are almost equally distributed (see Fig. 3.31(a)). This changes drastically when using the differentiation with the linear utility function. The small utility value differences clearly dominate (see Fig. 3.31(b)). The linear utility function decreases more moderately, hence, in general the difference variations are smaller and very similar for all message categories. This leads to a smoother distribution, however, this does not have a practical advantage.

3.5.5 Quantification of the Utility-based Information Distribution

The suggested utility-based information distribution scheme is a very promising approach to handle IVC in high load situations especially. The scheme helps to select and distribute the most relevant packets with low delay values between 50 ms to 3 s. Even though the selection is done at the sending ME the utility value estimation is sufficiently accurate to increase the global utility caused by IVC. The presented approach can be easily integrated into an IVC architecture and can coexist or even be used in combination with the IEEE 802.11p/WAVE communication standard or the IEEE 802.11e QoS extension for conventional WLAN. With an optimal benefit function setup the utility gain can be improved significantly, as shown in Fig. 3.24(a).

A very important factor in the configuration of the utility-based message differentiation is the selection and parametrization of the benefit functions. The benefit functions have the strongest influence on the utility values. However, parameters like the message queue length, network load, vehicle speed, and number of neighbors also influence the utility values as context information. Therefore, for the setup of the benefit functions the scenario context and the distribution requirements for the respective message category need to be incorporated. The given results show the influence of different benefit functions and can be used to support the configuration process for a real-life implementation.

In any VANET scenario the network load can get so high that an overload situation occurs and distribution delays increase significantly. Therefore, a message differentiation is needed to optimize the system's performance. Since many different V2V services will run in parallel, a utility-based message differentiation scheme promises to solve the overload situation in the best interest for the users. This has been demonstrated with the results in Sec. 3.5.4. The system can be used to increase the scalability of IVC systems in dense VANET scenarios, since the most important messages, in terms of utility, are preferred. Moreover, the approach is highly adaptable to future demands and new message categories, which is required for the dynamic field of VNs. Very promising is the fact that messages of different importance and type can be handled by the same MAC-layer, using the same communication device.

3.6 Information Distribution using the Content-Aware Mobile Data Request Protocol

Especially in a VANET, where most data exchanges are unrequested and broadcast-based, the use of a conventional routing protocol is useless and of not much help. This is also

mainly due to the anonymity of nodes, hence, no specific node can be selected for a route search. But solely relying on the broadcast-based data distribution is not sufficient in some situations, since a node might need specific information which however is not broadcast by any of the surrounding nodes at the same time. In this case it would be useful to be able to send a request to the neighbors or the network environment in general, asking for the required information.

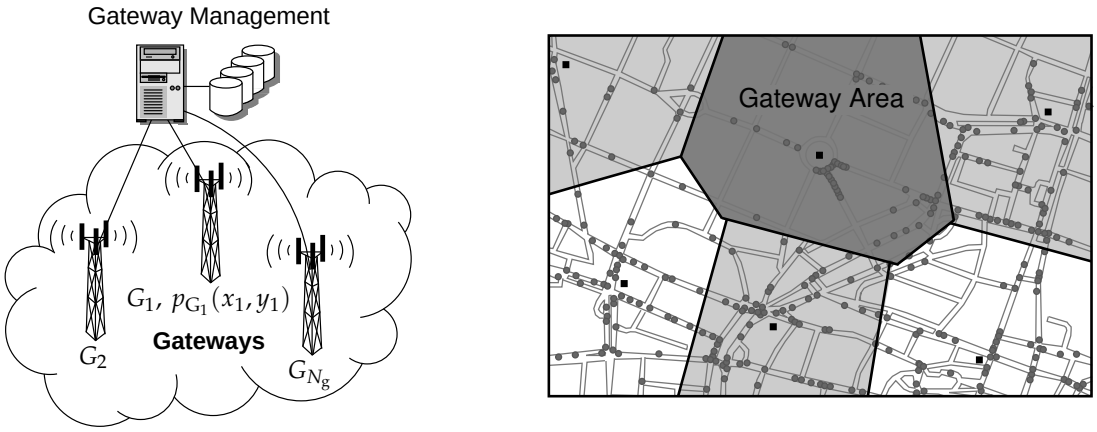
While most content is not directly addressable by nodes, since its existence is only known to nodes that already received the respective data, it can become addressable for example by introducing content categories. An example for such a category could be the *parking capacity* in a given city, the traffic conditions in a certain area, or the support and status information for the supporting security architecture needed in a V2V communication system. Therefore, introducing content categories, making information directly addressable for a content-aware protocol, will provide new ways of data distribution. This can reduce the load on the network, thus improve system scalability. Therefore, the content can be distributed in an optimized fashion compared to a simple broadcast solution.

An important feature for a protocol handling addressable information is *content awareness*. The term *content-aware* stands for any routing or data distribution protocol which itself relates to the content it transports. The Mobile Data Request Protocol (MDRP), which will be introduced in the following sections is one example for a content-aware protocol. It is content-aware in two ways. First, the content is used similar to a network address, and secondly the protocol is constantly aware of which content the node currently requested to be able to gather the respective data from incoming packets.

Basically, a content-aware message distribution mechanism can be realized in any MANET or VANET scenario. Nevertheless, a few requirements have to be fulfilled to implement such a mechanism. In addition, some scenarios are more beneficial for content-aware message distribution than others. In the assumed scenario, gateway nodes distribute content initially. Hence, the gateway nodes act as brokers for the information provided primarily by the content servers in the Backend network. Besides the initial content distribution, the gateways are also used to optimize the success rate of content requests. Each request is sent with a geographical direction. This direction is always towards the closest gateway of the requesting node. Since the gateways are omniscient due to the direct connection to the content servers, the requests do not have to be flooded throughout the whole network. However, to be able to find the closest gateway, an announcement mechanism is required, making gateway information available to nodes. Hence, before detailing MDRP, the addressing of gateway nodes in VN scenarios is discussed. The protocol MDRP has previously been published in [Eic07a].

3.6.1 Addressing and Identification of Gateway Nodes

As soon as gateway nodes are used to extend a VANET to a VN a management of these gateways as well as an efficient identification mechanism is needed. The main goal of adding gateways is to improve the functionalities of a VANET, for example, improve information diffusion or reduce message load in the network. Few studies on gateway usage in distributed environments have been made previously. However, the mechanisms for



(a) Gateway management **(b)** Gateway areas

Figure (3.32) Gateway organization in vehicular network scenarios

gateway announcement presented in [MB05, BWSF03, XB02] are viable approaches to solve this challenge.

The gateway notification presented in the following is similar to the techniques presented in [XB02, BWSF03], nevertheless, several adaptations have been made to improve the performance in VANET scenarios. In both publications, active and passive gateway discovery mechanisms are described from the MEs perspective. While the active discovery requires nodes to initiate a search request, which can lead to network congestion, the passive discovery passes the responsibility over to the gateways.

In the following, a gateway organization and discovery mechanism is suggested for VANETs. Each gateway node is responsible for a defined area of the scenario. It broadcasts notification messages within this *gateway area* only (see Fig. 3.32(b)). The partitioning of the gateway areas is done automatically. This is done by the *gateway management*, which is a server in the Backend as shown in Fig. 3.32(a). Each gateway G_i has an individual position $p_{G_i}(x_i, y_i)$, which is communicated to the management. The gateway management stores all positions and communicates the closest neighbors to each gateway node. Using this information, gateways can individually calculate their area of responsibility.

Each gateway continuously broadcasts its position and information on the closest neighbor gateways to the MEs in its vicinity, using notification messages. Similar to the dissemination areas described in Sec. 3.3.1 the mobile nodes forward the notification as long as they are still located within the respective gateway area. This notification mechanism makes sure that an optimal coverage can be achieved while still limiting the number of notification messages.

3.6.2 Evaluation of the Gateway Notification Mechanism

To get an idea of the gateway notification performance, several simulations have been conducted. Using the MGM model and the supporting VANET models described in App. A

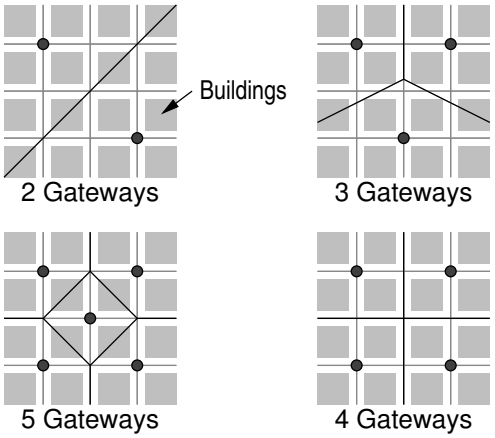
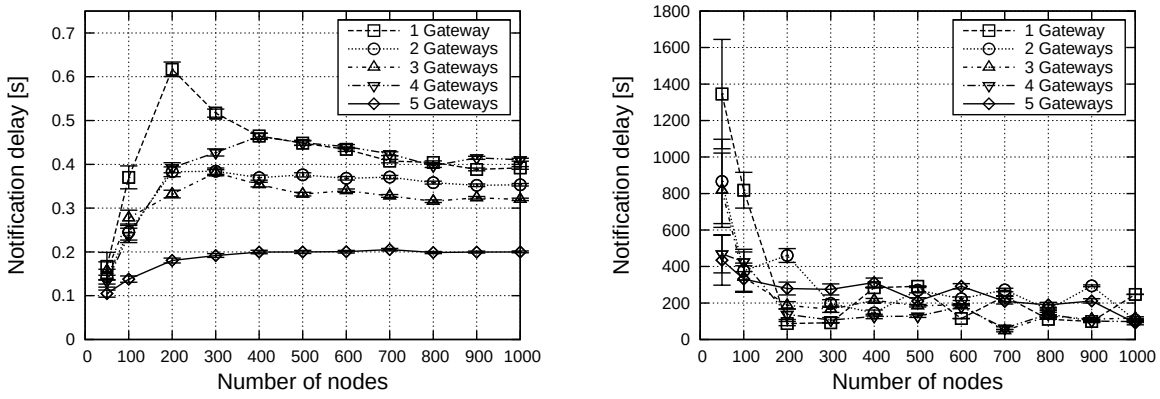


Figure (3.33) Gateway positions for different gateway densities on the Manhattan Grid



(a) Overall average notification message reception delay

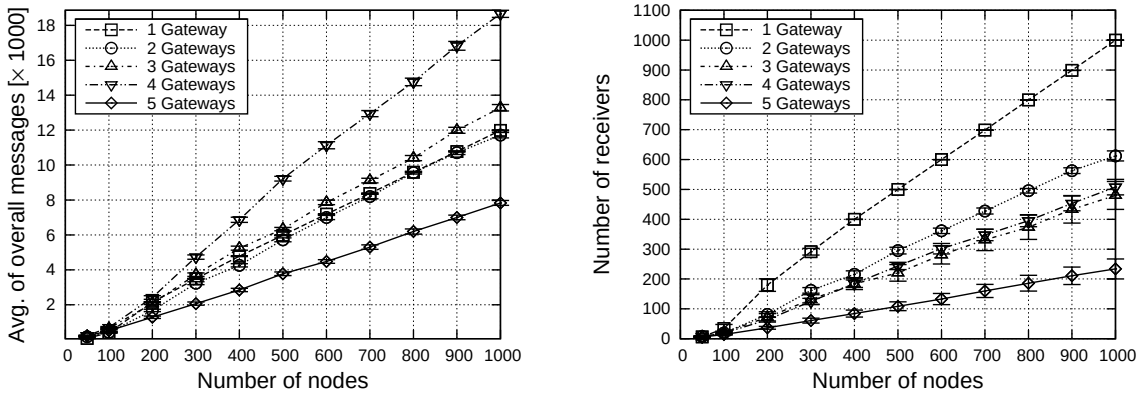
(b) Overall average gateway notification delay

Figure (3.34) Delay values of the gateway notification message distribution

various node and gateway densities have been simulated. All setups for the different gateway densities are depicted in Fig. 3.33, except the setting for just one gateway.

The single gateway was positioned at the center of the scenario and was responsible for the full scenario. In the other setups, the gateways were only responsible to serve their gateway area. Notification messages were only forwarded within the gateway area and not beyond.

The gateway nodes sent a notification packet of 600 B once every notification interval (t_n) of $t_\mathrm{n} = 5$ min. The simulation duration was set to $t_\mathrm{sim} = 1$ h. Hence, the number of initially sent notification packets can be determined by $N_\mathrm{m} = t_\mathrm{sim}/t_\mathrm{n} \cdot N_\mathrm{g}$. The MEs receiving a notification for the first time forward it to the neighboring nodes to increase the coverage range, however, this is only the case as long as the forwarding node is located within the respective gateway area. Several parameters have been evaluated in the course of the simulations. The most important ones are discussed briefly in the following.



(a) Overall number of messages caused by gateway notification distribution

(b) Number of receivers depending on node density

Figure (3.35) Message and receiver statistics for the gateway notification message distribution

Above all, three main issues have to be evaluated: Distribution delay of notification messages, number of reached nodes, and packet load caused by the notification process. To evaluate these values several simulations with different node densities have been run. The delay values for the packet distribution as well as the initial gateway notification delay are shown in Fig. 3.34. In Fig. 3.34(a) the average distribution delay of a notification message is depicted. All plots start at a low delay value which increases for higher node densities, irrespective of the number of gateways. The reason for this is the fact that for low densities the network is not fully connected. Hence, only the nodes in the direct vicinity of the gateway receive the message. In scenarios with $N_\mathrm{n} \geq 200$ the network can be assumed to be connected in most cases. At this point the graph for $N_\mathrm{g} = 1$ reaches a maximum, which can be explained with the long multihop connections leading to the long average delay. In Fig. 3.34(b) the absolute notification delay, which is the time measured from the start of the simulation until a node receives a notification message for the first time, is shown. This value strongly depends on the mobility of the MEs, hence, the variability is much bigger than in Fig. 3.34(b). However, the absolute delay is below 300 s in connected scenarios. The results prove that even in the scenario with $N_\mathrm{n} = 50$ all nodes will receive a notification at least once during t_sim.

The number of messages and the related distribution degree are depicted in Fig. 3.35. The number of messages, shown in Fig. 3.35(a), is an important value to estimate the network load caused by the notification mechanism. The number of messages for one to three gateways are quite similar and range from just above zero to about 13 k messages. Interestingly an increased gateway density does not automatically reduce the number of messages. This is proven by the values for the scenario with four gateways, which has by far the most sent messages. In contrast, the scenario with five gateways has by far the least number of messages. This result proves that the placement of gateways has a significant impact on the number of messages required to distribute the notification. The street intersections in the MGM model are key positions for message forwarding. No

obstruction due to buildings exists at intersections, since vehicles are only positioned on roads. The more streets and intersections are located directly on the border of a gateway area the bigger is the increase in messages. All vehicles located on the respective road practically belong to two different gateway areas, hence, they forward messages for both areas.

The number of receivers shown in Fig. 3.35(b) gives an idea how many nodes are reached by a specific notification message. In a scenario with one gateway the optimal value is equal to N_{n}, where all nodes are reached. In all cases where $N_{\mathrm{g}} \geq 2$ the optimal value is equal to $N_{\mathrm{n}}/N_{\mathrm{g}}$ assuming an equal node distribution in the scenario and an equal distribution of nodes to the existing gateway areas. The scenario with five gateways almost reaches this optimum, while the scenario with four gateways performs the worst. The respective results for $N_{\mathrm{g}} = 4$ are doubled compared to the optimum.

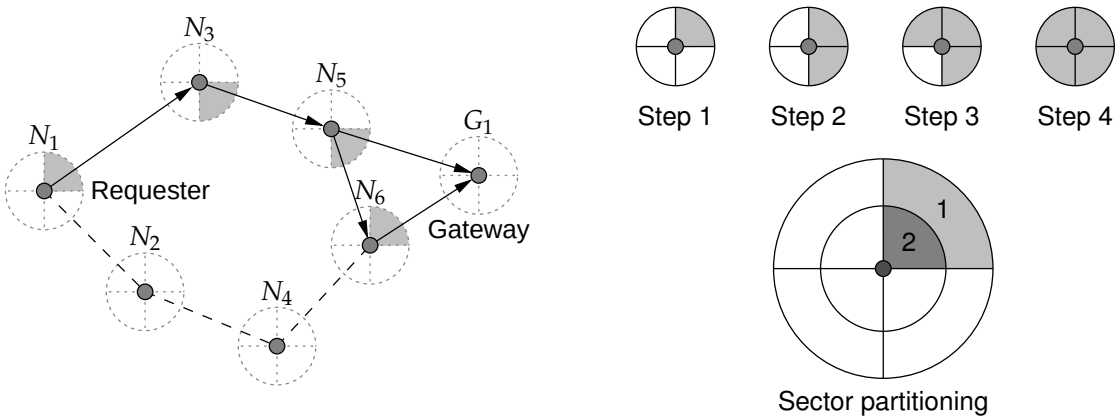
The results demonstrate that the gateway density has an immediate influence on the notification quality. Moreover, the position of gateways can lead to negative distribution effects due to the symmetry of a scenario, as seen for the case with four gateway nodes. The gateway management helps to improve the information status on gateways, since neighboring gateways can also be included into the notification. Overall can be summarized that a flooding-based notification distribution works with reasonable delay and high delivery success.

The difficulties of gateway usage in MANET scenarios have not been fully solved with the presented results, however, they give an idea on the performance and are a valuable basis for related research issues. Moreover, the presented results on gateway usage in VANETs are an important prerequisite for several remaining topics in this thesis. The Mobile Data Request Protocol (MDRP) relies on gateway usage as well as some of the security mechanisms described in Chap. 4.

3.6.3 Request-based Data Dissemination using MDRP – Protocol Description

In the following section the functionality of the Mobile Data Request Protocol (MDRP) will be introduced, which was previously published in [Eic07a]. The main peer types differentiated by MDRP are *nodes* and *gateways*. Since a gateway is the origin of content it will distribute data initially. In addition, it can be the final peer in a content request process, the peer providing the content at last. The regular nodes can either be a requesting node or a forwarding node.

Neighborhood Discovery: Before a node can start to request content or handle incoming requests, it needs basic knowledge of its current neighborhood. Each node holds a table with the neighbor information containing the respective node positions. The table is updated based on both incoming messages and the use of *Hello*-messages. If no messages have been received for a given time and the table entries are older than the specified MDRP aging interval (t_{age}), the node starts sending *Hello*-messages to update the neighbor information. The *Hello*-messages also contain information on the gateway responsible for the area, which improves the gateway notification. The neighborhood is divided into four sectors (see Fig. 3.36(b)), which helps to reduce the number of nodes that shall react to a request. In dense scenarios ($N_{\mathrm{ne}} \geq 16$), each sector can be divided further. Two sector parts are suggested, an outer part (1) and an inner part (2),



(a) Communication example for MDRP

(b) Sector selection process using multiple steps

Figure (3.36) Introduction to the MDRP content request process

as shown in Fig. 3.36(b). Thus, only the nodes farthest away from the sender will react to the request in the first step.

Request Process: If a node needs to request content, it first looks up the closest gateway from its database and identifies the gateway's sector, relying on a gateway notification mechanism. In the second step the neighbor database of the chosen sector is evaluated. If the sector contains nodes, the request is sent, being relevant only for the nodes located in the selected sector. In case the sector contains no nodes or a first request has not been answered after the MDRP wait interval (t_{wait}), the adjacent sector to the right is added to the list of recipients. If necessary this process is continued until all sectors are in the list of recipients (see Fig. 3.36(b)). A node receiving a request, first checks if it can provide the requested content. If this is the case it replies directly, addressing the sector where the request came from. Otherwise the request is forwarded to neighbors in the direction of the gateway. Again the sector management introduced above is used. In Fig. 3.36(a) an example for a communication process is shown. A reply is always sent directly after the request has been received, while the message forwarding is slightly delayed. This enables the protocol to detect the replies and halt the forwarding process to reduce the network load.

Reply Process: Any node in the network along the path towards the gateway can reply to a received request, provided that it holds the requested content. If none of the nodes along the path towards the gateway can reply, the gateway is the assured content source, being able to provide any requested content. During a request process, like the one shown in Fig. 3.36(a), each intermediate node saves the incoming requests and its originating sector in a reply table. This table is used to forward the replies to the designated sector in the network. As soon as the corresponding content reply has been forwarded, the entry in the reply table is removed. The reply table is reviewed constantly, to age entries and delete them accordingly. In the case where a request is replied to by both an intermediate node (e.g. node N_6) and the gateway, the next

upstream node(s) sharing both reply paths acts as a filter. Only the first reply received will be forwarded. Due to the deletion of the respective reply table entry after the first reply has been handled, all following replies can not be routed and are therefore deleted. This reduces the occurrence of multiple replies, therefore, the network load is reduced.

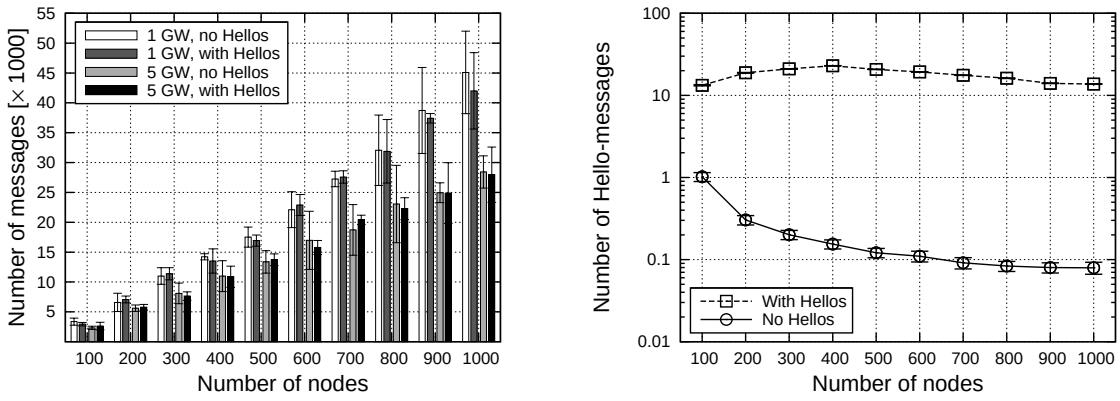
Expansion of Request Messages: A request can travel multiple hops until a gateway is reached (see Fig. 3.36(a)), therefore, it is likely that intermediate nodes also have requests to send. MDRP allows to expand received requests and attach additional content requests. This has no implications on the request process itself, however, the reply process has to be adapted slightly. Each node expanding a request has to scan the incoming replies for the desired content. If the requested content is contained in a received reply, the node extracts the content and forwards the reduced content reply message towards the remaining receivers. This feature makes MDRP content-aware.

Large Content Replies with Fragmentation: Content replies can become larger than the Message Protocol Data Unit (MPDU) of the used wireless transmission technology, for example IEEE 802.11 WLAN uses a MPDU-size of 2346 B [LAN99, CWKS97]. Hence, a fragmentation mechanism is required, to be able to send larger content messages fragmented into several smaller packets. MDRP handles each request and reply pair coherently. The requesting node selects a request-ID, which then identifies all messages during the full protocol process until its completion. The first packet of a reply process contains status information on the reply. This includes the request-ID, content size, and the number of fragments. Each of the following packets contains the request-ID and the current fragment sequence number. The requesting node can then collect all fragments, re-request potentially lost fragments, and combine the parts to the desired content.

3.6.4 Performance Evaluation of MDRP in VANET Scenarios

To evaluate the performance of MDRP, the protocol has been implemented for the OMNeT++ simulation framework. Again the simulation models described in App. A have been used as a basis. Several different settings and node densities have been simulated to test the characteristics of the protocol. Two variants of the protocol, one with continuous use of Hello-messages (With Hellos) and one with Hello-messages used on demand only (No Hellos), have been compared. In addition, two different gateway configurations have been used: One gateway and five gateways. They have been placed according to Fig. 3.33. Thus, the performance results on gateway notifications presented in Sec. 3.6.2 can be used as a basis for the gateway usage of MDRP. The simulation time was set to 1800 s. The continuous Hello-message interval was set to 60 s. The data to be requested had a size of 50 kB. With an assumed signature size of 200 B the data was fragmented into 30 pieces of 2 kB each. The gateways used a notification interval of 300 s. Once the simulation starts and the nodes are aware of at least one gateway they try to obtain the data. Based on the following simulation results the performance and scalability characteristics of MDRP can be determined.

In Fig. 3.37(a) the overall number of MDRP-related messages existing during a full simulation run are depicted. The diagram shows that increasing the gateway density reduces



(a) Comparison of overall MDRP messages

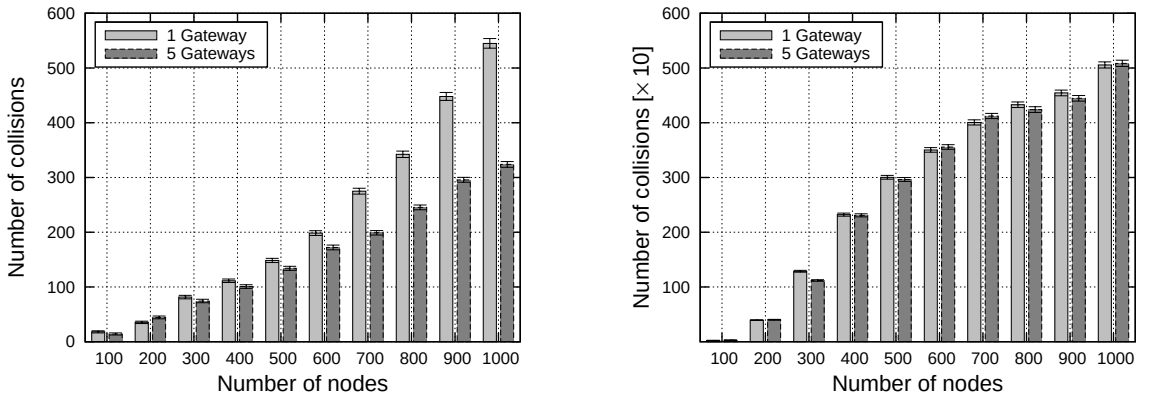
(b) Hello-message evaluation for MDRP

Figure (3.37) Average number of overall MDRP messages and Hello-message influence on the protocol performance

the number of messages, while introducing continuous Hello-messages practically has no effect. Moreover, the relative number of MDRP messages per node required to fully distribute the data remains between 35 and 45 messages, which shows the efficiency and high performance of the protocol. The number of Hello-messages sent per node in a scenario with one gateway is depicted in Fig. 3.37(b). The result shows that Hello-messages are only needed in scenarios with low densities. For higher densities the graph for "No Hellos" shows a significant decrease of Hello-messages. The other MDRP messages are sufficient to keep the neighbor lists up-to-date. Therefore, no continuous notification mechanism is required for the protocol. This result is also supported by following results.

An important characteristic for any broadcast-based dissemination protocol is the number of collisions it causes. The results for MDRP are shown in Fig. 3.38. The figures show that the continuous traffic due to Hello-messages results in a significant increase of packet collisions (up to a factor of 10). In addition, the result in Fig. 3.38(a) supports the statement that a higher gateway density improves the protocol performance by reducing the number of messages, since the number of collisions is decreased significantly by increasing the number of gateways. This effect is predominantly true for dense scenarios, since the reduction of nodes per gateway is much bigger when adding a gateway to a dense scenario.

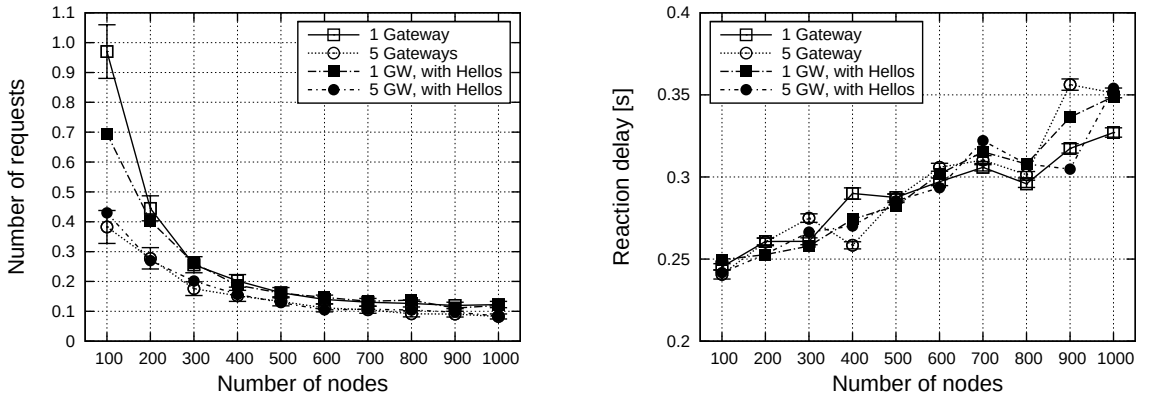
After having looked at the more general performance parameters, the request and reply phases of MDRP are evaluated. In Fig. 3.39 the results for the request process are depicted. Both plots prove that continuous Hello-messages have no benefit for the request phase whatsoever. The only case where a better result is achieved is for one gateway and 100 nodes. In all other cases the performance of both variants is equal. In Fig. 3.39(a) the average number of requests per node is shown. The caching and distribution techniques of the protocol are very efficient, hence, the number of requests per node can be lowered to 0.1 requests. In the scenario with five gateways, nodes need less requests to obtain data. The reaction time of MDRP is an important value besides the number of requests. In Fig. 3.39(b) the reaction times are depicted, which is the time elapsing between the sending of the



(a) MDRP using Hello-messages on demand

(b) MDRP using continuous Hello-message sending

Figure (3.38) Average number of overall detected collisions using MDRP



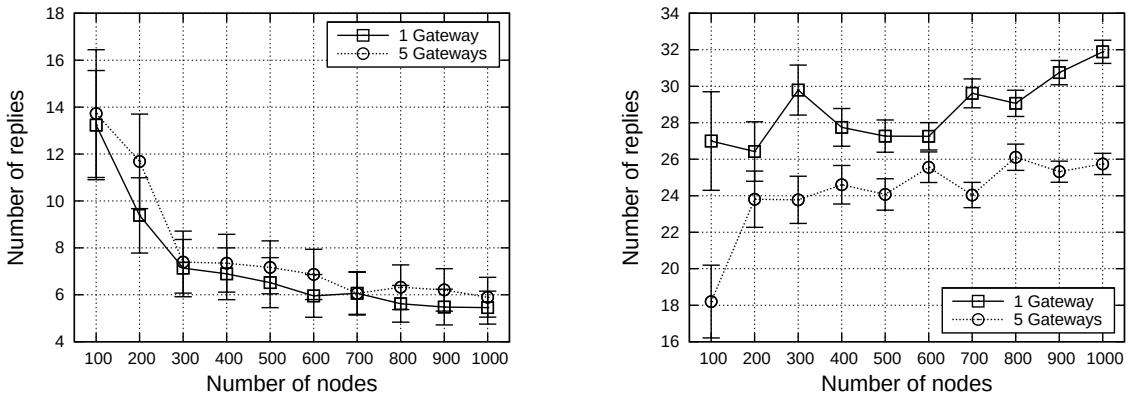
(a) Average requests per node

(b) Protocol reaction delay

Figure (3.39) MDRP request process evaluation

request and the reception of the first reply. They range between 250 ms and 350 ms, which is a very good value for a distributed protocol. The longer reaction time for scenarios with more nodes can be explained by the higher data traffic and the increase in data loss due to packet collisions.

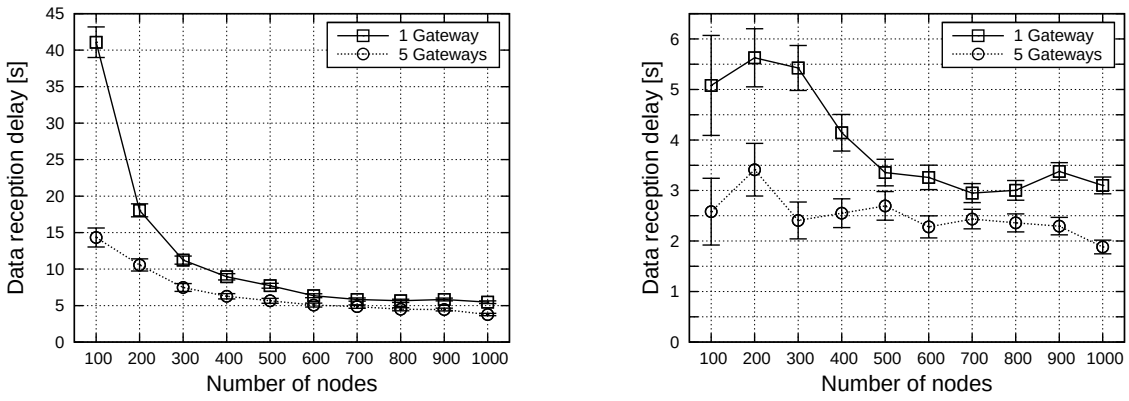
Besides the request process also the reply process needs to be analyzed to evaluate MDRP. The number of replies should not increase significantly more than linear with increasing node density, otherwise the protocol is not scalable. This characteristic is confirmed by the results shown in Fig. 3.40. The number of received replies is shown in Fig. 3.40(a). Irrespective of the number of nodes, on average six to ten replies are received per node. Throughout the simulation each node sent a number of reply messages (see Fig. 3.40(b)). It is not significantly influenced by the node density, however, an increased gateway density can reduce the number of replies. This result shows the good scalability of the protocol concept.



(a) Average number of received replies

(b) Average number of sent replies

Figure (3.40) MDRP reply process evaluation



(a) Average absolute reception delay for the full content

(b) Average delay between first and last data fragment

Figure (3.41) Evaluation results of reception delays caused by MDRP data distribution

Finally, the distribution delay required by MDRP shall be analyzed. The information with 50 kB is fully distributed in all network densities, however, the distribution time varies. In the slowest case (100 nodes) the distribution takes about 40 s, as shown in Fig. 3.41(a). This delay is reduced to a delay between 5 s to 10 s in scenarios with more nodes. Within this delay, the full data with all its fragments has been received. The delay elapsing between the first and the last fragment reception is shown in Fig. 3.41(b). Especially for dense scenarios the delay has a reasonable value between 2 s to 3.5 s. The scenarios with more gateways again show a lower delay proving the better protocol performance.

The presented results prove the scalability and high performance of the Mobile Data Request Protocol (MDRP). The request-based message dissemination is an efficient way to distribute information to MEs in a VN, since many nodes can be reached within few seconds and with few messages. Therefore, MDRP is a valuable concept completing broadcast-

based data diffusion protocols in VNs. It is very well suited to efficiently distribute any information which can be made addressable like status information, for example Certificate Status Information (CSI), or even regular software updates.

3.7 Conclusions

In this chapter the communication aspects of VNs have been discussed. After a detailed overview on the related work several of the existing challenges in respect to IVC have been addressed with different concepts. The presented approaches and simulation results can be used to design future IVC systems in a scalable and efficient way.

The first discussed aspects were the difficulties of message distribution in the distributed network setting of a VANET. Several strategies to improve scalability and performance of distribution schemes have been suggested, for example data aggregation, content priorities, and dissemination areas. The evaluation of the data aggregation approach proved its benefits over regular flooding-based approaches clearly. Aggregation is a very suitable concept for VANETs, to reduce the overall number of messages. This increases the scalability of message distribution especially for large scenarios with a high node density. The presented aggregation scheme is capable of modeling hazard areas using ellipses and disseminate their location throughout the network. Hence, it is the first aggregation scheme for VANETs capable of fully using the benefits of message aggregation.

Since the communication standard IEEE 802.11p is designated for the application in IVC scenarios its strengths and weaknesses should be known. Thus, the presented analytical and simulative evaluation of the standard helps to get an idea of the standard's capabilities and limitations. Especially the method to analytically calculate packet collision probabilities for different access classes is a very important result to be able to configure the standard according to the application scenario. The standard is a viable approach, on the other hand, its four priority queues are not sufficient to provide QoS in dense scenarios. The delays are in the range of seconds.

Prioritization of information or packets promises to be a valuable approach to improve the scalability of IVC in dense network scenarios. The presented concept for packet prioritization using content-utility as the priority index has proven to be very valuable to increase the overall as well as the individual information utility for nodes in a VANET. The extensive simulation results help to assess the benefits gained by using the suggested cross-layer approach in combination with different benefit calculation functions. The newly suggested system proved to be a successful filtering approach prioritizing the most important messages. The concept helps to use the scarce channel capacity in an optimized way, maximizing the overall utility of using IVC.

Besides the discussion of different content dissemination approaches the new topic of content-aware protocols in the context of VANETs has been introduced and the content request protocol MDRP has been presented. In this respect the integration of fixed gateway nodes has been evaluated. The results support the use of gateway nodes as well as content request protocol schemes in VNs. They are a valuable complement to active dissemination approaches. With MDRP an scalable protocol for request-based information distribution has been suggested. It is a candidate protocol to support content distribution, for example, CSI.

The different V2V communication protocols presented in this chapter and the respective simulative evaluations contribute to VANET research. They can be used to improve for example the scalability, performance, and the support of different node densities for future VANET realizations. However, for a real-life implementation security mechanisms are required in addition. Therefore, specific security solutions are required, which are adapted to the services and protocols in VNs. Some potential solutions closing this gap are discussed in the next chapter.

4

Security and Privacy Mechanisms for a Reliable and Trustworthy Vehicular Network System

Communication technology is the first and probably the most important requirement to realize a Vehicular Network (VN). However, to use communication technology in a commercialized setting the use of appropriate security mechanisms is equally important. The scenario setting of VNs and especially the mobile network part require security mechanisms which are specifically designed for the characteristics of the scenario. Thus, existing security solutions, known from e.g. the Internet can not be applied directly to VNs and Vehicular Ad Hoc Networks (VANETs) especially.

New mechanisms and strategies, as well as existing approaches with adjustments have to be found to fulfill the security demands of VANET scenarios in particular. The decentralized and distributed nature of VANETs suggests that a centralized security approach is not appropriate to provide the necessary trust. However, the already mentioned commercial application of these networks, as well as legal aspects of the Original Equipment Manufacturers (OEMs), call for a controllable trust architecture. Therefore, no fully decentralized and self-controlled solution can be applied in these mainly distributed scenarios. In order to successfully introduce Inter-Vehicle Communication (IVC) systems into future vehicles the specific security and privacy needs need to be addressed. Node authenticity, message integrity, and content reputation are the most important security features for IVC. They need to be integrated with the trust solution of the full VN setting.

In this chapter several security aspects of VNs are discussed and analyzed. The fundamental question of how to introduce trust in a VN, and how to use the same trust architecture in the VANET part of the system as well, will be discussed first. A semi-centralized trust architecture approach is suggested, which meets the requirements of the scenario and its commercial background. All other security mechanisms suggested in this thesis can be based on this trust architecture.

The security of practically centralized platform services needs to be realized differently than the security of distributed and session free services. A transparent and generic security

integration for platform service systems is presented, which has been developed in the context of the European Project Global System for Telematics (GST). This architecture can be used for various types of services and allows the use of privacy mechanisms as well.

In addition to the platform security mechanisms, the introduction of security mechanisms for IVC is discussed and analyzed. Adding security always results in overhead: Data overhead as well as additional processing requirements. This limits the scalability and performance of message distribution mechanisms like the ones discussed in Chap. 3. Hence, specifically adapted mechanisms need to be used to optimize the performance in distributed network settings. Moreover, different security features are important for session free IVC services, these are mainly authenticity and content validity. Therefore, a specifically adapted protocol called CoRS is suggested, realizing content reputation for message dissemination in VANETs.

Besides security another important requirement has to be met to realize VANET systems: The privacy of Mobile Entities (MEs). The use of pseudonyms is a promising approach to encounter this requirement, thus, it has been evaluated in the course of this thesis, suggesting how to manage pseudonyms in a VANET setting efficiently.

The security chapter is organized as follows. In Sec. 4.1 the related work on security and privacy is presented. The underlying trust architecture suggested for VNs is presented and evaluated in Sec. 4.2. The realization of security and privacy aspects for platform-based services in VNs is discussed in Sec. 4.3. In Sec. 4.4 the security of vehicle-to-vehicle (V2V) message exchanges is discussed and its efficiency is evaluated. To realize a content-based trust scheme a reputation mechanisms has been used in the new Content Reputation System (CoRS), which is introduced and evaluated in Sec. 4.5. In Sec. 4.6 an evaluation of privacy in VANETs based on pseudonyms is discussed. The chapter closes in Sec. 4.7 with conclusions.

4.1 Overview on Existing Security and Privacy Concepts

Security and Privacy have been an important part of system design since many years, however, they really have moved into the center of attention with the success of the Internet and the exponential increase in registered software vulnerabilities (refer to the Computer Emergency Response Team (CERT) statistics [CER08]). Especially the introduction of wireless communication technology increased the possibilities for attackers, since the communication medium can not provide any physical security. This leads to many different passive and active attack possibilities.

4.1.1 Trust Establishment, Communication Security, and Reputation

Trust Environments, Certificates, and PKI

Trust has been a research topic since many years. In [Deu58] an experimental evaluation of the trust concept has been presented, which can easily be transferred into the technical world. Trust and security in general has become a very important and prominent research topic in the context of communication networks. The importance of a trust management for network services has been discussed in [BFL96]. The authors point out that one unified trust

management is desired rather than trust on a service application level. Further they show several principles for such a system: A unified mechanism, a flexible system design, local control possibilities for trust mechanisms, and strict separation of mechanisms and policies. The suggested trust system concept is still very relevant for today's networks, especially for VNs.

In technical systems trust is typically introduced using public key cryptography and a Public Key Infrastructure (PKI), issuing certificates as trust associations. A selection of different PKIs and trust concepts has been analyzed in [Per99]. The paper gives an overview on the trust management problem. The authors advocate a convenient solution for the user which fulfills the most important requirements. Since no perfect trust management solution exists today, certain compromises have to be made. A mechanism needed for a reliable and trustworthy PKI is the revocation of trust associations as soon as a member has become unreliable. Hence, certificate revocation is a very crucial mechanism related to PKIs.

In the course of Mobile Ad Hoc Network (MANET) research the topic of trust management has become important as well. In [HBC01, CBH03b] the trust establishment in MANET environments has been discussed. A self-organized and distributed PKI was suggested to solve this challenge. The approach is very well adapted to the decentralized nature of MANETs, since it follows a similar concept like the Pretty Good Privacy (PGP) Web-of-Trust [Eck06, p. 375, p. 764]. However, this approach is easy to infiltrate and attack, which makes it unusable for a VN. Using a PKI as a trust basis has been suggested in other publications. In [CD03] a brief overview is given.

As stated above, PKI trust management is strongly connected to certificate revocation. The different technologies of certificate revocation have been discussed in many publications. In [Mic97, Mic96] the efficiency of different revocation approaches has been evaluated and a new approach enhancing both functionality and efficiency has been introduced. The so-called Certificate Revocation System (CRS) can provide validation as well as revocation information for certificates in a very efficient way. This concept has been extended to the NOVOMODO concept in [Mic02], which is a suited revocation and validation approach for decentralized systems such as VANETs. It will be discussed in Sec. 4.2.2 in more detail. An overview on several certificate revocation system has been presented in [Woh00] giving an overview on different approaches. But so far, no evaluation on revocation information distribution in wireless environments has been presented.

A specifically designed revocation scheme for MANET environments has been specified in [CD03]. The main challenges in respect to certificates and revocation in MANETs given in the paper are the certificate issuing process, certificate management and validation, and revocation of certificates. The suggested scheme uses a reputation-based revocation approach. Nodes accuse each other of misbehavior and generate an accusation profile. This is used for the revocation decision. This distributed approach is especially applicable in small MANETs which are fully distributed, therefore, it is not suited for a VN with commercial background.

The use of certificates to handle trust in VNs has been suggested in several papers. In [GM02] several application scenarios for certificates in the vehicular context have been suggested: fleet management and vehicle localization based on Global Positioning System (GPS) and digital certificates, payment systems for service applications triggered by the vehicle, and infotainment services. This vision is still very relevant and applicable for today's

VN concepts. In a second rather elementary paper on security issues in VANETs the use of certificates and a PKI has been promoted [ZMTV02]. The authors argue that a PKI needs to be adapted to the unique characteristics of VNs.

The high relevance of trust in future VNs has also been emphasized in [RH05]. Since especially the safety services require a high reliability and liability, trust is the key mechanism for the success of VNs. In addition, misbehavior is very likely in VN scenarios due to the many independent network members. Trust can only be realized if the different involved authorities cooperate [RH05]. The authors further argue that important building blocks are digital certificates and tamper-proof devices. Several certificate handling mechanisms are discussed and the use of a PKI in the vehicular context is considered as deployable, unfortunately without supporting these arguments by simulation results.

The use of a PKI with certificates and infrastructure support for VANET scenarios is also suggested in [RPH06]. The revocation of certificates should be done based on lists in combination with a specific handshake protocol between the vehicle and the infrastructure components. This seems to be a somewhat questionable approach, especially since the authors do not present simulation results to support the viability of the idea.

The trust establishment and a respective architecture integration for IVC systems is presented in [Ger07]. The author suggested to tag information with a trust value (trust interval of [0,1]) based on a sociological trust model. The different trust definitions (for example system trust and situation trust) mediate between entities and applications. This approach is very similar to a reputation system and not capable of providing a liable trust concept for all kinds of VN applications.

Overall can be summarized that security in future VNs will most likely be realized with a PKI and certificates as the trust basis as suggested in many publications. However, an evaluation of the performance has not been done so far. Especially for the revocation of certificates such a performance evaluation for distributed environments is necessary, before using the technology. Some basic evaluations have been presented in [SE04, Sch05b]. In this thesis suggestions and evaluations for different revocation mechanisms are presented. The revocation can be considered as the performance limiting part of a PKI concept in VANET scenarios, as long as certificates do not have to be renewed using the IVC.

Security for Communication Mechanisms

The main feature of a VN and especially a VANET is IVC. Besides the performance requirements on the communication, security is a very important feature which is a prerequisite to realize VNs. Several concepts have already been proposed for this challenge, the early ones in the context of MANET scenarios.

One of the first papers on security in MANET scenarios is [ZH99]. The main threats and security challenges are discussed and some early approaches presented. The concept of threshold cryptography was considered as a new mechanism especially well suited for the characteristics of distributed networks. These distributed security mechanisms are also seen as an important building block for security in MANET environments in [SH02]. A survey on distributed security and threshold cryptography is given in the paper. Nevertheless, these concepts are not generally applicable in VANET scenarios since liability and centralized

control can not be provided. However, these characteristics are desired features in future VN realizations.

A detailed concept for security in MANETs has been presented in [Kar03]. Besides the overview on threats and attacks, a security architecture is proposed, including secure routing, intrusion detection, and a trust basis using a PKI. The general ideas and threats presented in the thesis can be applied to a VANET scenario. However, especially the secure routing is not applicable since routing will practically not be used in a VN. A security concept for MANETs has also been presented in [Sch05b]. The presented security framework provides secure communication for distributed networks using certificates and a PKI. Like in most MANET-related publications point-to-point (P2P) communication is considered rather than the more likely broadcast communication. The presented evaluation results do not include a performance evaluation of the PKI, only the suggested P2P communication protocol has been evaluated using network simulations. The suggested security framework (LKN-ASF) targets secured P2P communication in MANETs using routing. Hence, this concept is not suitable for the requirements in VANET environments.

An introduction of security and privacy implications for VANET scenarios has been presented in [HCL04]. Besides the importance of privacy integration the authors stress the necessity for a reliable and trustworthy positioning of vehicles for IVC services. Two mechanisms are suggested, tamper-proof GPS devices or a triangulation using known base-stations in the vicinity of the respective vehicle. Again a PKI is suggested to be used as a trust basis for the VANET.

A detailed overview on the requirements and challenges of the security integration in VNs has been given in [PP05]. The specific scenario requirements of VNs call for very reliable security mechanisms. Besides an intensive discussion of potential adversaries and attacks the authors stress the challenge of key distribution for a used PKI and that a trade-off between full authentication and node privacy has to be found. In addition the concept paper lists possible security primitives as building blocks for secured V2V applications. Authenticated location information, anonymization services for temporal identities, and secure data aggregation mechanisms are some examples.

The security of IVC has been discussed in [RPH06] with great detail. Based on the vulnerabilities (jamming, forgery, impersonation, privacy violations) and challenges (liability vs. privacy, network scalability, robust security) an architecture design concept is presented. Based on a PKI trust basis each vehicle is equipped with an event recorder and a tamper-proof device. The use of Elliptic Curve Cryptography (ECC) is suggested to provide the public key cryptography since it is more efficient than conventional techniques. As the main open issues in the field secure positioning, Denial of Service (DoS) resilience, and data verification are pointed out. The data verification issue is tackled in this thesis with a content reputation protocol in Sec. 4.5.

Security systems can be evaluated with attack trees. In [ABD+06] an attack model for IVC systems based on attack trees has been suggested. It has been used to refine the communication concepts of the Network on Wheels (NoW) research project. Nevertheless, the results highlight new challenges for V2V systems. The security assessment showed that a trusted platform providing trusted communication services, privacy features, and tamper evidence are very important features for an IVC system.

Reputation Mechanisms in Mobile Networks

To realize a data verification mechanism for IVC the use of reputation mechanisms is a promising approach. Reputation mechanisms have been used in many different contexts of communication systems, wireless as well as fixed networks or service architectures. It has to be strictly differentiated between *node reputation* and *content reputation*. While node reputation is only helpful as long as nodes interact many times, content reputation is beneficial for all content receivers right away. For VANET scenarios content reputation is more important, since nodes will not necessarily re-interact multiple times. On the other hand, mainly node reputation schemes have been suggested so far. Use cases for reputation systems are manifold. In the context of communication networks they are mainly used for opinion management in various settings: Internet-based commercial platforms such as eBay, service selection, routing decision making in mobile networks, and others. A review of reputation usage for commercial Internet trading was presented in [RKZF00].

To support the design of node reputation mechanisms a reference model has been suggested in [Rei05]. The author used the Universal Markup Language (UML) to define the model and include different points of view. However, the model is primarily targeted at commercial systems with suppliers and customers. A similar contribution is presented in [CHD05], where an ontology description of reputation systems and their application areas is introduced.

With the upcoming of research for MANET protocols the concept of reputation has been discovered to motivate node collaboration. In [MGLB00] a first approach using behavior observations to improve routing decisions is presented. In a way this is the starting point for the use of reputation mechanisms in MANETs. This concept has been extended and improved by Michiardi et al., who presented a reputation system called CORE in [MM02], which tackles the problem of misbehavior and selfishness. Based on the collaboration rate that nodes observe of their neighbors, a reputation table is filled. Several different reputation values are measured and combined before making, for example, a routing decision. CORE has been applied and evaluated with the Dynamic Source Routing (DSR) protocol in the paper. A very similar approach has been suggested in [BB02b], the so-called CONFIDANT protocol.

In [BB03, BLB05] a reputation system especially applicable to the CONFIDANT protocol was introduced. It uses reputation for nodes, which is generated mainly by observation of the neighborhood. In addition, a filter for second-hand information based on a modified Bayesian model is suggested. It helps to improve the reputation information by including observations of other nodes. Again the reputation concept is applied to MANET routing. A very similar approach is presented in [RMVS05].

An elaborate concept for node reputation is presented in [LI04]. A detailed description of reputation characteristics and important parameters is introduced in the paper. The authors stress the subjective nature of reputation concepts. Further, they present a method to differentiate between truth-telling and lying nodes by using the new concept of *recommendation reputation*. It helps to differentiate truth-telling and lying agents, thus, increasing the degree of reliability.

The idea to use reputation to improve content distribution has been used to realize the Reputation, Opinion, Credibility and Quality (ROCQ) protocol [GC05]. A global reputation

on peers, generated on first-hand opinions, is used to reduce corrupted content transfers. A distributed database management is used to provide global reputation values for all peers. After a data transaction, peers rate each other base on the success of the transaction.

The combination of threshold cryptography and reputation mechanisms is rarely used in the literature. However, in [Nam05] the use of threshold cryptography to realize a reputation system for trustworthy service selection in cooperative Web-service environments was suggested. In this concept the reputation is connected to services and their providers, hence, a reputation for nodes and not for content has been presented.

A reputation management scheme for VANETs has been presented in [PJFY06]. It uses node behavior as well as content accuracy to obtain reputation values for nodes. The scheme requires persistent node identities to be able to handle the reputation values. The reputation is used as a basis for an epidemic data exchange and to encourage cooperative behavior between nodes.

Another concept proposing the use of a reputation mechanism for VANETs is presented in [DFM05]. The authors suggest a protocol called Vehicle Ad-Hoc Network Reputation System (VARS). It is a reputation system not based on node behavior but on opinion about distributed content. Hence, VARS is not a reputation system for nodes but for message content, similar to the concept presented in Sec. 4.5. In VARS forwarding nodes form an opinion on the content of a given message. This opinion is attached to the message before forwarding it to other nodes. Therefore, receivers can evaluate the opinion of other receivers and use it as a basis for their own decision making. Only important messages will be processed further. The VARS concept has no sufficient protection for the attached opinion values, which is provided by the protocol specified in Sec. 4.5.

4.1.2 Privacy Mechanisms

Besides trust establishment and communication security, privacy has been identified as a very important feature for IVC systems. To be able to assess the level of anonymity and privacy of VANET nodes a measurement approach for anonymity is required. Several different approaches using anonymity-sets or information theory exist in the literature [DSCP02, SD02, Hua06]. They can be used in the context of VNs to estimate the level of node anonymity and to optimize privacy protecting protocols.

As service applications started to become location aware the challenge of implementing privacy protection into these applications was brought up. The location privacy challenge in pervasive computation environments has been studied in [BS03]. The authors point out that controlling access to information is the key issue when implementing privacy protection. Frequently changing pseudonyms are used for privacy protection in their approach. Further, they suggest entropy measures or anonymity sets, which are the number of nodes in the vicinity at the same time, to be used as anonymity metrics. The information theoretic metrics proved to be most effective, hence, they should be used for future approaches. An entropy-based anonymity metric based on [SD02] has been suggested to be applied for VANET evaluations in [HCL04], motivating for privacy awareness in respect to VNs. The survey and position paper [PKHK06] gives an overview on different privacy and identity management schemes usable in VANET scenarios. The authors argue that architectural decisions have to be made before privacy enhancing mechanisms can be defined and selected.

A first system concept providing privacy has been outlined in [Döt05b]. The author presents the challenges connected to node anonymity in VANET scenarios and stresses the conflict between the need for node authentication and node privacy. Simply by combining practically innocent data pieces, the privacy of a respective node can be breached. The suggested architecture approach uses a trusted third party to hide the mapping between real identity and the connected pseudonyms. This idea can be realized using a PKI and certificates, hence it is a valid approach for VANETs. In addition the author points out that pseudonym changes of a node can be observed, therefore, unveiling that they belong to the same node. Thus, the question arises how pseudonym changes should be handled to remain hidden. This issue is discussed in Sec. 4.6.

To reduce the correlation possibility between addresses or pseudonyms of a node the authors of [HMYS05] suggest to use a random silent period. Their simulation results prove the assumption and show that a silent period can be used to increase location privacy for nodes in mobile scenarios. This result has been used in [SHL+05] to define the VANET location privacy system CARAVAN. In CARAVAN nodes are grouped, which helps to reduce the necessity for message forwarding of each individual member of the group. Therefore, the silent period can be increased, which helps to increase the location privacy. A more advanced version of CARAVAN called AMOEBA has been presented in [SLHP07].

Protecting the nodes' privacy by using changing pseudonyms is a common idea [Kar03, HCL04, SKL+06]. In [GG07] this approach has been evaluated in a VANET scenario. The authors use the context mix model, which helps to identify the ideal point in time to change a pseudonym. For example the number of neighbors is a mix criteria. As soon as enough neighbors are present, exceeding a specified threshold, the pseudonym is changed. The usage of the context mix model is supported by the simulation results presented in the paper.

The influence of node mobility on its privacy is discussed and evaluated in Sec. 4.6. The identified parameters relevant for the pseudonym change can be used as input values for a context mix model, using the presented results as a basis for the required parametrization of the pseudonym change protocol.

4.1.3 Cryptographic Building Blocks

A large selection of cryptographic mechanisms exist in the literature. Only a small selection shall be introduced and referenced in the course of this thesis. No new cryptography mechanisms have been developed, however, existing mechanisms were used and combined to solve specific issues in the context of VNs. Without cryptography the concepts presented in this chapter can not be realized. However, since the thesis does not focus on designing new cryptographic mechanisms, an introduction in this broad field is out of scope for this chapter. Very detailed introductions can be found in the literature, for example, [DH76, RSA78, DH79, MvOV96, Sch96, Eck06].

The main aspects discussed and used in this thesis are asymmetric cryptography mechanisms, for instance, Rivest, Shamir, & Adleman Public Key Encryption (RSA) or ECC, and their application for trust establishment. In addition, symmetric encryption, hash functions, and threshold cryptography are important building blocks used in the following.

4.2 Realizing Security and Privacy in Vehicular Networks with a Semi-Centralized PKI Approach

Before discussing the technical concepts of security, the fundamental social background of the topic security and trust is reviewed. The notion of trust has been studied in many different contexts in the literature. As a starting point the definition of the term *trust* given in the Oxford English Dictionary [SW89] shall be used:

> "Confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement."

Therefore, trust has to do with confidence and reliability. It can be related to persons, things, or information. Moreover, trust is a predominantly social construct. A detailed discussion on trust and its social importance has been given by Deutsch in [Deu58]. He connects the term trust with the terms *expectation* and *predictability*. Therefore, trust has to do with the expected occurrence of a specific event or the prediction of a certain behavior in a distinct situation. In addition, the author argues that trust is in close coherence with the concept of risk. Whenever there is a risk, trust can help to cope with or even mitigate it.

In [KC98] the notion of trust has been examined in the context of electronic commerce, hence, trust in a technical environment. The authors define trust as a belief in system characteristics and its dependability. The use of a technical system is always an act of balancing risk, trust, and usability.

However, since trust is a predominantly social construct its introduction to technical systems is complicated. In addition, models to map trust between machines and humans are required, since the perception of trust differs significantly between machines and humans. By reason that trust for machines comes down to a mathematical definition, it can be rated unambiguously. This, however, is not the case for human beings. Therefore, the technical representation of trust is only a method to map the social construct *trust* into the technical environment and back. This has to be kept in mind when talking about trust and security in the context of technical systems such as VNs. The technical trust can not be equated with the social trust, however, it should optimally be an adequate mapping.

4.2.1 Mapping of Trust in Technical Systems

The mapping of trust from the social level into the technical system level includes several steps. An appropriate mapping mechanism is needed to handle trust within the system. This is usually done by using asymmetric cryptography, which provides a sufficient degree of security and reliability. Thus, the trust represented by the cryptographic mechanisms can not be altered by any unauthorized entity in the system. Therefore, the mapping mechanism is appropriate, since a receiver can rely on its trust information value.

Besides this mapping mechanism usually a contractual basis is needed, specifying how the mapping is done, under which circumstances it is done, and which entity is executing the mapping. Thus, an entity, usually a corporation which is considered reliable and trustworthy, performs the mapping or the certification of trust. This entity is the so-called Certificate Authority (CA). Each CA has certain policies which are followed to issue a certificate. The

certification process can for example be used to connect a cryptographic key to an identity and make this information available to technical systems.

Last but not least a broad acceptance of the respective CAs on a social level is required. Otherwise no user of the technical system will rely on the service provided by the CAs, therefore, the social trust in the reliability and trustworthiness of the mapping entity is not existing.

In [LN98] the question of how to certify trust has been discussed. The authors present the term *security policy* which describes the contractual basis mentioned above. A security policy is a definition of well-accepted security-related practices. It is used as the basis for decisions on the trustworthiness of organizations.

4.2.2 PKI and Revocation – Definitions and Realization Approaches

The pure mapping of trust is usually not sufficient. A database has to exist to manage the mappings and provide information on them to any third party. A commonly used trust database concept is the so-called Public Key Infrastructure (PKI). The concept of a PKI has been first introduced by Merkle in the form of public key protocols [Mer79, Mer80]. Based on the initial idea for public key cryptography by Diffie and Hellman in [DH76] a new mechanism of trust handling and managing was possible. In a PKI a trusted third party, usually the CA, manages trust using public key cryptography. Hence, each member of the PKI holds a key pair with a public key (K_{pub}) and a corresponding private key (K_{priv}). The CA issues a certificate which binds the identity of the member to its respective key pair.

If the key pair is used in a communication session, the peer can use the certificate to verify to whom the key pair has originally been issued. Assuming that keys are handled in a secure way, thus, they will not be handed to anyone and can not be stolen easily, this verification allows to identify the originator of the signed messages. This example points out that the use of a PKI as such does not necessarily make a system secure or trustworthy. Many influences and system properties need to be included into the trust consideration. Nevertheless, the PKI helps to support the mapping of trust into the technical world and represents it sufficiently if certain rules are followed.

In [ES00] a risk overview for PKIs has been presented. The authors present 10 risks of a PKI and try to motivate a sensible usage of the PKI concept. The most crucial issues related to a PKI are:

- Which entities are using certified keys and how secure are they stored and handled?
- What is used as the certified identity and is this an explicit match or are dupes possible?
- How is the entity identified and is this done by the CA directly or is a Registration Authority (RA) used?
- What is the degree of system security?

Depending on the answers to these questions, the level of trust for a PKI can be assessed. However, events that breach the security of a system can occur, even if the trust level is considered as high. Therefore, mechanisms have to exist that help to restore or maintain the trust level.

The trust provided by a PKI can also be described as a circle-of-trust. All members of this circle can interact without any third party, since the trust relations have already been set up using the certificates of the PKI, which have been issued by the CA.

The Need for Revocation in PKI Trust Environments

In most PKIs the certificates are issued with an expiration date. Therefore, a certificate is only valid within a given period, after that it is no longer considered as trustworthy and will not be used anymore. The public key cryptography usually used for PKI certificates uses limited key sizes to keep the cryptographic calculations as scalable and efficient as possible. Hence, their level of security is limited (see [Gir07] for current key size recommendations). This limitation needs to be reflected in the validity period of the certificates. In addition, the expiration date prevents certificates from potentially being valid forever, which would be considered as insecure.

Even though an expiration date exists, it must be possible to declare a previously valid certificate as invalid. This might be required due to the identity change of the certificate owner, the revelation of a private key, or simply the loss of a private key. The process of invalidating a certificate and the corresponding key pair is known as *Revocation* [NN00, Woh00]. The revocation process includes two main steps: The identification of invalid certificates and the communication of revocation decisions, also known as Certificate Status Information (CSI), to all active members of the PKI. Many different revocation mechanisms and strategies have been suggested in the literature. Nevertheless, not all of them are suitable for the specific characteristics of VNs. The three main revocation mechanisms are:

Certificate Revocation List (CRL): The CRL is probably one of the oldest revocation mechanisms presented in [Mic95, HFPS99]. It is a list containing all certificates or at least their identification numbers. The list has an issuing date and can contain the reason for each individual revocation decision. In addition, the date for the next version of the CRL is given. A variant of CRLs is the so called delta-CRL. In the delta-CRL concept a full CRL is only distributed occasionally. Between the publication dates differential lists are published which contain only the revoked certificated since the last full CRL has been issued. This variant mainly reduces the size of CRLs, hence, makes them more usable in bandwidth limited scenarios.

Certificate Revocation System (CRS): The CRS has been introduced by Micali in [Mic95] and improved in [Mic96, Mic97]. The CRS uses both positive and negative revocation information lists, in contrast to the CRL concept which only uses negative information. A constant update interval is used by CRS to keep the validation (positive) and revocation (negative) information up-to-date. A user can ask for the validity or revocation of certificates and CRS will return a respective list which is signed by the CA. The CSI is transported in form of a single value, which can be verified by anyone using a cryptographic hash function ($\mathcal{H}$). The concept of hash functions has been introduced in [DH76]. A hash function has two important properties, required for the CRS: Compression of the argument with arbitrary length to a fixed length, it can be calculated in one-way only (the inverse function is practically impossible to calculate), and no collisions for two arguments leading to the same hashing result can

be found [Pre99]. Within each certificate of a CRS architecture two values are contained: A validation target (t_V) and a revocation target (t_R), where $t_V = t_{V,x} = \mathcal{H}^x(t_{V,0})$ and $t_R = \mathcal{H}(t_{R,0})$. Hence, the validation target is generated by hashing the initial value x times, where x resembles the maximum validity in discrete update periods. The revocation target (t_R) is computed by a single hashing operation. Every CSI period the CA publishes either a validation or a revocation value. For the validation values the CA starts with $t_{V,x-1}$, iterating through the hash chain backwards until $t_{R,0}$ is reached. If the CA publishes $t_{R,0}$, rather than a validation value, the certificate and its key pair is revoked. The publication of CSI can be done using a small ticket, containing only the CSI value and the certificate ID. The CRS variant using CSI tickets is also known as NOVOMODO and has been presented in [Mic02].

Certificate Revocation Tree (CRT): To reduce the complexity and flat hierarchy of CRLs Kocher suggested the CRTs in [Koc98]. The concept uses the idea of hash trees first introduced in [Mer79]. The tree contains both validation and revocation information and provides an efficient way to transport both types of CSI in one data structure. The certificate information is contained in the leaves of the tree by grouping valid certificates into groups, separated by the invalid certificates. This structure allows to construct a tree using a hash function. The verification of the distributed tree only requires one signature verification and several hash function evaluations, making it an efficient and scalable revocation mechanism.

Many other revocation approaches have been suggested in the last years [NN00, Woh00, CD03, Zhe03, EGR04]. Practically, they all are based on one of the three main approaches presented above and are adapted to specific characteristics of the application scenario. This adaptation is an important step, since only an efficient revocation mechanism will fully conserve the security of the system.

4.2.3 PKI Performance Issues

Several performance studies of PKIs have been made [Zhe03, Per99, EGR04, KAN99, Mic02, Mic97], however, none of these evaluations has considered the typical characteristics of a VANET scenario. A first concept using a PKI and certificates to secure communication in mobile environments, especially MANETs, has been introduced and evaluated in [SE04, Eic04], These preliminary results on the performance of certificate-based protocols in MANET scenarios were the basic motivation to evaluate the use of a PKI and the connected revocation challenges in the context of VANETs. The main question to answer is: Can the distribution of CSI in a VANET scenario be realized with reasonable delays and high reliability? A selection of the following results have previously been published in [EMR05] and [SEMR06]. Besides the CSI distribution challenge a second challenge is introduced by using certificates and a PKI: The exchange of certificates and the related data overhead. This issue will be further addressed in Sec. 4.4 since it mainly affects the IVC.

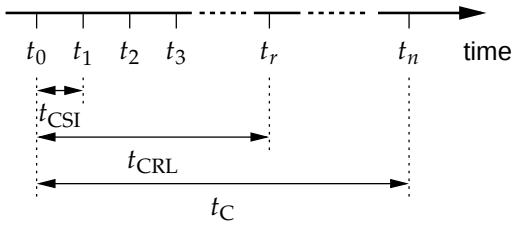


Figure (4.1) PKI intervals and time periods visualized with a timeline

Analytical Estimation of PKI Scalability

To be able to estimate the performance of a PKI and decide if it is adequate for VANETs, an example scenario can be used. The example scenario is characterized by the following values: The number of certified participants N_n, the estimated average rate of certificate revocation per year R_r and the certificate status information period (t_CSI). Moreover, the certificate validity period (t_C) and the certificate revocation list distribution interval (t_CRL) are required.

The numbers for a national PKI in Germany can be estimated as follows. In the year 2006 $N_\mathrm{n} = 55$ million vehicles had been registered, with a rate of 3.5 million vehicles being newly registered in that year. If a revocation rate of $R_\mathrm{r} = 10\%$ is assumed, 5.5 million revocation events will occur every year, roughly 15000 every day. In addition, the following values are used: $t_\mathrm{CSI} = 1\,\mathrm{day}$, $t_\mathrm{CRL} = 30\,\mathrm{days}$, and $t_\mathrm{C} = 365\,\mathrm{days}$. The relations of the different periods can be seen in Fig. 4.1.

The CSI for a PKI is updated and distributed once every t_CSI. In a conventional CRL scenario a full revocation list is sent, however, in a delta-CRL scenario a full list is only sent at the beginning of each t_CRL. Hence, in the given example this happens once every month. In between only differential lists are distributed. If the average revocation rate of the PKI is known the size of a delta-CRL can be determined using Eqn. (4.3). The values for $\mathrm{CRL}_\mathrm{base} = 271\,\mathrm{B}$ and $\mathrm{CRL}_\mathrm{C} = 48\,\mathrm{B}$ have been assumed [EMR05].

$$N_\mathrm{year} = N_\mathrm{n} \cdot R_\mathrm{r} \tag{4.1}$$

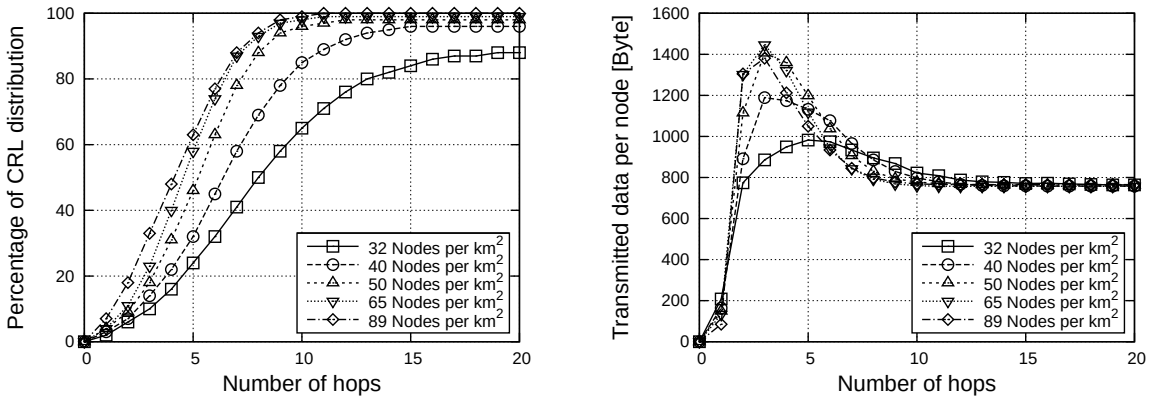
$$N_\mathrm{day} = \frac{N_\mathrm{year}}{365} \tag{4.2}$$

$$\Delta\mathrm{CRL}_\mathrm{size}(t) = \mathrm{CRL}_\mathrm{base} + (t \bmod (t_\mathrm{CRL} + 1)) \cdot N_\mathrm{day} \cdot \mathrm{CRL}_\mathrm{C} \tag{4.3}$$

Using the example values above, a delta-CRL will grow roughly 750 kB per t_CSI. Keeping in mind that a typical Message Protocol Data Unit (MPDU) is roughly 2500 B large, this increase is very significant. Therefore, a delta-CRL can only be distributed in fragments, rather than one single message.

Evaluation of Certificate Status Information Distribution in VANETs

To evaluate the performance of CSI distribution in VANETs, several simulations with different settings were used. In a first step a very general simulation setting was applied to get a



(a) CRL distribution effectivity

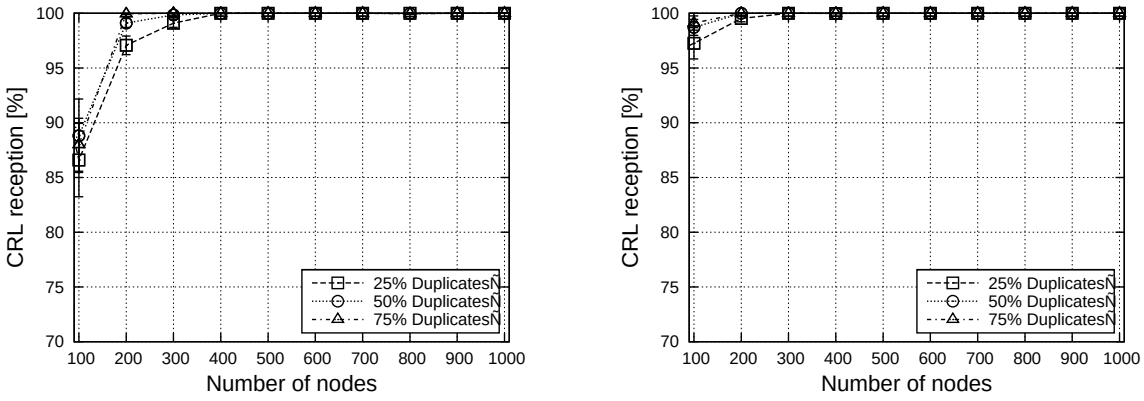
(b) Average transmitted CRL data per node

Figure (4.2) CRL distribution on a per-connection basis for different node densities [EMR05]

more general idea of the distribution possibilities in ad hoc communication environments. A squared simulation area of 2000 m × 2000 m with a fully random node placement was used. The nodes communicated randomly with other nodes, relying on the Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 Wireless Local Area Network (WLAN) communication standard at a data rate of 11 Mbit/s. One gateway node, which was placed in the center of the scenario, distributed CSI. After an initial distribution of the latest CRL the nodes communicated and exchanged the CSI on request, since a communication is only allowed as long as both peers have the latest CRL available.

This scenario led to the results in Fig. 4.2. The higher the node density the faster is the distribution of CSI. Moreover, the dissemination percentage increases with the node density (see Fig. 4.2(a)). After around ten hops the network is fully flooded with the CSI, mainly depending on the node density. In this scenario a CRL of 750 B was used, which led to the CSI data transmission profile shown in Fig. 4.2(b). Significantly more data per node than the size of the CRL has to be sent to reach the nodes for the first five to ten hops. This effect is caused by the forwarding of the CRL by the nodes. After ten hops only one packet needs to be transmitted, mainly since on average one additional node receives the CSI.

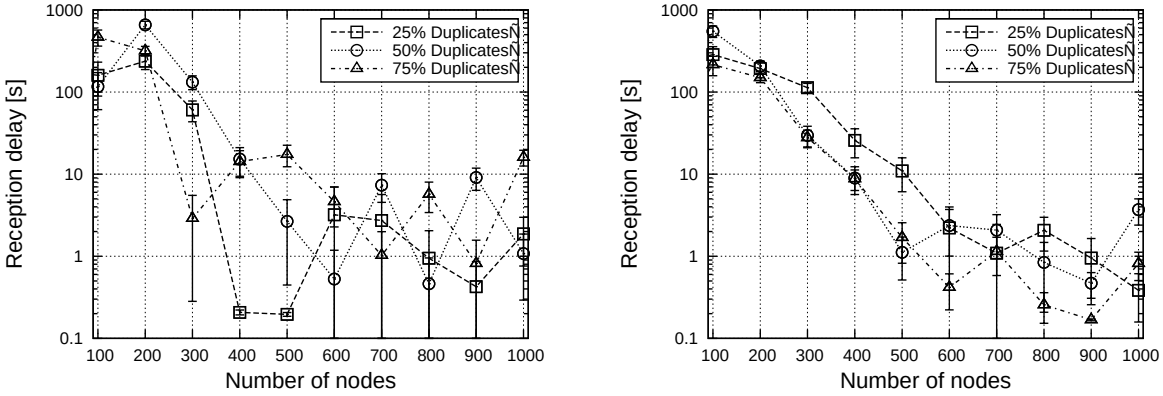
To verify these results and broaden the view on CSI dissemination a more elaborate scenario was set up. Using the OMNeT++ simulator with the simulation models introduced in App. A, several additional settings were simulated. One or four gateways, organized as shown in Fig. 3.33, were used to disseminate CRLs with the sizes of 10 kB (6 fragments) and 50 kB (30 fragments). The fragmentation process of the CRL cuts the data into individually usable and verifiable fragments. Therefore, even if not all fragments of a CRL have been received, the individual fragments can still be used. Each node receiving a new CRL fragment forwards it to neighboring nodes with a certain probability. The probabilities of 25%, 50%, and 75% have been used. The gateway nodes disseminated the CRL fragments once every 600 s for a simulation duration of 3600 s.



(a) One gateway in the scenario

(b) Four gateways in the scenario

Figure (4.3) Reception percentage for CRLs with a size of 10 kB



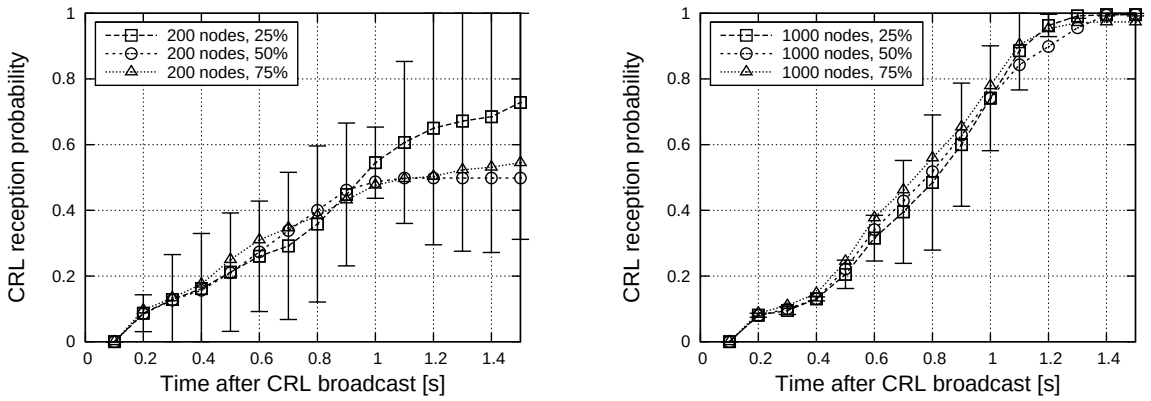
(a) One gateway in the scenario

(b) Four gateways in the scenario

Figure (4.4) Reception delay between the first and last fragment (CRL size 10 kB)

The reception percentage for CRLs with a size of 10 kB is shown in Fig. 4.3. In scenarios with 400 or more nodes the CRL is fully distributed to all nodes. Increasing the forwarding probability improves the distribution slightly. However, a greater improvement is achieved by using more gateways (see Fig. 4.3(b)).

Besides the overall dissemination percentage the delay of the dissemination is an important performance parameter. Since the CRL is fragmented some time passes between the reception of the first fragment until all fragments have been received. This reception delay is shown in Fig. 4.4. The delay until the full CRL is received can be significantly larger than 10 s in the scenarios up to 300 nodes, which is mainly due to the loosely connected network. With increasing node density this delay decreases below 10 s or even below 5 s, when using a higher gateway density. These results are average results. Many nodes receive the CRL in much shorter time. This can be seen in Fig. 4.5.



(a) 200 nodes

(b) 1000 nodes

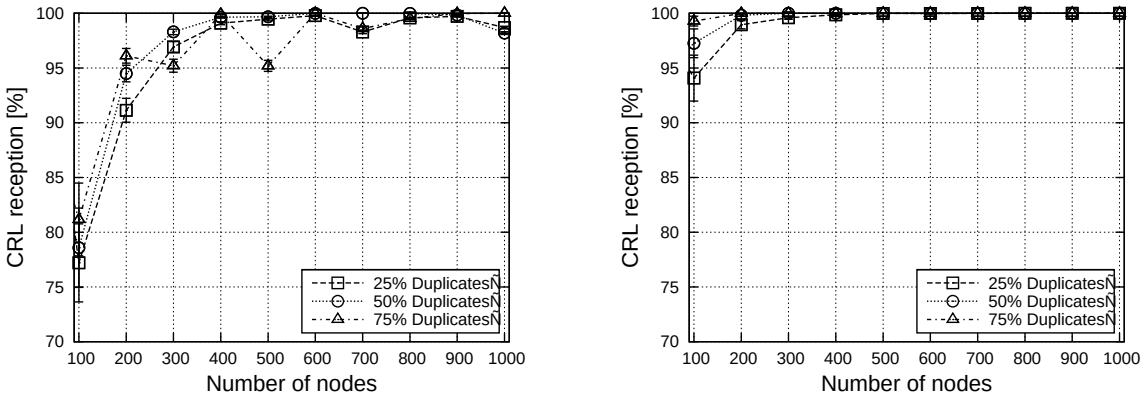
Figure (4.5) Reception delay after the first CRL dissemination (1 gateway, CRL size 10 kB)

As soon as the gateway sends CRL fragments, the receiving nodes forward the packets to other nodes, supporting the dissemination to nodes several hops away. Fig. 4.5 shows the global reception probability for a full CRL during 1.5 s after the dissemination started. In a low density scenario with 200 nodes (see Fig. 4.5(a)) the reception probability varies a lot due to the loosely connected network. This results in large confidence intervals. However, on average up to 75% of the nodes receive the full CRL within 1.5 s. In scenarios with higher node densities the variability is reduced significantly and the reception probability reaches almost 100%. This is shown in Fig. 4.5(b).

To evaluate the performance for larger CRLs the same simulations have been done using a CRL with a size of 50 kB. The reception percentage is shown in Fig. 4.6. The distribution is not as successful as for the smaller CRL, however, for node densities of 400 nodes and above the distribution degree almost reaches 100% in all cases. Especially when using more than one gateway node the distribution of larger CRLs works almost with the same performance (see Fig. 4.6(b)).

But a performance degradation can be seen when looking at the distribution speed for the full CRL. In Fig. 4.7 the reception delay between the first fragment reception until all fragments have been received is depicted. Notice the change in scale of reception delays as compared to Fig. 4.4. Especially when using one gateway only, the delay increases significantly compared to the smaller CRL (see Fig. 4.7(a)). Increasing the number of gateways helps to reduce this delay as shown in Fig. 4.7(b).

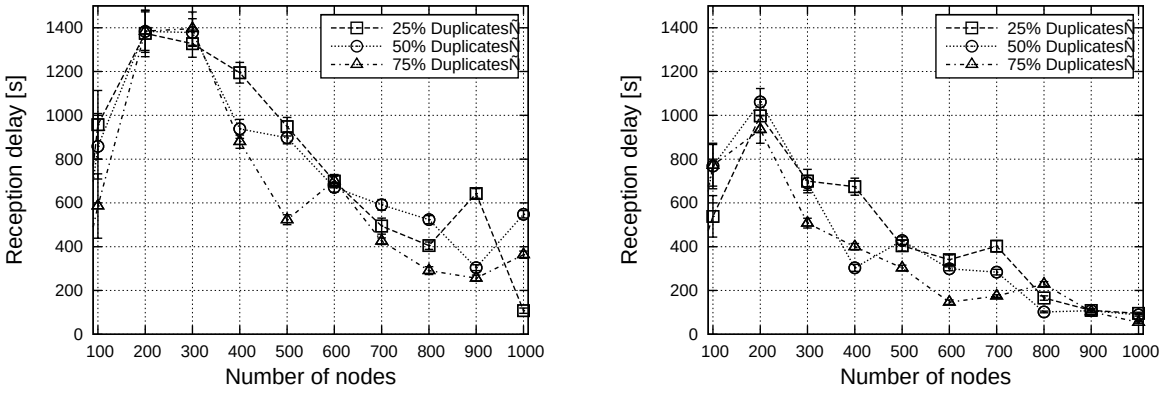
The presented results help to estimate the performance of CSI distribution in a VANET. Based on the results a distribution even of larger CRLs appears to be feasible using V2V broadcasts. The distribution quality will be further increased as soon as some vehicles receive CSI information for example using a platform-based service. Moreover, the distribution mechanism can be complemented with a request-based data exchange protocol such as the Mobile Data Request Protocol (MDRP). This will compensate the packet loss caused by collisions and reduce the reception delay significantly. Overall the presented results show that the use of a PKI in a distributed network setting like a VANET is feasible and especially



(a) One gateway in the scenario

(b) Four gateways in the scenario

Figure (4.6) Reception percentage for CRLs with a size of 50 kB



(a) One gateway in the scenario

(b) Four gateways in the scenario

Figure (4.7) Reception delay between the first and last fragment (CRL size 50 kB)

the crucial CSI distribution process can be realized with reasonable protocol and system complexity. This is an important result supporting the realization of a trust basis using a PKI for VNs.

Besides the usage of a CRL-based revocation the CRS suggested by Micali promises to be a very effective mechanism to exchange CSI with low data overhead. The simulations on this so-called NOVOMODO ticket-based revocation presented in [EMR05] show that the additionally exchanged ticket data does not influence the V2V communication significantly, assuming a ticket size of 138 B. Organizational aspects for the PKI use in VNs are presented in Sec. 5.2.3.

4.2.4 Semi-Centralized PKI Trust Architecture for Vehicular Networks

Using a PKI is a well suited approach to provide a trust environment in VNs. The concept supports basic security features like authentication as well as commercial applications

requiring complex invoicing procedures. In addition, the operator has the full control over the membership and subscriptions of distinct MEs. PKI certificates can be used as a basis for all required security features, in Backend-based service operations as well as distributed VANET interactions. In fact, the trust basis is provided without continuous access to the PKI servers. An important requirement for the application of a PKI-based trust architecture in VN settings is the availability of at least one revocation scheme suitable for distributed and decentralized VANET environments. The performance evaluation on revocation techniques presented above supports this availability.

For VNs a semi-centralized trust architecture should be used. It is based on a centralized PKI as the trust basis, in combination with distributed revocation and validation mechanisms for certificates. All certificates and CSI can be used and distributed throughout the VANET without any Backend server interaction, while keeping the same level of security. Hence, only the trust providing PKI is centralized, whereas use and trust preservation based on revocation can be done in both centralized and distributed environments. This setup fulfills the requirements for Backend and VANET of a VN.

4.3 Introducing Service Platform Security for a Global Telematics System

As introduced in Chap. 2 [pp. 13] future VNs will include a multifunctional service platform besides the V2V communication services. To be able to open the platform to several distinct service providers it needs to be standardized. This standardization includes the security functionality, which is required to run services on such a platform. The security mechanisms need to be integrated into the platform concept, hence, it is an integral part of the system architecture. The work presented in this section has mainly been done in the context of the research project GST [Glo04], which was funded by the European Commission (EC) within the 6th Framework Programme (contract no. 507033). The GST security concept has previously been published in [RKEC04, EBM+05, EB05, GST05]. A special security solution based on the GST architecture for the Floating Car Data (FCD) service has been suggested in [Eic06]. The security concept provides a way to actively participate in a service, after being authenticated but still remaining anonymous.

To be able to understand the platform security concept, a detailed understanding of the actual platform design is not needed but helpful. The platform concepts which are needed to understand the security concept will be discussed briefly, a more detailed architecture introduction is presented in Sec. 5.2.2.

The GST Architecture Concept

Within the course of the GST project a full architecture concept has been specified, including the actors and their communication relations. The architecture is based on a key component, the *node*. All actors in the system are nodes, therefore, they all share the basic node structure.

A GST node is made up of several building blocks which interact. The organization of the block is based on the Open Systems Interconnection (OSI) layer model. Each node is a distinct hardware unit. It can interact with other nodes by means of communication. Both

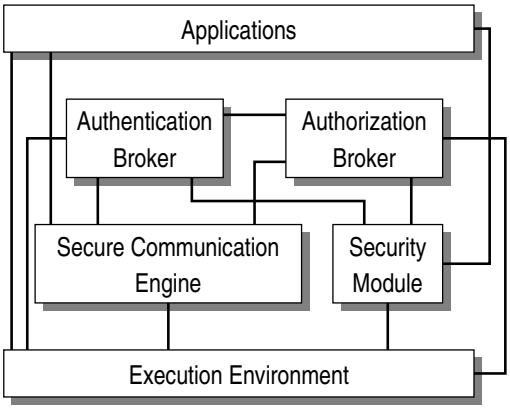


Figure (4.8) The intra-node security model used in GST [EB05]

the node architecture and communication between nodes can be secured using the intra- and inter-node security architecture presented in the following. A unique characteristic of this node concept is that one hardware entity can be used to realize one or several logical entities, for example, platform management and service provider, required for the platform architecture. The node specific building blocks are only required once and can be shared between the logical entities.

4.3.1 Providing Node Security with a Layered Security Concept

A layered security concept is used to integrate security into the node. This intra-node security is depicted in Fig. 4.8. The figure shows the relevant node building blocks and their interactions. The central building block for the node's security is the Security Module. It is a tamper-proof hardware security component, similar to a Trusted Platform Module (TPM) [Tru08], which amongst others provides cryptographic mechanisms and secure storage. The Security Module will be introduced in greater detail in Sec. 4.3.3. Besides the Security Module five other building blocks are part of the node (see Fig. 4.8).

Applications: All software components running on the platform, including services, which provide functionalities. Software components required to provide the basic node functionalities are not part of this building block.

Brokers: The two brokers provide the security functionalities of authentication and authorization. They are software components which interact closely with the Security Module. The brokers are responsible for the security of the node and its communication sessions. To fulfill this task they interact with the other node building blocks as well as distant brokers of other nodes.

Secure Communication Engine: Communication on the GST platform is dominated by session based node interactions. All security issues related to communication are handled by this building block. This includes session management, key exchanges,

		Confidentiality	
		No	Yes
Authenticity	No	*insecure*	*confidential*
	Yes	*authentic*	*secure*

Table (4.1) The four security classes used within the GST platform

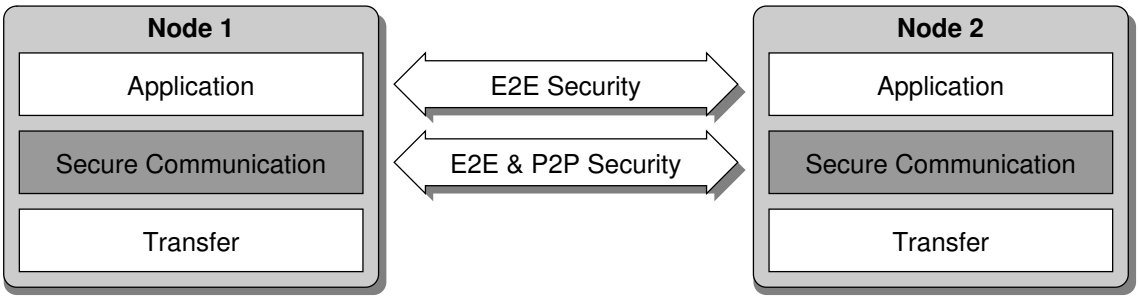


Figure (4.9) Security integration in different OSI layers of a GST node [EBM+05, EB05]

data handling, and management of the communication interfaces. The component interacts closely with the brokers.

Execution Environment: This building block represents the hard- and software execution architecture. It handles and schedules all software components running on the platform.

Four different classes of security are applied in the platform system (see Tab. 4.1). Any communication session will use one of the classes. Therefore, between fully insecure and secure the messages can only be encrypted to provide confidentiality or they can be digitally signed to provide authenticity. The security for communication is integrated into the layer stack of the system, which is depicted in Fig. 4.9. Moreover, two different security modes can be differentiated: end-to-end (E2E) security and P2P security. The E2E security mode is originating from the application layer, while the P2P security mode is initiated and terminated from the underlying secure communication layer. Since they can be combined, both modes are shown between the secure communication layers in Fig. 4.9.

The secure communication layer provides several functions to the adjacent layers. A detailed setup of this layer is presented in Fig. 4.10. The figure shows the components of the secure communication layer and its interactions. As described above, the Security Module is the fundamental building block for all security mechanisms. In addition, the layer bundles all security mechanisms and their interactions required for platform services. This enables a transparent security usage for the bordering layers. In theory, four different settings can be differentiated when using the secure communication layer.

1. Using only the Secure Communication Engine alone. This setting secures the communication sessions only: In this case, only the security class *confidential* can be applied since no authentication is included.

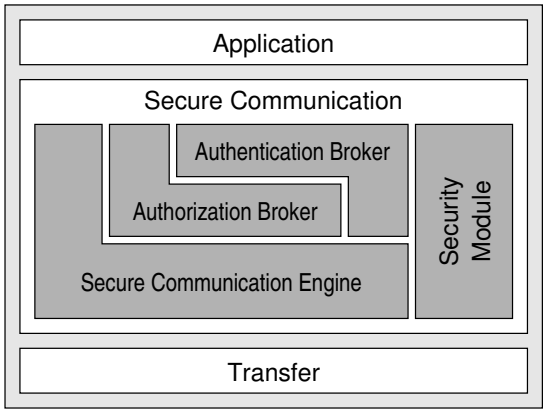


Figure (4.10) Component setup of the secure communication layer used in GST [EBM+05, EB05]

2. Using the Authentication Broker in combination with the Secure Communication Engine: Besides the secured communication sessions the node can interact in an authenticated setting, where the nodes of the E2E communication session know verifiably the identity of the communication peer.

3. Using the Authorization Broker in combination with the Secure Communication Engine: The Authorization Broker is used to provide an improved access management for users of platform services. In cooperation with respective brokers of the communication peer it makes features like Single Sign-On (SSO) feasible. Hence, based on previously exchanged access tokens, the brokers grant or deny access without any required user interaction.

4. Using all three blocks in combination: This will be the setting used in most cases. It can provide the full security functionality to the user of a node.

4.3.2 Security for Heterogeneous Communication Sessions

Communication between peers of a service platform is not limited to E2E communication links. Practically in most cases the communication can not be routed without intermediate nodes, therefore, heterogeneous communication chains with several peers are used. The intermediate peers can be simply used as a router, however, it is also likely that intermediate peers need to influence certain packet information. Hence, the communication security needs to be adapted to these cases and provide mechanisms allowing several different communication security settings.

The security variants of heterogeneous communication sessions are depicted in Fig. 4.11. Five different cases can be differentiated. All of them use one of the respective security classes (see Tab. 4.1) and can be set up with the functionalities of the secure communication layer.

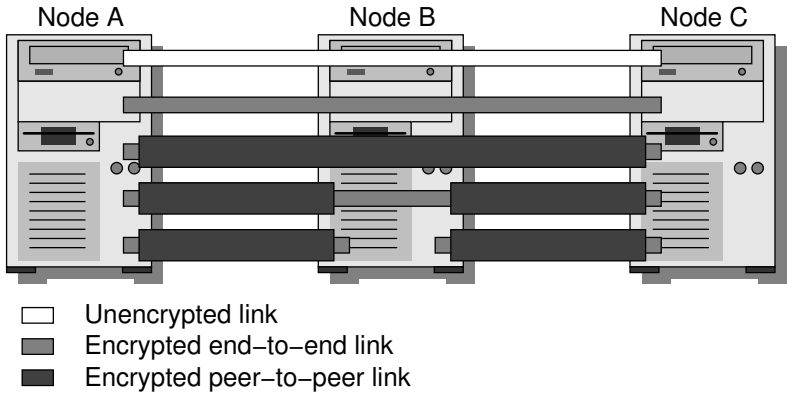


Figure (4.11) Overview on security configurations for heterogeneous communication relations [EBM+05, EB05]

Insecure communication: The communication is not secured whatsoever. All participating peers can access and alter the data.

Secure E2E data communication: The data of the communication is secured with at least one security class. The protocol information is not secured and can be altered by intermediate nodes. Thus, intermediate nodes can only do a packet forwarding.

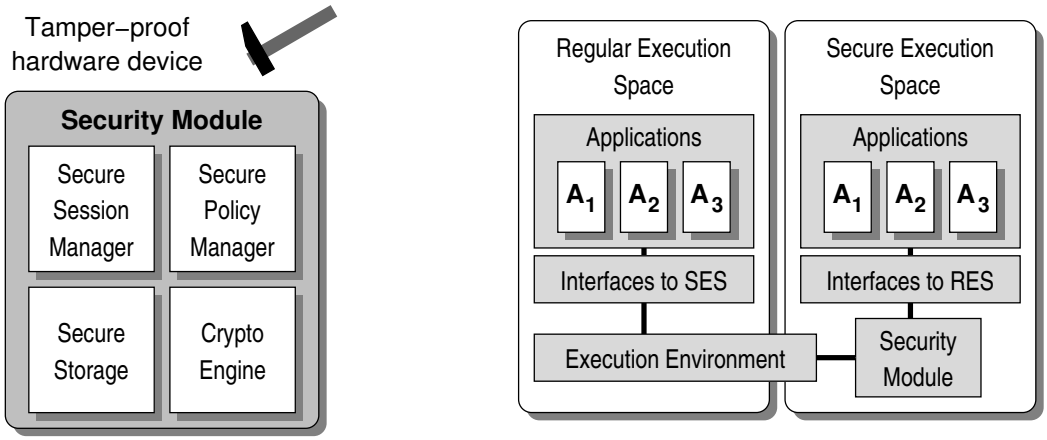
Secure E2E communication: Both data and protocol information are secured and can not be changed or even accessed by an intermediate node.

Secure E2E data communication with intermediate converters: The data, which is E2E secured, is transported in a P2P secured container. Therefore, an intermediate node can decapsulate the E2E data from the container and send it over a differently secured P2P link to its destination. The intermediate node can not access the transported data due to its E2E protection, however, a protocol conversion can be made.

Secure P2P communication with intelligent intermediate nodes: In certain situations a forwarding node needs access even to the transported data. In this case the E2E security needs to be interrupted to allow the forwarding node to transform the data and make it accessible for the final destination node.

4.3.3 Using Hardware Security as Basis for a Secure Platform

Comparable to the trust anchor in a whole system (see Sec. 4.2) the security implementation of each system component needs to be trustworthy. After all, the security of a system is determined by its weakest component. Hence, the integration of security in a node is a very delicate issue, especially since the nodes operate in the field, not controllable by the system operator. Therefore, the security needs to be protected in each node, such that a manipulation can be detected or better makes the node useless. This can be realized with a tamper-proof hardware device. Although a tamper-proof hardware like a smartcard is



(a) The tamper-proof hardware security module

(b) Secure execution environment configuration

Figure (4.12) Security Module and secure execution environment for nodes

considered to provide a very high level of security, it still has weaknesses [AK96]. They have to be kept in mind when designing a system relying on a hardware security module.

The Tamper-Proof Security Module

The Security Module introduced in Sec. 4.3.1 is such a tamper-proof device, for example, a cryptographic smartcard. It provides the basic security and cryptographic functions to the node. A block diagram of the module is shown in Fig. 4.12(a).

Besides the basic cryptographic functions signing, de-/encryption, key generation, and random number generation, it provides a secure storage for certificates and security tokens. The hardware device is an important basis to be able to unambiguously identify a node in any platform interaction. It stores the node's identity in an unchangeable fashion. In addition, the module provides the basic security management features required in the secure communication layer. Especially for the commercial applications offered on a service platform the Security Module is needed. It provides secure session logging, which is used to prove the usage of a specific service. This undeniable usage proof is a prerequisite for a solid invoicing procedure.

Splitting the Execution Space to Provide Shielded Security Execution

For some applications the use of a tamper-proof security device is not sufficient. For example applications including online payment can have certain software components which should be executed in a secure way. Since these functions can not be constricted distinctively, they can not be included in the Security Module. Hence, a secure execution space shall be used as one part of the execution environment. This shielded execution space for specific security functions strongly relies on the Security Module and provides secure processing of data, including secured memory. Therefore, critical data can not be accessed by other processes since they are shielded within the secure execution space. The integration of the secure execution space into the system is shown in Fig. 4.12(b).

4.4 Securing Vehicle-to-Vehicle Message Exchange

Besides the security for service applications communicating with Backend nodes primarily, the communication between vehicles needs to support security features as well. Especially the authenticity and integrity of messages is crucial for any message dissemination scheme in a VANET. Taking the PKI concept presented in Sec. 4.2 and Sec. 5.2.3 as a prerequisite, the security for V2V message exchanges is based on asymmetric cryptography and certificates issued by a CA.

In this section the integration of security into the message format of V2V messages is addressed. A generic message format is suggested and the overhead of the security is briefly analyzed. This is a prerequisite to judge the limitations and performance degradations caused by security mechanisms.

4.4.1 Realizing Message Security with Asymmetric Cryptography and Certificates

In a data dissemination system the receivers of messages are interested in two main things besides the validity of the data itself.

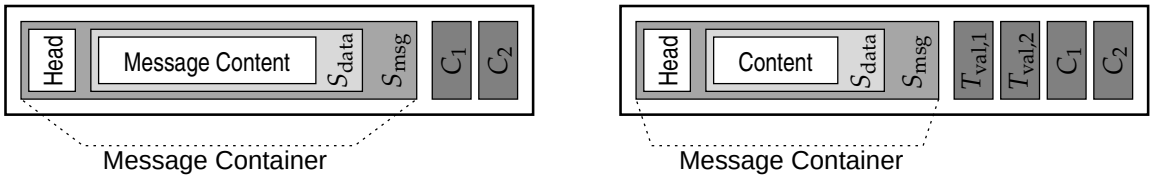
- *Authenticity*: Who sent the data and, even more crucial, is this sender a valid member of the trust environment.
- *Integrity*: Is the data unchanged, like the sender meant it to be.

A feature like confidentiality is usually undesired in a dissemination system, as long as no limited receiver group is used. However, in case confidentiality is required in certain cases it can be realized by means of encryption, using the asymmetric cryptography mechanism to transport the used symmetric key. In this case the sender needs to know the receiver and its respective public key.

Dissemination of Secured V2V Messages

As proposed in Chap. 3, messages can be distributed in VANETs with many different approaches. Adding security to these distribution mechanisms is practically identical in all cases. Therefore, a very general description how to secure V2V messages is given in the following.

- *Messages with static header*: The original sender of a message generates the content and the respective header for the dissemination message. Together they are signed using the private key of the sender. The obtained signature value is disseminated in combination with the header, the message content, and the certificate for the signature key. Any receiver will decompose the incoming message and validate the included certificate. In case the certificate is valid it is used to validate the header and message content. Otherwise, the message is dropped. Every verified message is processed further and potentially forwarded unchanged to neighboring nodes.
- *Messages with variable header*: In some settings the content of a message remains the same, while the header changes at each hop. In this case the content is signed



(a) Message format based on conventional certificates

(b) Message format based on certificates with validation ticket

Figure (4.13) Message formats for V2V data dissemination schemes

individually. Therefore, each individual disseminating ME can alter the header, sign the full message, and include its additionally required certificate. A receiver of such a message needs to verify two certificates and two signatures before the message can be processed.

These two variants to secure V2V messages lead to certain message formats, depending on the applied certificate scheme.

Message Formats for Secure V2V Messages

The suggested message formats for a secure V2V data dissemination scheme are shown in Fig. 4.13. Depending on the linkage between CSI and certificates, the message format has to be defined.

If conventional certificates are used, where the CSI is linked by means of the certificate ID only, the message format shown in Fig. 4.13(a) can be used. The signed message content and the signed header information are included in the message together with the used certificates. The certificate required for the content signature (S_{data}) is always the first certificate (C_1), while the certificate used for the message signature (S_{msg}) is always the second certificate (C_2) in the message.

In case a validation-based CSI is used, the message format can be different. This is especially the case if a ticket-based validation system such as the NOVOMODO scheme [Mic02] is applied (see Sec. 4.2.2 for details). In this case the ticket (T_{val}), required for the verification of the certificate, needs to be included into the message additionally. This is regarded in the message format shown in Fig. 4.13(b).

These are the basic message formats which can be applied for any type of dissemination application in a VANET scenario. They can be applied to all communication schemes discussed in Chap. 3 and are similar to the message format suggested in [IEE05] for the Wireless Access in Vehicular Environments (WAVE) standard.

4.4.2 Overhead of Secured Messages

Integrating security into any communication system generates overhead. Besides the data overhead for security information and certificates, additional processing power is required to generate or verify the security information. Knowing that especially V2V communication

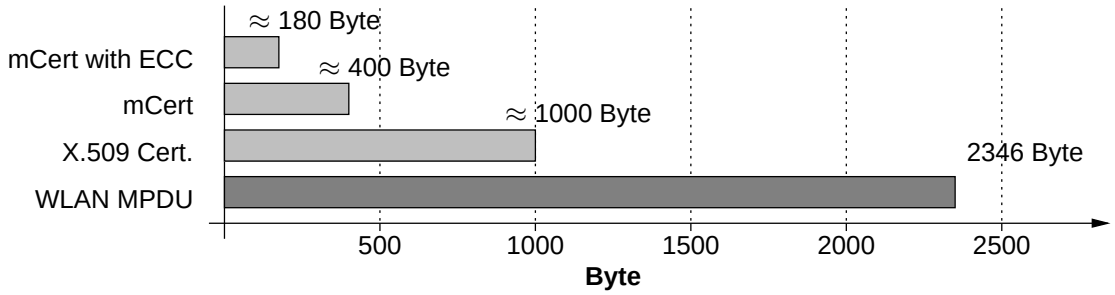


Figure (4.14) Comparison of data overheads caused by different certificate formats

suffers from large message sizes and long delays (see Sec. 2.6.1 and Chap. 3), these overheads need to be reduced as much as possible. A promising approach is to use ECC, which allows to reduce the key sizes and therefore the overhead without reducing the level of security compared to conventional mechanisms like RSA [Eck06, p. 340].

Signature information data sizes as well as certificate sizes vary strongly depending on the used certificate format and the cryptographic mechanisms. A data size comparison of two certificate formats (X.509 [HFPS99, HPFS02] and mCert [ER06a, ER06b]) using ECC or RSA is shown in Fig. 4.14. A conventional X.509 certificate contains information on the user like name and affiliation in addition to the required key pair information and the signature. This overhead has been excluded in the mCert certificate, which only contains the absolutely necessary information on the key pair, the CA, the used algorithms, as well as a node ID and the signature.

The comparison in Fig. 4.14 also includes the size of the MPDU available in the IEEE 802.11 standard family. The conventional X.509 certificate using a RSA key with 1024 bit consumes almost 50% of the MPDU, which is an absolutely unacceptable overhead for any system. As long as the security is protected sufficiently, hence, the key lengths are chosen according to the latest recommendations [Gir07], the security information size shall be kept as small as possible.

The same is true for the processing overhead introduced by cryptographic mechanisms. A very general performance comparison between RSA and the ECC-based Digital Signature Algorithm (DSA) signing mechanisms using OpenSSL [Ope08] is shown in Fig. 4.15. The performance evaluation has been done using an Intel Pentium 4 machine and an Intel Pentium 3 machine running a Linux (GNU Debian 3.1) with OpenSSL version 0.9.8c. The diagram shows the values for both *signing* (left) and *verification* (right). The value for signing with RSA using a key size of 4096 bit amounts to rather large values (141.7 ms and 536.8 ms). They are not shown in the diagram for clearness reasons.

Combining the performance analysis with the overhead comparison leads to a clear choice of technology. Preferably ECC should be applied in VANET scenarios. Nevertheless, a hybrid solution could also be considered, especially since a PKI is used to introduce trust. The CA uses RSA to issue certificates for the nodes. The CA will have enough processing power, thus, the long processing time for the signature creation can be neglected. However, the fast verification time will reduce the processing time in the MEs. A selected certificate

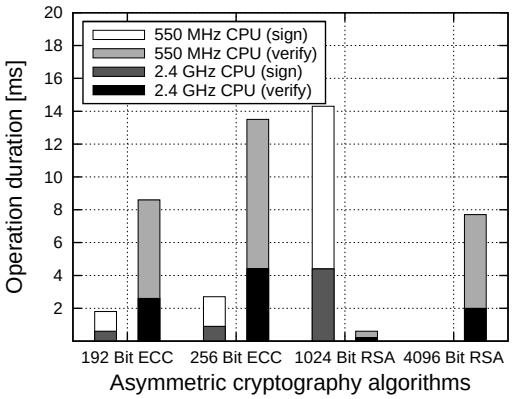


Figure (4.15) Processing overhead caused by RSA and ECC authentication mechanisms

format should contain only the most important pieces of information, for instance, the key ID, the public key, validity information, and the CA's signature. This reduces the data overhead and keeps all protocols relying on the certificates as scalable and efficient as possible.

4.5 Increasing Trust with a Content Reputation Mechanism

Multiple security architectures and protocols have been suggested for VANETs. They focus on many different aspects: secure data exchange, authenticity, confidentiality, and node privacy are just a few examples. In addition, the concept of reputation has been suggested to be used in VANET scenarios. In [SS01] *reputation* is defined as "the opinion or view of one (agent) about something". This is an adequate definition for the context of this section as well. Reputation is a subjective opinion on nodes or information, which is valid in a differentiated context only. A reputation system uses these individual opinions to improve reliability and trust in network operation mechanisms, for example routing. A determined reputation is not constant, however, it can change over time. Thus, it provides an adaptive way of assessing trust, in contrast to static cryptography-based mechanisms, for example digital signing.

A well-known example of a reputation system is the rating system used by the Internet auction portal eBay (`http://www.ebay.com/`), where each participant of an auction can rate the counterparty. Any third person can review these evaluations and deduce the reliability and trustworthiness of the participant. The use of reputation mechanisms is a promising approach to increase reliability and trust in VANET scenarios. Since VANETs are distributed and usually decentralized networks, it is very common that information is exchanged between nodes not knowing each other. Therefore, the question arises, how can a node rate the accuracy and reliability of messages rather than communication peers and still use reputation mechanisms? The use of digital signatures can provide a certain degree of authenticity and message integrity. But the significance of the message content, its accuracy, and its relevance can not be certified this way, at least not if several nodes shall participate in the opinion making. Hence, a specially designed system using reputation is suggested to

fill in this gap. In the following the new Content Reputation System (CoRS), designed for application in VANETs, is introduced.

4.5.1 Motivation for Content-based Message Reputation in VANETs

Besides the standard security functionalities, *authentication* and *integrity*, a third issue is very crucial in VANET environments: the reliability of disseminated information. The reliability or trustworthiness of information messages can be transported, for example, using a reputation mechanism. The idea is that a message's content, which has been checked and acknowledged by a group of nodes, is more reliable than a previously unchecked content. Nevertheless, simply the verification is not enough. The positive verification result has to be permanently attached to the message before starting the dissemination.

Nodes usually don't know each other in a VANET scenario. Hence, the sender of an information message will most likely be a new node to the receivers. Thus, using reputation on a node basis could be a first step. This idea has been already followed and addressed in several publications, for example [DFM05, PJFY06, RMVS05, YM03]. Nevertheless, reputation on a node basis which is stored locally is mainly of no use. Nodes will not interact on a very regular basis in a VANET. Moreover, one or very few interactions will be the norm. Hence, node reputation would have to be self-managed and provided on request. However, such a system is very hard to secure, for example to prevent manipulation. Usually, nodes participating in a VANET do not care about the identity of communication peers, as long as they belong to the system's trust architecture and hold a valid certificate. In fact, the reliability of the disseminated information is much more important, the trustworthiness of the communication peer is secondary. Therefore, the reputation mechanism should address the content information rather than the originator, especially since the trustworthiness of the peer is provided by certificates.

Content reputation can be used for many different tasks in a VANET environment. First of all it is used to certify the validity of the respective content. However, it could also be used to rank different messages based on their reputation level and select messages with higher reputation over others. This could especially be beneficial in high load scenarios. Further, it can be the basis for a message filter. The reputation can be a measure for different characteristics of the content, for instance, validity or importance. In fact, each message category can have an individual configuration for the reputation usage, for example, the level of required reputation can differ. Only the information messages reaching the required level of reputation will be disseminated through the network. All other messages will not be sent, therefore, they do not block valuable network capacity.

4.5.2 Cryptographic Requirements and Building Blocks

Before the protocol CoRS is introduced in Sec. 4.5.3 the cryptographic requirements are discussed. The Content Reputation System (CoRS) is not a stand-alone security system, however, it relies on other security mechanisms. It mainly is an additional component which complements previously installed message authentication- and integrity-preserving schemes. The two main cryptographic mechanisms required for the content reputation system, a PKI and a threshold cryptosystem, will be introduced in the following.

Supporting PKI Infrastructure and Node Security Architecture

To be able to use CoRS a PKI like the one discussed in Sec. 4.2 is required. Only MEs having a valid PKI key pair can participate in the secured VANET environment. This includes the CoRS protocol. The integration of security components into the vehicle architecture is described in Sec. 5.3 in greater detail and has previously been published in [Eic07c]. It also provides a solid basis for the integration of CoRS into the vehicle's architecture.

Secret Sharing as a Basis for Threshold Cryptography

A fundamental building block of the content reputation system is a threshold cryptography system. Many different variants of threshold cryptography mechanisms have been proposed in the literature. For all of the main public key cryptosystems, such as RSA or Diffie-Hellman Cryptosystem (DH), a threshold scheme has been proposed [DF89]. The main idea of these systems goes back to a publication by Shamir [Sha79].

In [Sha79] Shamir presented the concept of how to divide a piece of data D into n parts $P_1, P_2, \ldots, P_n$. By knowing any T or more parts P_i the full data D can be computed. However, knowing only up to $T-1$ parts leaves the data D completely undiscovered. This initial (T, n) *threshold scheme* is based on polynomial interpolation. It can be used for any data D which can be represented by a number. This is the case in practically all digital systems. The idea uses the fact that given T distinct points in the 2-dimensional plane, $(x_i, y_i) \mid i \in [1, T]$, exactly one polynomial $f(x) \mid f(x_i) = y_i$ with the degree $T-1$ can be found. The parts P_i of the data are generated with Eqn. (4.4), where $a_0 = D$ and all other a_i are selected randomly. The n different parts are determined by evaluation of $f(x)$: $P_1 = f(1), \ldots, P_n = f(n)$.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{k-1} x^{T-1} \tag{4.4}$$

By knowing any subset of T parts P_i and their respective indices, the coefficients of $f(x)$ can be calculated by interpolation. The data D can be recovered by calculating $D = f(0)$. This basic threshold scheme is the starting point for many different threshold cryptosystems suggested in the literature [DF89].

RSA-based Threshold Signatures

In a regular public key cryptosystem only two parties are required: The private key (K_{priv}) owner and the public key (K_{pub}) user. Any message which has been encrypted using K_{pub} can only be decrypted using the corresponding K_{priv} and vice versa. In addition, a message signed with K_{priv} can be verified by the corresponding K_{pub}, therefore, every receiver of this message knowing K_{pub} can authenticate its sender.

A threshold cryptosystem generating threshold signatures works very similar compared to a regular public key cryptosystem. However, a few new players and parameters exist. The following description is based on the system design presented in [Sho00]. In a threshold cryptosystem the nodes hold a threshold public key ($K_{\text{ts,pub}}$), a corresponding certificate C_{ts}, and a key share (s), which is a partial private key, rather than a private key. A single s is useless, only if enough nodes cooperate and combine their shares a digital signature can be computed. The system has several shares, $s_i \mid i \in [1, N_s]$, which are generated by the share

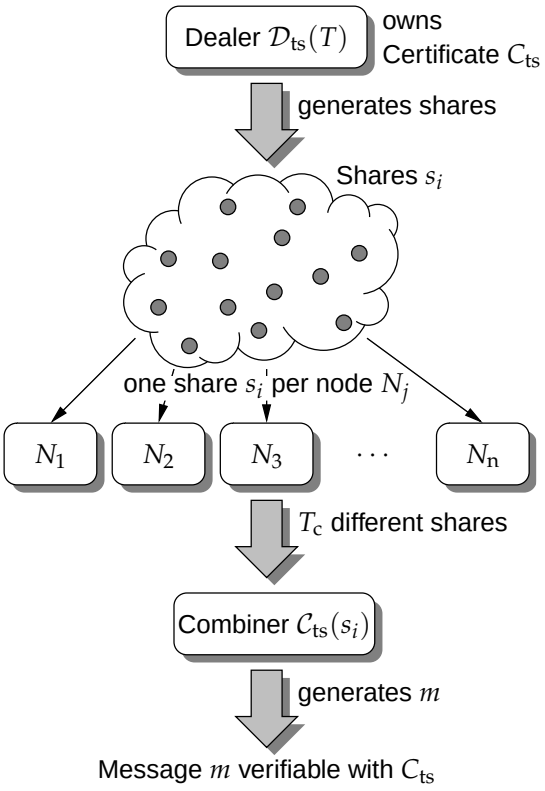


Figure (4.16) Setup and usage of a threshold cryptosystem

dealer $\mathcal{D}_{ts}$ in combination with $K_{ts,pub}$ and C_{ts}. The number of shares (N_s) is a characteristic parameter of a threshold cryptosystem. Usually each node N_j holds one individual share s_i. A very important system parameter is the threshold (T), which defines how many shares are required at least to successfully sign a message. The combination of any number of key shares smaller than T does not lead to a successful signature whatsoever. In addition, the system is robust against corrupted nodes. As long as the number of attackers (N_a) follows $T \geq N_a + 1$, the system is secure and usable. Nodes which do not follow the protocol rules or intentionally try to sign false content are considered as compromized nodes. It is important to remember that a single attacker can not harm the system, in fact, T or more nodes need to collaborate to generate a valid signature for false content.

The setup process for a threshold cryptosystem and the usage of the shares is shown in Fig. 4.16. Any T nodes $N_j \mid j \in [1, N_n]$ can sign the document using their individual key share s_i and the threshold signing operation $\mathcal{S}_{ts}(m, s_i)$ to successfully sign a document. The combiner $\mathcal{C}_{ts}$, which can be one of the nodes, uses a combination algorithm to combine the individually signed data pieces. The combined pieces result in a valid signature, which can be verified with the corresponding certificate C_{ts}. The individual key shares (s) remain secret in this combination process.

This threshold cryptosystem has been proposed by Shoup in [Sho00]. It uses the well known RSA cryptosystem [RSA78] as basis and has already been implemented (see `http:`

//sourceforge.net/projects/threshsig/). It is used in the Content Reputation System (CoRS), presented in the following section.

4.5.3 General Protocol Outline for Content-based Message Reputation

The CoRS protocol relies on a threshold cryptosystem as a basis for the reputation mechanism. Further, it needs a PKI-based security infrastructure to provide mechanisms such as node authentication and message integrity. These required mechanisms have been explained in Sec. 4.5.2 in greater detail.

Protocol Initialization Phase

An initialization is required, before the content reputation can be used in a VANET. The CA of the PKI also is the dealer ($\mathcal{D}_{\text{ts}}$) for the threshold cryptosystem and generates N_{s} key shares (s_i). Since a VANET is a constantly changing system with a varying number of nodes (N_{n}), it is not reasonable that the number of shares (N_{s}) is equal to N_{n}. To fulfill $N_{\text{n}} = N_{\text{s}}$ at all times, the dealer would have to re-initialize the threshold system with every change of N_{n} and generate new key shares as well as the connected public key. Hence, to keep the system scalable and prevent constant redistribution of shares, not every node will receive an individual s, $N_{\text{s}} \ll N_{\text{n}}$ is valid. Thus, in a group of nodes combining their shares in a transaction, collisions of equal shares can occur. For this reason two requirements need to be fulfilled for N_{s}: Enough shares must exist to achieve a reasonable T ($3 \leq T \leq 10$) and the collision probability of two or more equal shares, used for one transaction, should be as small as possible. This also depends on the node density of the scenario.

The shares s_i are equally distributed to the nodes to minimize the collision probability. Hence, each new node will receive one share, such that each share is handed out equally often. In addition to the share, each node receives at least one conventional key pair $\{K_{\text{priv}}, K_{\text{pub}}\}$ and a corresponding certificate of the underlying PKI, which is used for node authentication.

The Protocol Mechanisms of CoRS

After the initialization of the system, the content reputation protocol can be used. The use of CoRS starts with an event detection at the *generator*. CoRS handles all events which can be detected by several nodes, for instance, traffic events, local danger warnings, or weather events. The protocol steps followed in the generator are shown in Alg. 4.1. Suppose the node N_j detects a traffic event, for example the end of a traffic jam. It will generate an information message m_{info}. However, before disseminating the message to other nodes the content reputation protocol is applied. Refer to Fig. 4.17 for the protocol players and their message exchanges and to Fig. 4.18 for the protocol steps (1–8) inside the players.

Node N_j is in the role of the *generator*. It detects the event, generates the information message (step 1) and needs support of surrounding nodes located inside the *reputation area*. Since the direct neighbors will most likely also have detected the event, they can act as *verifiers*. After N_j generates a reputation request, it will digitally sign the request message

if *event detected* **then**
  collect reputation parameters r_p from context information;
  generate request m_{req};
  sign m_{req}: $\mathcal{S}_{\text{pki}}(m_{\text{req}}, K_{\text{priv}}) \rightarrow Sig_{\text{pki}}$;
  send signed request: $m_{\text{req}} \circ Sig_{\text{pki}}$;
end
while $t \leq t_{\text{thresh}}$ **do**
  receive m_{rep};
end
if $\text{Num}(m_{\text{rep}}) < T$ **then**
  resend signed request: $m_{\text{req}} \circ Sig_{\text{pki}}$;
else
  for $i \in [1, N_{\text{msg}}]$ **do**
    extract $Sig_{\text{pki},i} \hookleftarrow m_{\text{rep},i}$;
    if *verify* $Sig_{\text{pki},i} =$ true **then**
      extract x_i;
    end
  end
  generate Sig_{ts}: $\mathcal{C}_{\text{ts}}(x_i)$;
  sign $m_{\text{info}} \circ Sig_{\text{ts}}$: $\mathcal{S}_{\text{pki}}(m_{\text{info}} \circ Sig_{\text{ts}}, K_{\text{priv}}) \rightarrow Sig_{\text{pki}}$;
  send $(m_{\text{info}} \circ Sig_{\text{ts}} \circ Sig_{\text{pki}})$;
end

Algorithm (4.1) CoRS protocol steps in the generator node

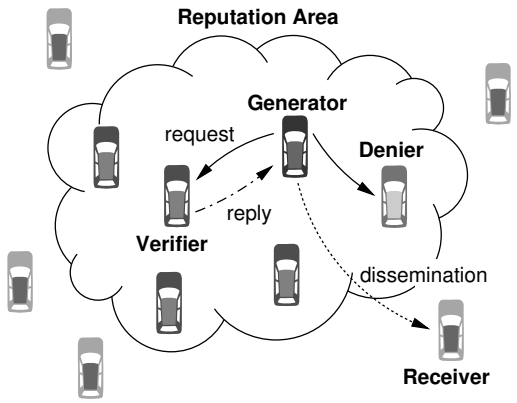


Figure (4.17) Actors and protocol steps of a CoRS message reputation scenario

using its individual private key (K_{priv}). The signed verification request, also containing the information message itself, is send to the *verifiers* (step 2).

```
1  if received m_req then
2      extract Sig_pki ↩ m_req;
3      if verify Sig_pki = true then
4          extract m_info ↩ m_req;
5          evaluate m_info ⇒ R_info(m_info);
6          if R_msg(m_info) ≥ 0 then
7              generate threshold signature share x_i: S_ts(m_info, s_i) → x_i;
8              encapsulate x_i: x_i ↪ m_rep;
9              sign m_rep: S_pki(m_rep, K_priv) → Sig_pki;
10             send (m_rep ∘ Sig_pki);
11         end
12     end
13 end
```

Algorithm (4.2) CoRS protocol steps in the verifier node

Each receiver of the verification request message is in the role of a *verifier*. The protocol steps of CoRS followed in a verifier are shown in Alg. 4.2. At first the request's digital signature is verified using the corresponding public key (K_{pub}) and its certificate. Only request messages with a valid signature are processed any further. Hence, the signature helps to identify a valid request, sent from a node which is part of the underlying PKI. Only members of this circle-of-trust are allowed to participate in the reputation protocol. Due to the fact that nodes do not have an individual share s, they can not be identified unambiguously using s. Therefore, the additional signature verification provides a sufficient authentication verification instrument. After the signature verification, each *verifier* checks the content of the information message and tries to validate several pieces of information to determine the values for the reputation parameters $r_p \in \mathcal{R}$ (step 5). The reputation parameters are used to rate selected event and message characteristics. The rated characteristics are combined to determine a message reputation value. Each message category can have its own reputation parameters in addition to the following general parameters:

Event (r_{e}): The verifier has to check if it detected the same event as the generator. If this is the case, the node can evaluate the message m_{info} further. Otherwise the protocol ends and the node will not sign the request with its key share s_i.

Event detection time (r_{et}): The generator saves the detection time and uses it as a time stamp for the information message. The verifier checks that the time stamp is in the past and the time delay between detection and verification is smaller than a given threshold. Thus, only recent events are handled by CoRS.

Location of the event (r_{el}): Each event has a distinct location or an event area, where the event is actually present. The verifier checks if the stated location is identical or within a certain range of the event it detected itself.

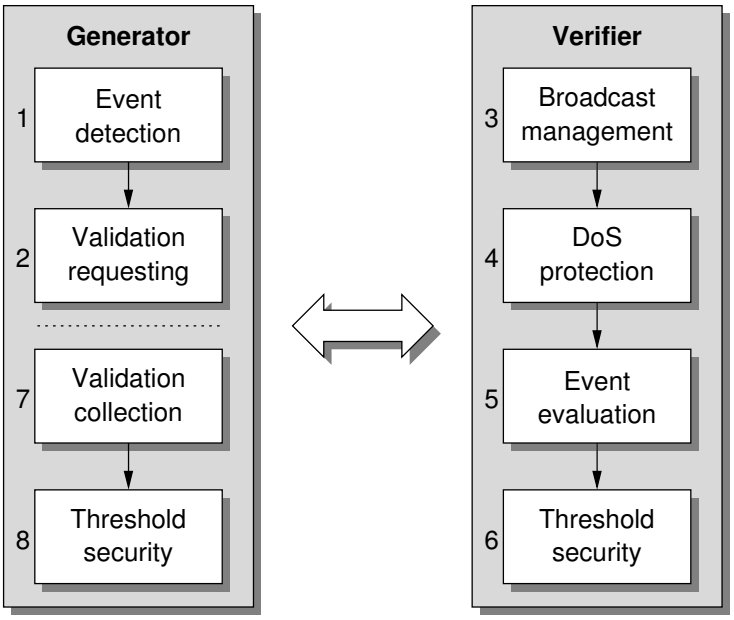


Figure (4.18) Protocol steps and interactions between generator and verifier used for CoRS

Location of the node (r_{sl}): The location of the generator and of the event can differ, mainly due to the mobility of nodes and events. Thus, a collator has to check if the location of the generator is within a reasonable distance to the verifier and the event.

Sending time of the request (r_{st}): The sending time of a request can differ compared to the event detection time. If at the time of the event detection not enough collators can be found, a second request will be sent with a slight delay. Hence, the sending time can be different from the detection time. Therefore, the collator needs to check that the sending time corresponds to the actual system time, allowing a certain tolerance.

The stated parameters are the base parameters applicable to all type of events. Event specific parameter checks can be considered additionally. To rate the parameters and specifically the validation results, a reputation calculation mechanism is suggested. Each of the reputation parameters r_p of the parameter set $\mathcal{R}$ can adopt values of $+1$ for positive validation, 0 for neutral validation, -1 for negative validation. Thus, the overall message reputation value can be calculated with Eqn. (4.5).

$$R_{\text{msg}}(m_{\text{info}}) = \sum_{p \in \mathcal{R}} r_p \qquad R_{\text{msg}} \in [-|\mathcal{R}|, +|\mathcal{R}|] \tag{4.5}$$

As long as $R_{\text{msg}} \geq 0$ holds true the message is validated with a positive reputation and the verifier can send a reply to the generator. Otherwise, the request is deleted and remains unanswered. In this case the node is a denier (see Fig. 4.17).

Suppose $R_{\text{msg}} \geq 0$ holds true. Therefore, the *verifier* successfully validated the event and will send a reply message. To generate the reply message the threshold cryptosystem is used.

The *verifier* extracts the information message from the request and digitally signs it with its key share s_i (step 6) using the signing operation $\mathcal{S}_{\text{ts}}(m_{\text{info}}, s_i)$. The obtained signature value x_i is embedded into the reply message with some status information, linking it to the respective request. Before sending the reply, it is digitally signed with the private key (K_{priv}) of the verifier. This ensures the authenticity and integrity of the reply message.

The generator collects all incoming reply messages (step 7). Before processing them the digital signatures are validated, therefore, only authenticated replies are used, sent by member nodes of the PKI. The generator checks the status information and matches the replies to the respective request. Next, the signature value x_i, generated with the threshold scheme, is extracted. If enough distinct signatures shares x_i, $|i| \geq T$, generated with different s_i, have been obtained, the generator can combine them. This is done with the combination algorithm $\mathcal{C}_{\text{ts}}(x_i)$ and the signature shares provided by the verifiers. The combination algorithm generates a valid threshold signature (Sig_{ts}) for the event information message m_{info} (step 8). The threshold signature can then be verified by any receiver holding the corresponding threshold signature certificate C_{ts}. It certifies that at least T nodes have validated the event. The signed information message m_{info} is distributed to other receivers outside of the *reputation area*. The threshold signature (Sig_{ts}) represents the reputation value of the message. The larger T the higher the reputation, since it is harder to find enough verifiers. Hence, it is a possibility to use several threshold key settings in parallel, each having its distinct T. The generator can request a signature made by the key shares corresponding to a dedicated T. Therefore, the generator can selectively chose the level of reputation to be used for a message.

All steps of the protocol are also shown in Fig. 4.18. Steps three and four have not been discussed so far. They are not directly required for the reputation protocol, however, they adapt the protocol to the specific scenario requirements of VANETs.

The broadcast management (step 3) is a filter to prevent Broadcast Storms. Therefore, only previously unseen messages are forwarded to the lower components of the protocol stack. All other messages are deleted and not handled again, reducing the network and processing load.

The DoS protection block protects the system against irregular or false messages (step 4). The block maintains a list of peers that have not followed the rules of CoRS. All incoming messages which originated from one of these nodes will not be processed. A node will be added to this block list if it continuously sends unsigned messages, request messages with wrong content, or information messages with forged threshold signatures. Hence, the DoS protection prevents the node from using up its processing power for meaningless cryptographic operations.

4.5.4 Analytical Evaluation of the Share-Collision Probability

A very crucial issue is the possibility of share collisions. Since nodes do not hold individual key shares for the threshold cryptosystem, it is possible that replies to a reputation request contain a signature share made with the same key share (see Sec. 4.5.3 [p. 109] for details). Hence, combining them does not increase the entropy for the signature. To estimate the collision probability for different values of N_{s} and T an analytical examination of the issue

has been made. The calculation has been done in several steps using combinatorial logic. First, the special case of $T = 2$ is examined, before generalizing to arbitrary values of T.

Imagine the following scenario with N_n nodes within the reputation area. One node is the generator, therefore, the number of neighbors is determined by $N_\mathrm{ne} = N_\mathrm{n} - 1$. A threshold cryptosystem with N_s shares and a threshold T is used. What is the probability $\overline{p_\mathrm{coll}}$ receiving $n \geq T$ different s_i if N_s different shares exist and N_ne neighbors are within the reputation area? The problem can be solved in three steps using combinatorial analysis:

1. Calculation of the lower bound ($N_\mathrm{ne} = T$).
2. Calculation of the special case $T = 2 \mid N_\mathrm{ne} \in [T, \infty)$.
3. Generalizing the special case for $T > 2 \mid N_\mathrm{ne} \in [T, \infty)$.

Lower Bound and Special Case $T = 2$

The lower bound and the special case can be solved using the formula for variations V_n without repetitions (Eqn. (4.6)) and V_n with repetitions (Eqn. (4.7)), considering the order of elements [BSMM99, p. 745].

$$V_n = k! \cdot \binom{n}{k} \tag{4.6}$$

$$V_n = n^k \tag{4.7}$$

For the first step Eqn. (4.6) and Eqn. (4.7) can directly be used to calculate the lower bound for the collision probability (see Eqn, (4.8)).

$$p_{\mathrm{coll},N_\mathrm{ne}=T}(T, N_\mathrm{s}) = 1 - \frac{T! \cdot \binom{N_\mathrm{s}}{T}}{(N_\mathrm{s})^T} \tag{4.8}$$

The lower bound describes the worst case scenario of $N_\mathrm{ne} = T$, where no collision is allowed to fulfill the threshold. If N_ne exceeds T the collision probability can be reduced significantly. Fig. 4.19 shows the analytical results of $p_{\mathrm{coll},N_\mathrm{ne}=T}$ for different values of T and N_s.

Knowing how to determine the lower bound of the share collision probability is an important step, however, a universal equation shall be found to be able to calculate any combinations of T and N_ne, for any given value of N_s. The first step towards this universal solution is the special case where $T = 2$. This case is special since only the combinations where all key shares are equal are a collision. All other cases are valid combinations.

$$N(T, N_\mathrm{s}) = T! \cdot \binom{N_\mathrm{s}}{T} \tag{4.9}$$

$$\begin{aligned} N_{T=2}(N_\mathrm{ne}, N_\mathrm{s}) &= N(N_\mathrm{ne} - 1) \cdot N_\mathrm{s} + N(T) \\ &= N(T) \cdot \sum_{i=T}^{N_\mathrm{ne}} N_\mathrm{s}^{(i-T)} \end{aligned} \tag{4.10}$$

$$p_{\mathrm{coll},T=2}(T, N_\mathrm{s}, N_\mathrm{ne}) = 1 - \frac{N_{T=2}}{(N_\mathrm{s})^{N_\mathrm{ne}}} \tag{4.11}$$

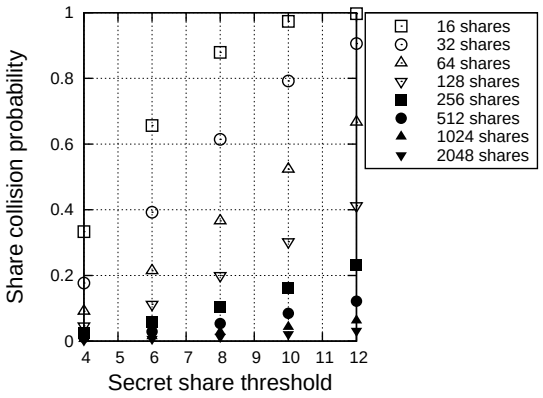


Figure (4.19) Share collision probability for $N_{\text{ne}} = T$

The equation for determining the variations without repetitions (Eqn. (4.6)) is used as a basis. Parameters are T and N_{s} (see Eqn. (4.9)). The number of collision free combinations for the threshold $T = 2$ can be calculated with Eqn. (4.10), the parameters are number of neighbors (N_{ne}), number of shares (N_{s}), and the threshold (T). The equation is a sum of two parts. The first summand determines all combinations that have been valid for $N_{\text{ne}} - 1$, they are still valid for N_{ne}. However, since an extra neighbor exists, N_{s} different combinations exist. The second summand stands for all newly valid combinations. The recursive nature of Eqn. (4.10) can be resolved and rewritten with a sum. The share collision probability for the case $T = 2$ is determined by Eqn. (4.11).

Generalized Calculation

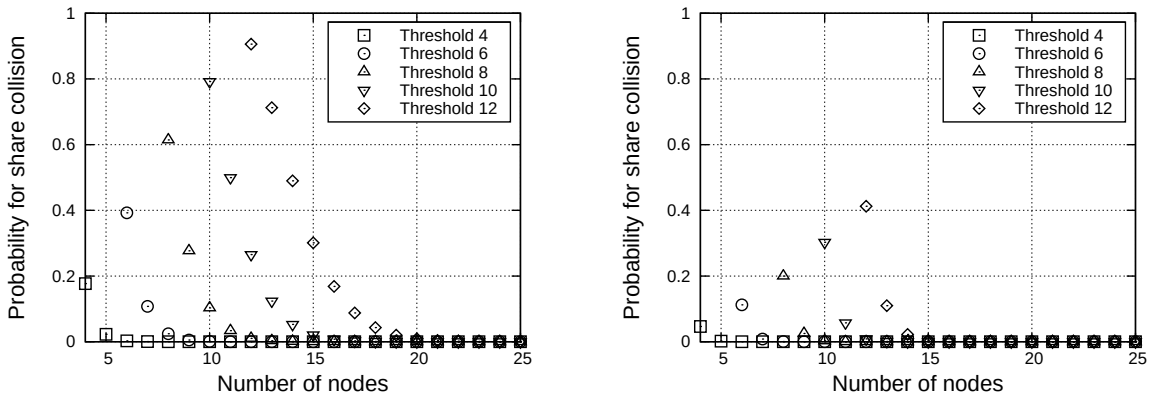
Having solved the special case for $T = 2$, a general solution for $T > 2$ can be found. A similar approach like the one used for $T = 2$ is used for the general solution, a sum with two summands (see Eqn. (4.12)).

$$N_{T>2}(N_{\text{ne}}, N_{\text{s}}) = N(N_{\text{ne}} - 1) \cdot N_{\text{s}} + [N_{T-1}(N_{\text{ne}} - 1) - N_T(N_{\text{ne}} - 1)] \cdot [N_{\text{s}} - T + 1] \tag{4.12}$$

$$p_{\text{coll}}(T, N_{\text{s}}, N_{\text{ne}}) = 1 - \frac{N_{T>2}}{(N_{\text{s}})^{N_{\text{ne}}}} \tag{4.13}$$

Eqn. (4.12) is the solution for the general case of $T > 2$ with the parameters N_{ne} and N_{s}. The first summand of Eqn. (4.12) are the combinations that were already valid for $N_{\text{ne}} - 1$ multiplied with its N_{s} different combinations of the additional share. The second summand is used to determine all new combinations that have not been valid before. This part of the equation uses a recursion, since terms with $T - 1$ and $N_{\text{ne}} - 1$ are used. The share collision probability can be calculated using Eqn. (4.13).

Using Eqn. (4.11) and Eqn. (4.13) a comparison of share collision probabilities can be done. Two cases have been calculated and the probabilities for *no* share collisions are shown in Fig. 4.20. These results support several things. If $N_{\text{s}} < 128$ the threshold (T)



(a) Number of shares $N_s = 32$

(b) Number of shares $N_s = 128$

Figure (4.20) Analytical comparison of probabilities for share collisions

dominates the signing success, hence, depending on T the probability for attaining a valid threshold signature varies. For $N_s > 128$ the number of neighbors (N_{ne}) dominates the signing success, since already the lower bounds shows a high probability of attaining a signature ($\overline{p_{coll}} \approx 60\%$).

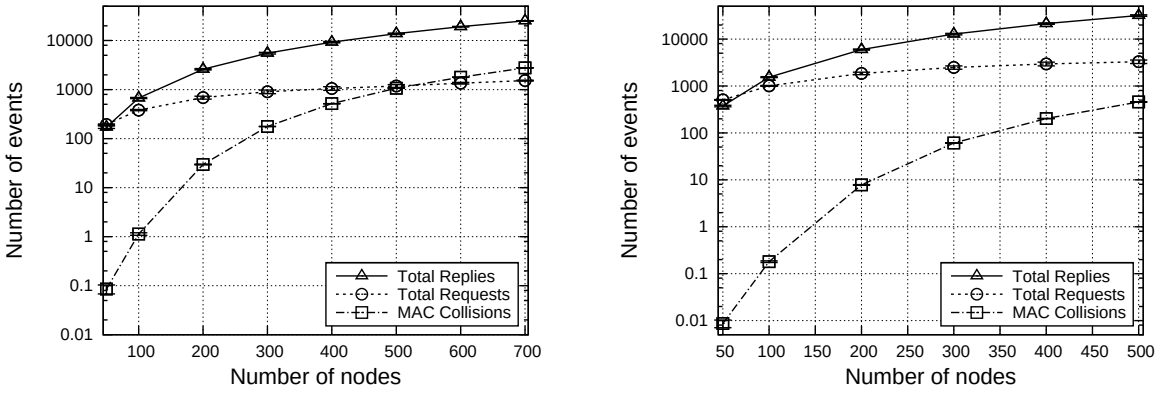
The analytical evaluation of the share collision probability helps to evaluate the viability and efficiency of the cryptosystem usage. The results support the idea that a relatively small number of shares ($128 \leq N_s \leq 2048$) is enough to provide the threshold signature functionality to CoRS. For a threshold of $T = 12$ this setting results in a collision probability of sightly over 40% for the worst case with $N_s = 128$ and $N_{ne} = T$.

4.5.5 Simulative System Evaluation

The content reputation protocol CoRS has been evaluated within a network simulator. Before integrating the protocol into the simulation environment a full system model using the Specification and Description Language (SDL) has been implemented and evaluated to identify crucial protocol parameters and timers. After improving the protocol using the SDL model, a simulation model has been designed. This has been used to run different simulation scenarios, to evaluate the protocol under different settings. The used simulation framework, the supporting models, and the respective parameters are presented in the following.

Simulation Framework and Parameters

The simulation model of CoRS uses several delays and timers. An average application delay of 100 ms and an average network layer delay of 5 ms were used. A processing time for newly detected events of 5 s was set. During this time the model waits for surrounding nodes to detect the same event as well to be able to react to a request. With a delay of 5 s a follow-up request was sent, if the first one remains unanswered. A packet size of 600 B was assumed for the reputation mechanisms. The nodes randomly moved on the Manhattan



(a) Fixed node group

(b) Changing node group

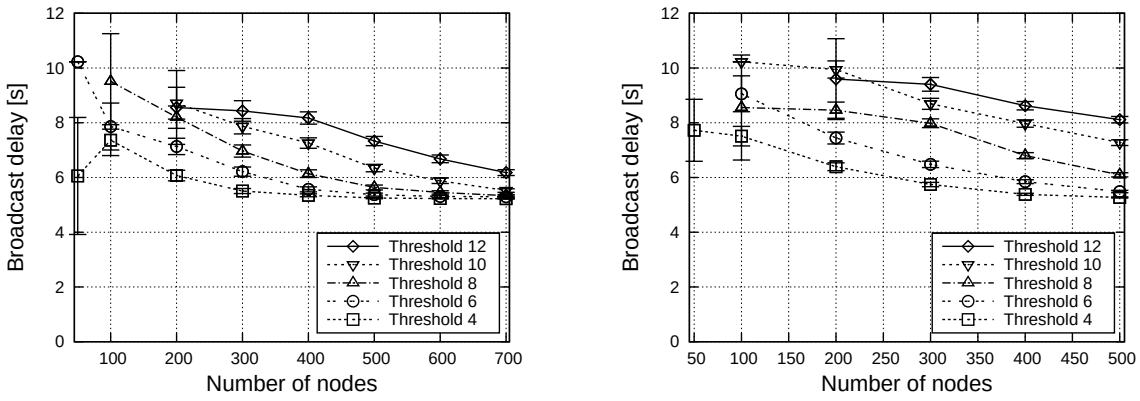
Figure (4.21) Evaluation results for collisions, requests, and replies in a CoRS scenario

Grid with its three horizontal and three vertical roads. In the center of the scenario a section of the middle road was the event area. Vehicles entering this area detect the event and used CoRS to acquire a signed information message. All neighboring nodes participated in the reputation protocol process. Nodes remembered a detected event for a time of 300 s after the last event update. Thus, only nodes not already aware of the event tried to broadcast an information message, while other nodes only participate in the protocol.

The simulation time for the scenarios was set to 1800 s. Several scenarios with different values of $T \mid T \in \{4, 6, 8, 10, 12\}$, $N_\mathrm{s} \mid N_\mathrm{s} \in \{128, 256, 512, 1024, 2048\}$, and $N_\mathrm{n} \mid N_\mathrm{n} \in \{50, 100, 200, 300, 400, 500, 600, 700\}$ were simulated. The Manhattan Grid Mobility (MGM) model provides the possibility to replace nodes if they reach the border of the simulation area. This simulates the continuous change of neighbor nodes in VANETs and was applied in some simulation scenarios, which are called *changing node group* scenarios in the following. For scenarios with the changing node group a maximum node density of $N_\mathrm{n} = 500$ was applied, in contrast to a maximum node density of $N_\mathrm{n} = 700$ for the fixed node group scenarios. The reason for this difference is the much higher simulation times required for the changing node group scenarios, due to the complexity of generating and deleting nodes in the model during the simulation execution. In the following, a selection of simulation results are given for both scenario types.

Simulative System Evaluation Results

A very important analysis for any wireless protocol is its throughput requirements and the tendency to cause packet collisions on the channel. The total number of requests, replies, and collisions are shown in Fig. 4.21. In Fig. 4.21(a) the results for the fixed node group scenarios are plotted. The graph shows the strong increase of packet collisions for increasing node densities. However, the number of replies dominate the events in the scenarios. This was to be expected since no limitation for replying exists other than not having validated the request. Moreover, the number of replies is depending on N_ne, which increases with N_n. This also explains the large difference between requests and replies.



(a) Fixed node group

(b) Changing node group

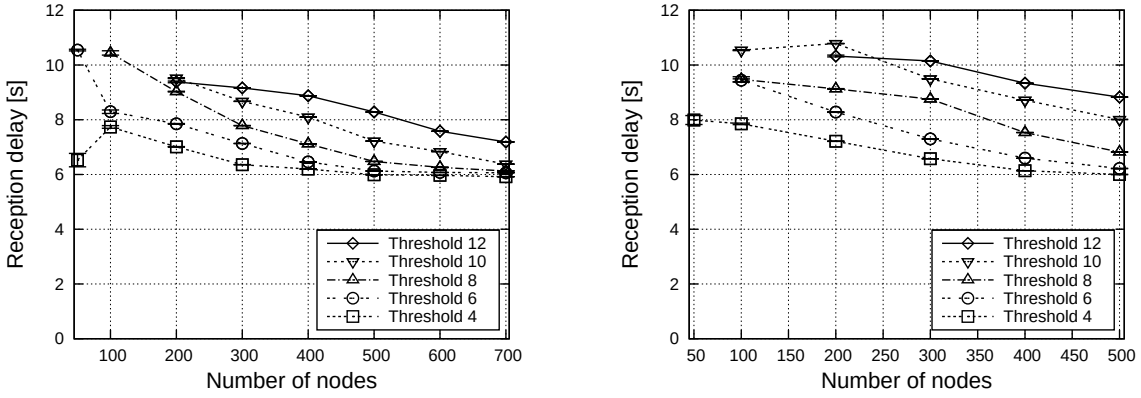
Figure (4.22) Average broadcast delay caused by CoRS for $N_s = 128$

The results for the changing node group shown in Fig. 4.21(b) are similar. However, the number of requests and replies increases stronger compared to the fixed node group results. The lifetime of a node in a scenario with $t_{sim} = 1800$ s is around 150 s. Hence, many new nodes will detect the event and send request messages, explaining the stronger increase.

The parameter N_s in combination with the node density and the threshold is most relevant for the collision free protocol operation. Since $N_s = 128$ is the worst case scenario the respective simulation results have been selected for the following protocol analysis of CoRS. The results for $N_s > 128$ are equivalent, in fact the protocol performance is better, showing shorter delays and less requests, due to a lower share collision probability. An important value for the performance of the protocol is the delay it introduces to the information distribution process. In Fig. 4.22 the broadcast delay, which is the time between event detection and information dissemination, is depicted. Remember that the processing delay for events is at least 5 s due to the model settings. Hence, the lower bound for the broadcast delay is 5 s. The missing data points in the graphs result from the fact that in these cases no broadcast has been made, due to missing signature shares.

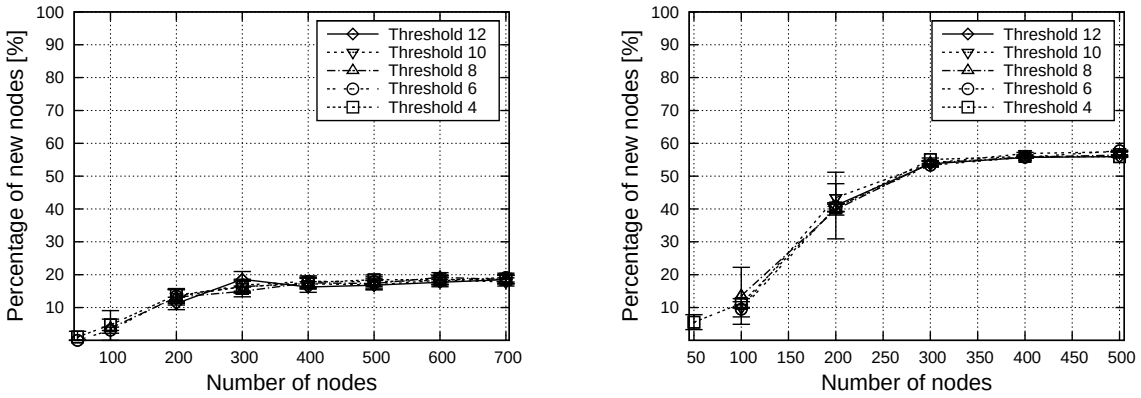
For the fixed node scenarios the average delay ranges between just over 5 s and 10 s (see Fig. 4.22(a)). The higher the threshold the higher is the average delay, which is to be expected. The graphs show all data points for scenarios with $N_n \geq 200$, therefore, these scenarios have enough nodes to reach up to 12 or more neighbors. The delay values approach the lower bound the more nodes are in the scenario. In Fig. 4.22(b) the broadcast delays for the changing node group scenarios are shown. The delays are slightly higher due to the fact that more nodes are new to the scenario, hence, they can not yet verify the event. In the closed scenarios the probability that a node has not seen the event is much smaller, therefore, the number of replies is higher. This leads to the shorter delays shown in Fig. 4.22(a).

Besides the broadcast delay the reception delay is important to evaluate the influence of the network load on the information distribution. The reception delay is the time between the event detection until the signed information message is received by a node. The corresponding results are shown in Fig. 4.23. The graphs are very similar to the broadcast



(a) Fixed node group **(b)** Changing node group

Figure (4.23) Average Reception delay caused by CoRS for $N_s = 128$

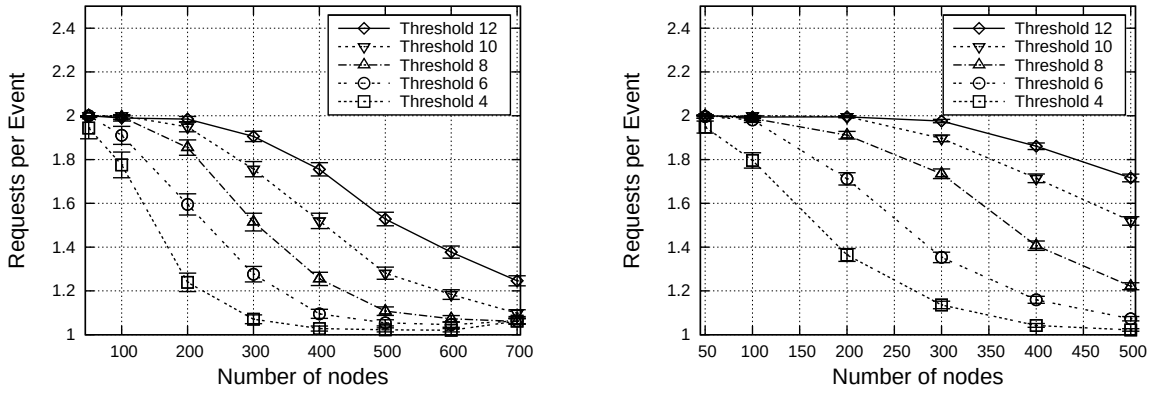


(a) Fixed node group **(b)** Changing node group

Figure (4.24) Broadcast impact of message dissemination using content reputation

delay graphs in Fig. 4.22. However, an increase of about 0.5–1.0 s compared to the broadcast delay can be seen. This is a very reasonable delay for a decentralized dissemination system.

The delays are an important criteria to evaluate the performance of the dissemination and the load on the network. Additionally, the impact of the dissemination process should be analyzed. The question is, how many nodes receive the information messages without knowing of the respective event. This impact has been analyzed and plotted in Fig. 4.24. The figure shows the relative impact of new nodes compared to all nodes receiving an information broadcast. The results for the two scenario types differ significantly. In the changing node group scenarios, shown in Fig. 4.24(b), the impact is much higher than for the fixed node group scenarios depicted in Fig. 4.24(a). In a real life scenario the impact would be similar to the results shown in Fig. 4.24(b), since the changing node group corresponds to the settings in a real scenario. These results prove that even if around 60% of the nodes



(a) Fixed node group

(b) Changing node group

Figure (4.25) Average number of requests needed per detected event to acquire $\geq T$ signature shares for $N_\mathrm{s} = 128$

are not familiar with an event, the content reputation is still usable and provides validated information with reasonable delays to surrounding nodes.

Finally, the number of requests that need to be send to acquire enough signature shares from neighboring nodes shall be evaluated. Optimally, only one request is needed to receive enough signature shares, however, as results in Fig. 4.25 show, this is not the case for most scenarios. For the low node density scenarios ($N_\mathrm{n} \leq 200$) on average almost two requests have to be sent before the information broadcast can be carried out. With increasing node density the average request number approaches the minimum of one request. The results for the changing node groups scenarios in Fig. 4.25(b) show a slightly higher value for the average requests per event. This is again caused by the higher number of nodes which are not aware of the event.

The presented results were are all done with $N_\mathrm{s} = 128$. The same simulations have been done with higher values for N_s. As expected the results are equivalent to the presented results. The scenarios with $N_\mathrm{s} = 2048$ showed slightly lower delay values, primarily due to the lower share collision probability. Therefore, the analytical evaluation of the share handling has been confirmed by the network simulations. In scenarios where $N_\mathrm{s} \geq 128$ the number of neighbors dominates the protocol behavior, rather than the threshold. Moreover, the potential and the performance of CoRS has been shown with the given simulation results.

4.5.6 Assessment of the Simulation Results for the Content Reputation Protocol

The proposed protocol provides a mechanism to realize a content reputation attachment for information dissemination in VANETs. It is designed for information which can be verified by several nodes individually and does not require instant dissemination, such as collision warning messages. The system uses a threshold cryptography technology, like the one presented by Shoup in [Sho00], to attach the content reputation level to an information message in a secure and verifiable way in form of a threshold signature. CoRS is robust due to the characteristics of the threshold signature scheme, hence, it can withstand a certain

amount of attackers in the system, depending on the threshold value (T). Attackers can only harm the system security in collaboration. A single attacker can not break the threshold scheme and generate a false signature.

The suggested content reputation system CoRS is a new and effective approach to increase security and especially trust in information dissemination systems such as VANETs. The mathematical analysis of the share collision probability proves that the reputation protocol works, even with a relatively small number of shares. Moreover, the use of CoRS does not significantly increase the delays for message dissemination. In combination with a PKI-based cryptosystem the suggested reputation scheme can be realized in a scalable and efficient way.

With the CoRS protocol a new way of combining threshold cryptography with reputation mechanisms is suggested. The threshold cryptography can be used to efficiently realize a secure and verifiable attachment of the reputation to an information message. The CoRS protocol is one possible realization for such a threshold cryptography-based content reputation system designed for VANET scenarios.

4.6 Privacy Evaluation of the VANET Communication Scenario

Besides security another very important feature needs to be realized for VNs, node privacy. Especially due to the integration of communication mechanisms into future vehicles, new demands for the privacy protection of such vehicles and their passengers arise. When communicating a peer automatically leaves traces, for example, its Medium Access Control (MAC)-address or a communication session identifier. In combination with the security demands for VANET scenarios these traces can be assigned to a specific node distinctively. Therefore, the combination of communication and security mechanisms enables a highly reliable traceability of vehicles and potentially even drivers, breaching their privacy. This situation demands to integrate privacy protecting mechanisms into the system architecture of future VNs. Otherwise, no broad acceptance will be gained in the public for such systems.

4.6.1 Privacy in VANET Scenarios

Before analyzing the use of pseudonyms a short introduction on privacy in VANETs is given. It can also be seen as a motivation to realize privacy in these scenarios to increase the users' acceptance.

Required Privacy Features

The term privacy stands for several distinct features. The Oxford English Dictionary [SW89] defines privacy as:

> "The state or condition of being alone, undisturbed, or free from public attention, as a matter of choice or right; protection from public knowledge or availability."

This definition can very well be applied to the VANET setting (also refer to [PH07] for terminology). The following privacy mechanisms are crucial for future VNs:

Anonymity: According to [PH07] anonymity means "a subject is not identifiable". This perfectly fits the VANET case. Nodes, their users, and their respective identities shall remain anonymous to any other entity in the system, unless desired otherwise. Hence, an entity can not identify an other entity without its approval.

Unlinkability: If two entities, messages, etc. can not be brought into relation with each other they are considered unlinkable [PH07]. This is a very important requirement for many different settings in VANETs. Generally it can be demanded that any combination of nodes, user IDs, events, or messages shall remain unlinkable for third persons. Hence, a message should not be directly linkable to its sending node at any given point in time.

Location privacy: The location of a node at a defined moment as well as the related movement pattern has to remain unknown to other nodes. Therefore, it shall not be possible to link a node to an event in time which occurred at a distinct location.

Mechanisms to Realize Privacy Protection

Privacy protection has been investigated in several contexts in the literature. Several different mechanisms have been suggested to realize privacy protection in real systems. The use of Chaum's "Dining Cryptographers Problem" leads to a mechanism for absolute sender and recipient privacy [Cha88]. The mechanism enables the computation of an XOR including more than two participating nodes. The nodes are organized in a ring network using encrypted links. Every node selects a random number $n \in [0, 1]$ and shows it secretly to the node on its right side. The nodes compute the XOR value using the own number and the number received from the neighbor at the left side. The result is published to all nodes in the network, unless the node wants to transmit a message. For a message transmission, the node inverts the XOR computation in all rounds corresponding to a 1 in the binary representation of the message. The XOR combination of the published results corresponds to the binary value of the data, as long as only one node tried to send a message. Otherwise a collision occurred and the sending needs to be restarted. The scheme protects the privacy of sending and receiving nodes, however, it is rather costly to implement and use in distributed networks.

A different mechanism is the use of one or several proxy servers which remove the identifying information of messages before forwarding them. This basic concept can be expanded to the so-called Mix concept [Cha81]. A mix is a special proxy server hiding the context of distinct messages in the cloud of all received and processed messages. A very sophisticated web traffic anonymizer is the Crowds concept [RR98], similar to the mix idea. In Crowds, the participating nodes are grouped and work together to anonymize Web-traffic. Each transfer is routed through a random number of group members, disguising the actual sender/receiver. Nevertheless, this concept is generating too much overhead traffic for a VANET scenario.

A simple yet effective way to protect privacy is the use of pseudonyms. While nodes using a single identity can not act anonymously, pseudonyms help to disguise the identity. Hence, a ME uses more than one identity by holding more than one key pair with a connected certificate. The different identities can not be linked to each other, therefore,

messages sent with different pseudonyms can not be linked either. A similar effect can be realized with short living certificates. They are only valid for a short time span before they are replaced with a new key pair and certificate. This however leads to a lot of overhead due to the continuously required connection to the CA to renew certificates.

The most promising approach to realize a privacy protection in future VNs is the use of pseudonyms. This concept can be used for the session-based communication as well as for the dissemination schemes. In addition, it doesn't require additional entities in the system to anonymize messages.

4.6.2 Using Pseudonyms to Increase Unlinkability

Two of the most important privacy mechanisms in VANETs are *anonymity* and *unlinkability*. The use of pseudonyms is one possibility to achieve these mechanisms while sustaining other security mechanisms like authentication. The nodes have to hold several pseudonyms to be able to substitute to a previously unused pseudonym. Several parameters have to be considered when analyzing and developing mechanisms for pseudonym use in VANET scenarios. Examples are the number of pseudonyms, the validity period of a pseudonym, and the influence of pseudonym use on network protocols. A very crucial parameter is the change rate for pseudonyms, how often does a node have to change a pseudonym to achieve a sufficient degree of privacy. This change rate is influenced by several aspects, for example, the communication activity or the node mobility.

The main impact of mobility concerning privacy is that nodes can interact more than once. Depending on speed, region size, and node lifetime the node re-interaction frequency can vary. Any re-interaction should happen preferably using a different pseudonym than before, concealing the re-interaction. However, as the number of re-interactions increases with time and the number of pseudonyms may be limited, a linkable node re-interaction becomes more likely. In case each node has an unlimited number of pseudonyms available the simplest and most effective strategy is to change pseudonyms after each encounter. But if the number of pseudonyms is limited, their use has to be optimized to achieve the highest possible degree of privacy.

Besides the mobility parameters, the rate of communication activities has to be considered for pseudonym changes. Concerning communication, especially the unencrypted messages have to be taken into account, since they most likely leak context information. This is a crucial aspect, since context information helps an eavesdropper to create linkability between events, hence, aggravating the node's privacy. This observation leads to an important requirement: During any open, most likely context leaking communication cycle, no pseudonym change shall occur. In fact, to increase the probability for unlinkability between pseudonyms of the same node a quiet-time (t_{q}) shall be introduced. After a communication process has been finished the node waits for the quiet-time until the next communication process is started, then using a different pseudonym. During the quiet-time the communication neighborhood changes due to the node mobility, therefore, eavesdropping nodes have a lower probability of linking different pseudonyms to the same node. The maximum reasonable quiet-time is defined by the duration of a full neighborhood change. In this case, non of the new nodes has the chance of linking the old and the new pseudonym of the sending node, since the old pseudonym is not known to any of the new neighboring nodes.

Both parameters, node re-interaction time and quiet-time, are analyzed and simulated. The respective results can be applied for the configuration of pseudonym management schemes in VANETs. Moreover, due to a well configured pseudonym management the privacy of MEs can be improved. The results presented in Sec. 4.6 have previously been published in [Eic07d].

4.6.3 Motivation for Node Mobility Analysis in Respect to Node Privacy

To design and evaluate anonymity preserving protocols the concept introduced in [Zug03] can help to evaluate the influence of mobility on the privacy of a ME. The conceptual model recognizes the four major actors and their interactions in a communication system: User, device, action, and location. Moreover, the concept helps to identify weaknesses of a concept in relation to anonymity preservation. This work motivates that *mobility* is a critical parameter regarding privacy and anonymity. Hence, mobility of MEs in a VANET shall be analyzed to identify crucial parameters for the use of pseudonyms to protect the nodes' anonymity.

Especially the node mobility is a crucial parameter for the configuration of pseudonym changes. In a mere static scenario a node has a group of N_{ne} neighbors. Any sent message can be mapped to a specific node with the likelihood of $1/N_{\text{ne}}$. The same is true for the linkability of pseudonyms to a certain node, as long as all nodes in the neighborhood change their pseudonyms simultaneously. However, if only one ME changes its pseudonym, while all other nodes continue to use their previous pseudonym, the new and the previous pseudonym of the respective ME can be linked easily.

As soon as mobility has to be considered, traceability of identities becomes much harder. In this case a newly detected pseudonym could derive from either a pseudonym change of an already known ME or from a new ME entering the neighborhood. This simple example gives a good impression on the influence of mobility on the privacy of nodes. Hence, in static scenarios the privacy of MEs can be calculated using information theory, which is not the case for the mobile scenario. Thus, mobility will have an influence on privacy and pseudonym changes in VANETs.

Practically, node mobility helps to support privacy in a wireless network. It adds an unknown and unpredictable component to the identity analysis problem. On the basis of mobility the isolation of identity changes bound to a distinct ME becomes much harder, especially if the node density is high ($N_{\text{ne}} \geq 5$). The respective influence on node privacy are examined in the course of the following node mobility evaluations. The goal is to maximize the degree of privacy for different VANET scenarios.

4.6.4 Simulation of Node Mobility affecting the Node Re-Interaction and Node Quiet-Time

In Sec. 4.6.2 the terms *node re-interaction* and *quiet-time* have been introduced briefly and their relevance to the challenging pseudonym management has been shown. In this section simulation results for the node re-interaction and the required quiet-time are presented based on the simulation environment presented in App. A. First, the simulation settings are presented briefly.

Simulation Parameters

The mobility analysis has been done based on the OMNeT++ simulator and the simulation models introduced in Sec. A.3. The underlying mobility model was the MGM model with a grid length (d_g) of $d_g = 500$ m. The mobility model updated the node's positions in an interval of 100 ms.

Several different settings for the node speed (v_n) were used, to analyze the influence of speed. The values for v_n were set to $v_n \in \{6, 9, 12, 15, 18, 21, 24\} \cdot$ m/s. The node density was varied to influence the number of neighbors (N_{ne}), the number of nodes (N_n) was set to $N_n \in \{100, 200, 500\}$. A simulation duration of 3600 s was used. The nodes were equipped with ten pseudonyms each, which were used equally often.

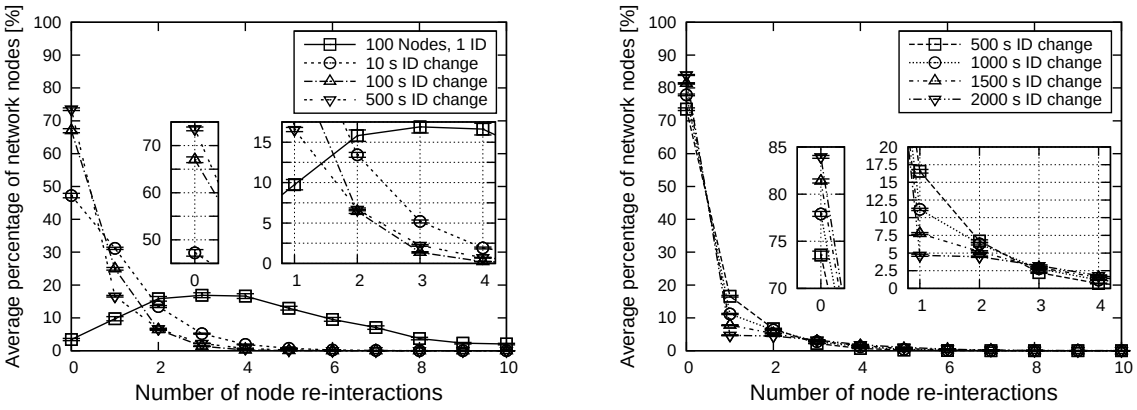
Evaluating the Pseudonym Change Interval in Reference to the Node Re-Interaction

The first results of the simulations are on node re-interaction. Concerning these results one weakness of the simulative approach has to be kept in mind. In reality node movement would not be limited to a defined and very limited area. However, since the primary interest is in the local and short term effects of mobility, the influence of the fixed dimensions is acceptable. One advantage of the limited simulation area is that the results can be interpreted as a worst case scenario, therefore, in reality the number of re-interactions would be similar but most likely much smaller.

The simulation model checks the communication neighborhood of each node after every position update of the MGM model to analyze the node re-interactions. Every newly detected node is logged with the respective pseudonym used at that moment. Using this log data the number of node re-interactions has been analyzed at the end of each simulation run. The reference runs used a scenario where each node had only *one identity*, thus, no pseudonyms where used. In all the other runs ten pseudonyms where used and randomly changed in a certain pseudonym change interval (t_{pc}). The results are depicted in Fig. 4.26. On the x-axis the number of node re-interactions are plotted, while on the y-axis the corresponding percentage of network nodes is shown, which detected the respective number of re-interactions.

In Fig. 4.26(a) the results for the reference model as well as for the pseudonym change intervals (t_{pc}) of 10 s, 100 s, and 500 s are presented. The result for the reference scenario is the flat curve with its maximum at around three node re-interactions, representing the worst case for the given settings. In this scenario many node re-interactions occur, since every interaction after the first encounter can be linked due to the missing pseudonyms. The other graphs (simulations using pseudonyms) have their maximum at zero node re-interactions, representing no or one interaction, but no re-interaction. Depending on t_{pc} the case of one-time re-encounters occurs between 15% and just above 30% of all cases. Node re-interactions of four and above have a very low percentage.

First of all these results prove that introducing pseudonyms reduces the number of detectable re-encounters quite significantly as could be expected. But which change interval t_{pc} is the best? To answer this question, several values for t_{pc} have been simulated (see Fig. 4.26(a) and Fig. 4.26(b)). The results show an interesting and for future implementations helpful result. Not the fast pseudonym change is the best, the 10 s change interval actually



(a) ID change intervals $\leq$ 500 s

(b) ID change intervals $\geq$ 500 s

Figure (4.26) Average node re-interactions for 12 m/s

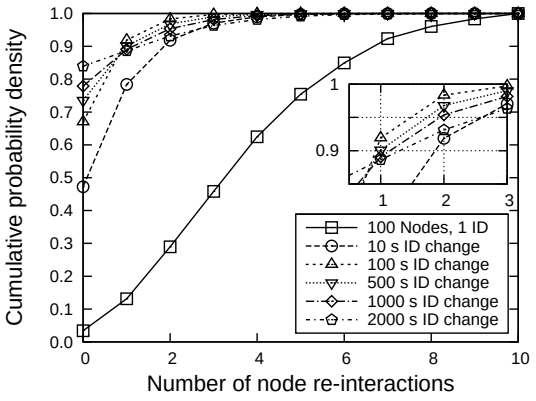
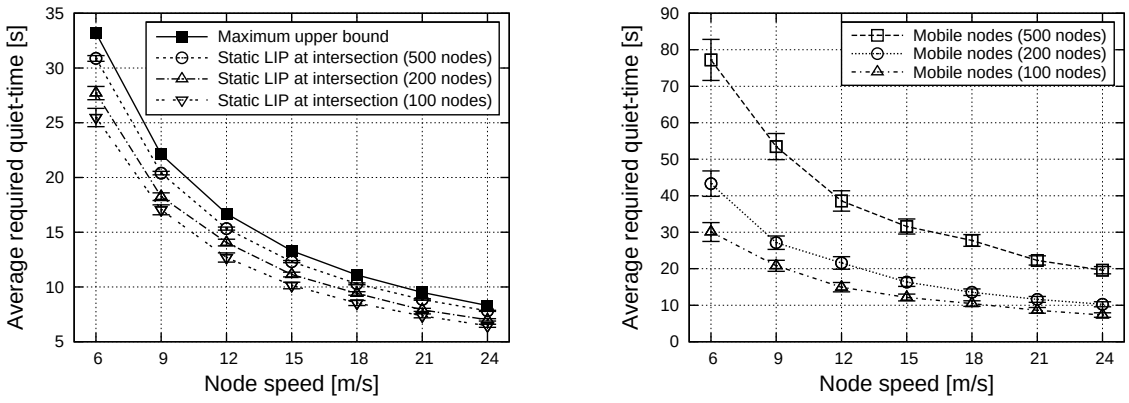


Figure (4.27) Re-evaluation of Fig. 4.26 using a cumulative PDF

performs rather poorly. The reason for this is the average node interaction time (t_{int}). The length of t_{int} is specific to the mobility model and the node speed, in our setting it amounts to $t_{\text{int}} = 13$ s. The change interval should be bigger than t_{int}, therefore, the 10 s can not be the best parameter value. A second scenario specific parameter which is helpful to identify the best value for t_{pc} is the average time duration elapsing between a re-interaction (t_{ri}). For the given scenario t_{ri} amounts on average to 583 s with a 95% confidence of 10 s. Therefore, t_{pc} should be just smaller than t_{ri} to achieve an optimal result for the pseudonym change. Looking at the results plotted in Fig. 4.26(a) this claim seems to hold, the plot for $t_{\text{pc}} = 500$ s is decreasing faster. However, including larger values for t_{pc} into the analysis seems to counter the claim. In Fig. 4.26(b) the graphs for values of t_{pc} larger than t_{ri} are presented. They decline even faster.

However, the decline of the graphs is somewhat misleading. Larger values for t_{pc} definitely lead to higher probabilities for the low re-interaction values, but they also cause a higher probability for the re-interaction values of three and above. This can not be seen in Fig. 4.26 since the difference is very small. Re-plotting the same results to a cumulative



(a) Static node at intersection

(b) Mobile nodes only

Figure (4.28) Average required quiet-time t_q for a full neighborhood change

Probability Density Function (PDF) results to Fig. 4.27, where the effect is more visible. The larger t_{pc} gets, the higher is the starting value of the graph, but the inclination of the plot decreases at the same time. In Fig. 4.27 can be seen that $t_{pc} = 100$ s has the biggest incline, even slightly better than $t_{pc} = 500$ s which would be closer to t_{ri}. Most likely the selected scenario was too small to show the desired effect better. Even though $t_{ri} = 583$ s, still many samples of t_{ri} are smaller than this average, since the uncertainty of the random processes in a small scenario is rather big. That's why $t_{pc} = 100$ s leads to a slightly better performance in these results.

Simulation of the Required Node Quiet-Time t_q Before a Pseudonym Change

In the second group of simulations the quiet-time (t_q) has been evaluated. In these simulations t_q is defined as follows: The nodes of the simulation make a snapshot of their neighborhood at a random point in time. Then the simulator measures the time t_q until *all* neighboring nodes of the snapshot have left the radio range of the node evaluating its neighborhood.

The simulation results depicted in Fig. 4.28(a) have been measured using a static node (N_{static}) analyzing its neighborhood. The node N_{static} was located at an intersection in the simulation scenario while the other nodes where moving at one of the plotted node speeds. It can be seen that depending on v_n and N_n the value for t_q changes. In this rather simple scenario with N_{static} measuring t_q, an upper bound can be calculated. Since $d_r < d_g$, the maximum length any node has to move in the range of N_{static} is $2 \cdot d_r$. Therefore, regarding the node speed (v_n), the upper bound for t_q can be calculated: $\widehat{t_q} = 2d_r/v_n$. In Fig. 4.28(a) the upper bound for the given scenario is presented with the squared, solid points. In Fig. 4.28(b) the results for an all mobile scenario are given. In this case no upper bound can be calculated, since two nodes could move along the same path throughout the whole simulation time. It can be seen that in an all mobile scenario t_q is much longer, especially for simulation scenarios with many nodes moving at a rather low speed. But for

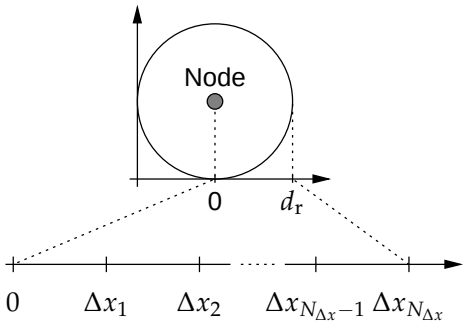


Figure (4.29) Separation of radio range d_r into distance increments Δx_i for the analytical calculation

common vehicle speeds in urban areas t_q is much shorter than one minute and can be as low as 10 s.

Overall, the results given in Fig. 4.28(a) and Fig. 4.28(b) can be used to identify parameters for a pseudonym change algorithm used in VANETs. To increase the effects of pseudonyms, a minimum value of $t_q \geq 15$ s should be set.

4.6.5 Analytical calculation of Node Quiet-Time

In this section a method to analytically calculate the quiet-time (t_q) for the static node scenario is presented. First, the parameters and the setup for the analysis will be described. In the second step the equations of the analysis are given. The analysis helps to set t_q for static moments of a ME, for example at an intersection.

The setup for the calculation is shown in Fig. 4.29. The calculation is based on the node density (N_n), the radio range (d_r), and the node speed (v_n). The node density can be transferred to an average N_{ne} that are within d_r (see Sec. A.4 [pp. 174] for details). The length of d_r is segmented into equally long distance increments Δx_i (see Fig. 4.29). The index parameter i equals the number of increments taken into account at the respective calculation step. The main principle of the calculation is the use of the binomial coefficient (Eqn. (4.14)) and its variants [BSMM99].

$$\left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{(n-k)!k!} \tag{4.14}$$

The analytical way of determining t_q uses the statistical characteristics of the mobility model, to determine an upper bound for a given N_{ne}. Due to the symmetry of the mobility model, the intersection with four directions can be reduced to one direction only. Since the nodes are distributed equally over the simulation area all increments Δx_i have the same probability containing a node at any random point in time. The upper bound of t_q is only influenced by the node having the longest remaining distance within the radio range. For example having only one node within radio range being located within increment Δx_i allows for two possibilities: Either the node has to travel the i increments towards the intersection plus the distance d_r to leave the radio range or it only has to travel the remaining distance

$d_\mathrm{r} - i \cdot \Delta x$ to the edge of the radio-range. Only the first case is relevant for the calculation, since it is setting the longer time.

The number of possible combinations to place N_ne nodes within the i distance increments closest to the intersection is defined by Eqn. (4.15). The maximum number of possible combinations including all distance increments is an important parameter to be able to determine the combination probabilities. They can be calculated using Eqn. (4.15) and setting $i = N_{\Delta x}$.

$$C_i\left(N_\mathrm{ne}, i\right) = \left(\begin{array}{c} N_\mathrm{ne} + i - 1 \\ N_\mathrm{ne} \end{array} \right) \tag{4.15}$$

$$C_\mathrm{max}\left(N_\mathrm{ne},\, N_{\Delta x}\right) = \left(\begin{array}{c} N_\mathrm{ne} + N_{\Delta x} - 1 \\ N_\mathrm{ne} \end{array} \right) \tag{4.16}$$

The probability to have at least one node placed within the i-th distance increment and all other nodes in an increment closer to the intersection is defined by Eqn. (4.17). The sum of all probabilities p_i has to fulfill Eqn. (4.18). Finally, the upper bound of the quiet-time for a static node is given by Eqn. (4.19).

$$p_i = \frac{C_i - C_{i-1}}{C_\mathrm{max}} \tag{4.17}$$

$$1 = \sum_{i=1}^{N_{\Delta x}} p_i \tag{4.18}$$

$$\widehat{t_\mathrm{q}}\left(N_\mathrm{ne},\, N_{\Delta x}\right) = \frac{d_\mathrm{r}}{v_\mathrm{n}} + \sum_{i=1}^{N_{\Delta x}} p_i \cdot \frac{i \cdot \Delta x}{v_\mathrm{n}} \tag{4.19}$$

Using Eqn. (4.19) to calculate the upper bound of t_q for a given N_ne generates the results shown in Fig. 4.30. To be able to compare simulation and analysis the simulation result for the scenario with 100 nodes is plotted again. A node density of 100 nodes is just above the neighbor density of one average neighbor. Therefore, the analysis using Eqn. (4.19) gives a good upper bound for t_q.

4.6.6 Using Node Re-Interaction and Quiet-Time for Pseudonym Management

In a final step the simulation results shall be used to define strategies for the change of pseudonyms in VANETs. The two parameters node re-interaction time and quiet-time, which are mainly influenced by the characteristics of the node mobility, have to be considered to optimize the privacy effects achieved by pseudonyms. The better the pseudonym change interval is adapted to the node re-interaction interval, defined by the mobility, the higher is the degree of unlinkability between different pseudonyms of a node. To optimize the results presented in Fig. 4.26 other scenarios (for example different size or including a driver behavior model) should be considered.

The second important parameter, the quiet-time (t_q), can be set using the results presented in Fig. 4.28 or using the analytical way presented in Eqn. 4.19. The results can either be used to define a fixed value for t_q, resulting in a compromise for all possible

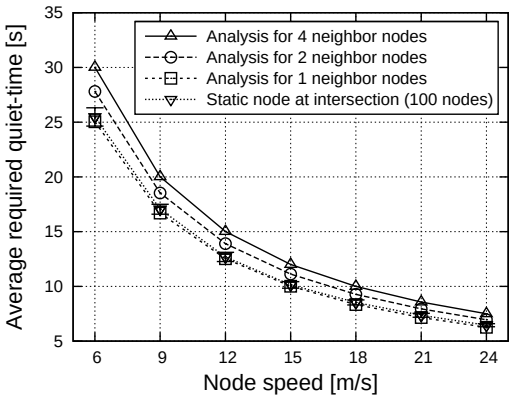


Figure (4.30) Upper bound for the required quiet-time $\widehat{t_q}$ using statistical analysis

scenarios. However, this would lead to a quiet-time that is not satisfying the scenario in most cases. Either the time is too short or too long. Moreover the results can be used within an algorithm used to change the pseudonyms of a node. In this case the algorithm can determine or estimate the parameters *speed* and *node density* and set t_q accordingly. This would achieve the highest level of unlinkability in any scenario.

A method to measure the degree of anonymity has been suggested in [DSCP02] and [SD02]. Based on the concept of information entropy [Sha48] the degree of anonymity for a distinct ME in a given scenario can be calculated. An observer listens to communication and assigns probability values to MEs, which determine the likelihood that a distinct ME sent a certain message. This basic concept can be used to realize a pseudonym management system for VANET nodes.

Such a system needs to monitor several parameters and evaluate the current degree of anonymity. Based on this evaluation system parameters like t_q can be adapted to reach the highest possible degree of anonymity.

4.7 Conclusions

In this chapter security and privacy aspects for a future VN have been discussed and evaluated. Security and privacy are considered key elements if VNs shall be successful and widely adopted in the future. The most important building block in a secured VN is the trust architecture. Before discussing the semi-centralized trust architecture applicable in VN environments the meaning and mapping of trust in technical systems has been outlined. To maintain the security of the PKI-based trust architecture certificate revocation is a very crucial mechanism. Different concepts for revocation and validation have been reviewed and their performance in the distributed network part of a VN evaluated. The results support the use of a PKI even for the distributed VANET part of VNs. This is an important result, since one trust environment solution should fit all settings in a VN.

Several other security concepts have been presented, based on this trust architecture. The infrastructure supported telematics platform services rely on heterogeneous communication links, which require a specific security setup. In addition, to be able to support services from

multiple service providers the security concept needs to be adaptable and fully integrated into the core system architecture. Both requirements have been taken into account in the outlined platform security concept. To realize the security mechanisms in a reliable way and have a trustworthy hard- and software environment a tamper-proof hardware device is used. Using this platform concept multiple services can be realized, having individual security needs.

Security for V2V messages needs a specified packet format. A possible format has been suggested and the generated overhead evaluated. This evaluation helps to select the appropriate security mechanisms for IVC, Elliptic Curve Cryptography (ECC). The suggested message format in combination with ECC can provide the required features like message authenticity and integrity, however, the trust level of the disseminated content can not be attached in this way. The suggested protocol CoRS realizes this content reputation feature using threshold signatures. The protocol relies on node cooperation to generate reputation information for message contents. The evaluation results of CoRS prove the feasibility of the approach. The additional delays introduced by the content reputation protocol are in the range of few seconds, depending on the node density. The analytical calculation of signature share collisions can be used to configure the protocol and minimize share collisions. CoRS provides an important building block for future VNs, since the communication of trust levels for content is an important requirement for many IVC services.

Finally, privacy issues for VANET scenarios have been presented. As pseudonyms will be an important part of a VN trust architecture, since they are compatible with the suggested trust architecture, their usage needs to be investigated. The node mobility influences pseudonym changes significantly which has been shown by the presented simulation results. Parameters like node re-interaction time and quiet-time need to be carefully configured to maximize the anonymity of nodes. A concept of doing this has been outlined in this thesis.

The results on security and privacy presented in this chapter are valuable results when realizing VNs in the future. With the PKI a trust architecture has been identified, which can serve both the centralized and the decentralized network part of VN scenarios. In combination with the communication mechanisms presented in Chap. 3 the security mechanisms can be integrated in a full-scale architecture concept. This architecture concept as well as organizational aspects of the trust architecture are introduced and discussed in the next chapter.

5

System Architecture for Vehicular Networks: Entities and Interactions

The introduction of a Vehicular Network (VN) in reality requires the definition of a complete architecture, including aspects like communication, security, and management. Such an architecture needs to be complete while still being very adaptable for future trends and enhancements. Therefore, a modular concept including modules for security and communication, applicable in centralized as well as decentralized settings seems to be the most promising approach.

Whereas a distinct mechanism, such as a message dissemination protocol, needs to meet only a few requirements, a system architecture for VNs needs to comply with many requirements from diverse contexts. Security and communication aspects play an important yet not an exclusive role in determining the cornerstones for such an architecture. Also organizational aspects, management issues, and practical implementation-related matters influence the layout for such a system architecture significantly. That is why the definition of one standardized architecture is extremely difficult. However, defining a modular concept allows to continuously adapt and improve such an architecture, keeping it compliant to the latest requirements.

In this chapter the results of Chap. 3 and Chap. 4 are used and integrated into the suggested system architecture for VNs. This clearly shows the modular setup of the architecture components. In addition, a potential security setup for Mobile Entities (MEs) is introduced in great detail, differentiating the integration of security and privacy mechanisms. Further, the interactions of modules and system components are explained. Several use case examples point out the versatile application areas of the modular architecture concept.

The chapter on the system architecture for VNs is structured as follows. In Sec. 5.1 an overview on related work is presented. The configuration of the Backend architecture and the main component interactions are detailed in Sec. 5.2. The architecture of MEs and the integration of security and communication modules is presented in Sec. 5.3. Finally, in Sec. 5.4 the usage of the described architecture is explained with some examples. The chapter closes with conclusions in Sec. 5.5.

5.1 Overview on System Architecture Concepts for Vehicular Networks

Since VNs combine many different features an architecture concept is required. Due to the increasing interest in Vehicular Ad Hoc Network (VANET) research several different architecture concepts have been proposed in the literature. The most important ones are briefly presented in the following section.

With CarNet, one of the first basic architecture concepts for large Mobile Ad Hoc Network (MANET) scenarios in the vehicular context has been presented [MJK+00]. The system uses a geographic forwarding approach to disseminate messages to given locations.

A much more elaborate architecture concept for telematics services has been introduced in [BBC+02]. The concept relies on open standards such as web services and can be seen as a middleware for many different telematics services. Nevertheless, the concept does not include Inter-Vehicle Communication (IVC) and its security. The Vehicular Communication Platform has been outlined in [CS02]. It is an architecture concept specifically considering IVC for safety messages. The concept takes the most important requirements into account, however, security features have not been considered at all. Moreover, the architecture does not include any considerations for platform-based services.

A very important challenge for the in-vehicle architecture is the integration and parallel usage of multiple different communication technologies. In [KBS+01, KVS02] service and system architecture issues in the vehicular environment are discussed. The authors present the heterogeneity of the scenario for both the different service applications and the required communication technologies. The integration of different access technologies is very important to realize VNs to their full extend. Hence, the presented communication gateway architecture is a valuable building block for an architecture concept. It manages the available communication devices of a ME and selects the best suited device for each communication session, regarding the availability of each technology individually.

The architecture for the Virtual City Portal introduced in [BKK+03] is a service portal providing infotainment applications from a server infrastructure to the vehicular nodes. In the paper a detailed description of the system setup for both the client and server architecture is presented and evaluated based on a prototypical implementation. Even though IVC and security aspects are not taken into account, the paper gives an overview on the complexity of such a system. The realization of such an architecture, should be done with a hardware independent concept as it is possible with the Open Services Gateway Initiative (OSGi) concept. In [GPZW04, GPZ04] the requirements for context-based services are discussed and mapped to the OSGi system layout. It proves to be suited to realize service oriented architecture concepts very well. Thereby OSGi provides hardware independence, level-based security integration, and parallel service provisioning and usage from different providers. The discussed architecture in the paper is targeted at smart home scenarios, however, it can be transferred to the vehicular context. The author use an ontology-based approach to realize a context model for the context-based services. Security as well as the distinct discussion of client and Backend architecture setups are missing in the paper.

A similar approach has been discussed in [ZWH04], however, as a starting point the automotive context is used. In the paper a context-aware service architecture based on

OSGi is suggested for telematics services. The authors differentiate the Server and the client system structures and give a detailed overview on the software structure of the OSGi service gateway inside the vehicle. The protocols Hypertext Transfer Protocol (HTTP) and Session Initiation Protocol (SIP) are suggested for communication with the Backend servers. Security is not discussed in detail, however, the Java security model in combination with Secure Socket Layer (SSL) is suggested to be used to realize communication security. A service-oriented middleware based on OSGi for automotive telematics services has been introduced in [GWP06]. The authors consider the crucial requirements security, privacy, usability, and reliability to design an OSGi-based middleware concept which can be used to ease application integration into VN clients. The concept is implementation driven, yet it leaves out security considerations for platform security as well as IVC integration.

A stack-based VANET protocol architecture especially suited for IVC has been presented in [FTTM+05]. The authors chose a conventional stack approach and integrated scenario specific components such as layer interaction as well as a single-hop and multihop communication layer. The authors outline a basic architecture concept not touching security aspects, platform services, or the integration of application specific protocols. Nevertheless, the results can be used as a starting point for future architecture discussions.

Besides the different communication architecture concepts, also several security architecture ideas have been suggested in the literature. In [DGL+02] a system architecture for data protection and privacy applied to automotive telematics has been introduced. The concept uses the so-called data protection manager, which is part of both the client and the server architecture, in order to handle application data with the respective security and privacy policies. In addition the concept relies on the platform protection manager, which provides the security mechanism to the system and checks its integrity.

An architecture concept primarily concentrating on security and privacy aspects in VANETs has been outlined in [PNM06]. The concept is designed for three types of services, warning messages, alarm signal, and value-added services. Similar to the communication concept discussed in [FTTM+05] a differentiation between single- and multihop has been made. The concept does not follow a strictly layered approach and does not differentiate between different scenario entities. The authors combined existing ideas to define the architecture concept, however, the suggested system does not take into account the specific requirements of both IVC and platform services. The specified block diagram of the architecture does not fully explain the interactions of the different mechanisms and leaves out platform security issues.

A security architecture concept for VANETs based on the existence of different stakeholders (for example Original Equipment Manufacturers (OEMs) or public authorities) and their responsibilities has been outlined and evaluated from different viewpoints in [GFL+07]. The integration of security mechanisms into the protocol stack is outlined in the paper, showing a general concept which can be further adapted to future challenges and vulnerabilities. The analysis of the different viewpoints helps to identify certain security functionalities as important building blocks. These are digital signing, use of pseudonyms, tamper resistance, secure positioning, and communication security. The in-vehicle trust model and the concept of the context mix model providing location privacy is discussed in greater detail in [Ger07].

Overall can be summarized that several different architecture concepts have been discussed in the literature. The aspects of security, communication, and distributed and centralized services have been discussed in several publications. Nevertheless most papers concentrate on very distinct aspects, leaving out holistic aspects. In this chapter an architecture concept is discussed, including communication, security, and privacy aspects using and including the results of the previous chapters.

5.2 Backend Architecture Setup of Vehicular Networks

Depending on the functionalities of a VN the Backend architecture needs to support different features. Basically a VANET could be operated without any support architecture whatsoever, but practically a Backend will always exist, even if only providing the trust basis. More likely is the use of a Backend with several features, hence, its setup needs to fulfill several requirements.

5.2.1 Features Provided by the Backend Architecture

A Backend architecture can provide a great variety of features to VN nodes. These range from management to service provisioning functionalities. A VN is a very heterogeneous network with many different stakeholders, however, the operator of such a network uses the Backend to control the network functionalities.

The most important feature located in the Backend of future VNs is the main component of the system's trust environment, the Public Key Infrastructure (PKI). Closely connected to this feature is the user and subscription management. A server infrastructure in the Backend network provides these features and issues certificates for new users in combination with a Registration Authority (RA) and a Certificate Authority (CA). Every VN using any form of trust environment needs a user management to provide the important feature of certificate revocation to maintain the system's security. Moreover, a platform operator will want to have a certain control and supervision on the activities and the members using the system.

The user management and subscription handling is also important from another point of view, telematics services offered by a service provider will be provisioned by server nodes located in the Backend. To ease system management on the one hand and usability on the other hand, service providers will have a contractual relationship to the platform operator. Hence, they can use the user data held by the platform management. Thus, the data needs to be maintained in one position only. Further, the users do not need to hold individual contracts with the operator and the service providers. Moreover, a transparent service consumption involving several stakeholders can be realized due to this user subscription policy (see Sec. 5.4.1 for the Single Sign-On (SSO) example).

For the billing of services the sole responsibility for user data and its administration at the operator is also beneficial. Having an extended user and subscription management in the Backend allows for the realization of a single invoicing procedure. Thus, no matter how many service providers are involved in the provisioning of applications for a single user, only one single invoice is issued. This increases the usability of the system for the users and makes the billing procedures much simpler.

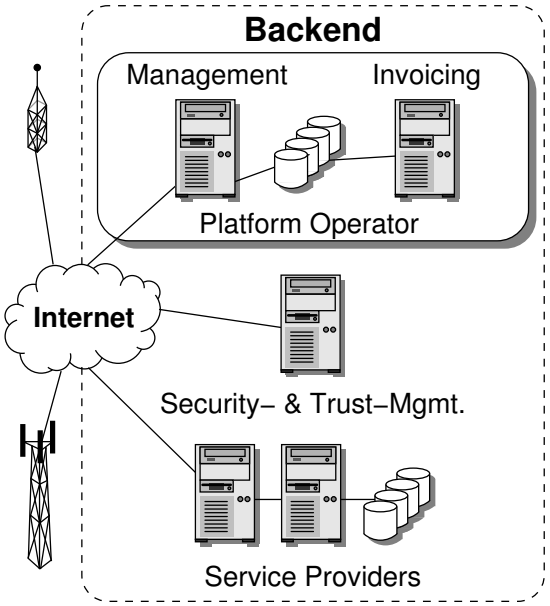


Figure (5.1) Components of the Backend architecture in a Vehicular Network

Besides these more or less administrative features, the Backend is required to realize any kind of server-based application or services relying on the Internet. The Backend can for example be used to provide Internet connectivity as a service to the MEs. In addition, all platform-based services for telematics and infotainment require the Backend architecture. They will make up a great deal of commercial applications needed to finance such a VN in the future.

5.2.2 Components of the Backend Architecture

The system setup of the Backend architecture consists of several components. Three main components can be differentiated: The operator servers, the security and trust servers, and the servers of the service providers. A general setup of these components is shown in Fig. 5.1. The components of the Backend are all connected using the Internet or a similar network and their functionalities are distributed as follows:

Platform operator: A VN using a Backend with services requires a user and platform management. This is done by a platform operator. The most important task of the operator is to handle the user data and the related service subscriptions. It closely cooperates with the trust management to provide security credentials to its users. Moreover, the operator itself or a closely related stakeholder will do the invoicing of the platform. Hence, all service consumptions are logged and accounted by or at least in close cooperation with the platform operator. Further, the operator has a contractual relationship to one or several service providers. A service provider gets access to the users of the operator and in return offers its services to the users. The providers do not have to manage their own user database, however, they can solely rely on the user

data provided by the operator. In a practical VN implementation not necessarily one operator needs to handle all users. Nevertheless, several operators can run in parallel. The users can interact as long as the trust environments are interconnected through cross-certification, establishing an extended circle-of-trust.

Trust management: Closely connected to the platform operation is the trust environment and its management. In most settings a RA will be used as a mediator between the platform operator and the CA. It handles all certificate and revocation requests and bundles them to be handled by the CA. The CA will provide the certificates and the up-to-date Certificate Status Information (CSI) for the platform. Further it will provide cross-certified certificates to realize trust relations to other CAs or trust environments. All members of the circle-of-trust, the platform operator, the service providers, and the users need to have a valid trust credential issued by the trust management to be able to interact with other platform actors.

Service providers: The functionalities experienced by the users are provided by the service providers. Hence, they are a crucial part of a VN. Their server infrastructure will be located in the Backend of the system. They are integrated in the platform on a contractual and technical basis. The providers offer their services through the platform of the operator, thus, they can not directly offer services to the users. However, as multiple platform operators can co-exist in a VN, a service provider can offer its applications to multiple platform operators and their users. They only need to adapt the software to the requirements of the respective MEs and their On-Board Units (OBUs).

Besides these three main components the Backend needs to rely on a network infrastructure which will most likely be the Internet. In addition, the connectivity to the MEs needs to be realized with both a IVC technology as well as a General Packet Radio Service (GPRS)/Universal Mobile Telecommunications Standard (UMTS) network provider.

The GST High Level Architecture as a Possible Realization

The European research project Global System for Telematics (GST) has defined a reference architecture for platform services in a VN. The main goal of the project was to suggest an architecture concept which is general enough to be acceptable by all players, the OEMs, the service providers, and the network operators. Further, the market access especially for providers should be eased. An overview on the architecture and the main component interactions is given in Fig. 5.2.

Similar to the Backend architecture setup three main components are differentiated in the GST high level architecture [VVM+04, FCD06], the Control Center (CC), the Client System (CS), and the Service Center (SC). The CC is very similar to the platform operator. It handles the subscription, the invoicing, and the general system management. Due to the fact that OSGi [MK01] is suggested as main operation environment within GST [FCD06, pp. 74], the CC is also capable of a CS management. Therefore, it can update the software components of CSs based on a remote update procedure. The SCs are the providers in the GST high level architecture. They operate the services and provide service applications

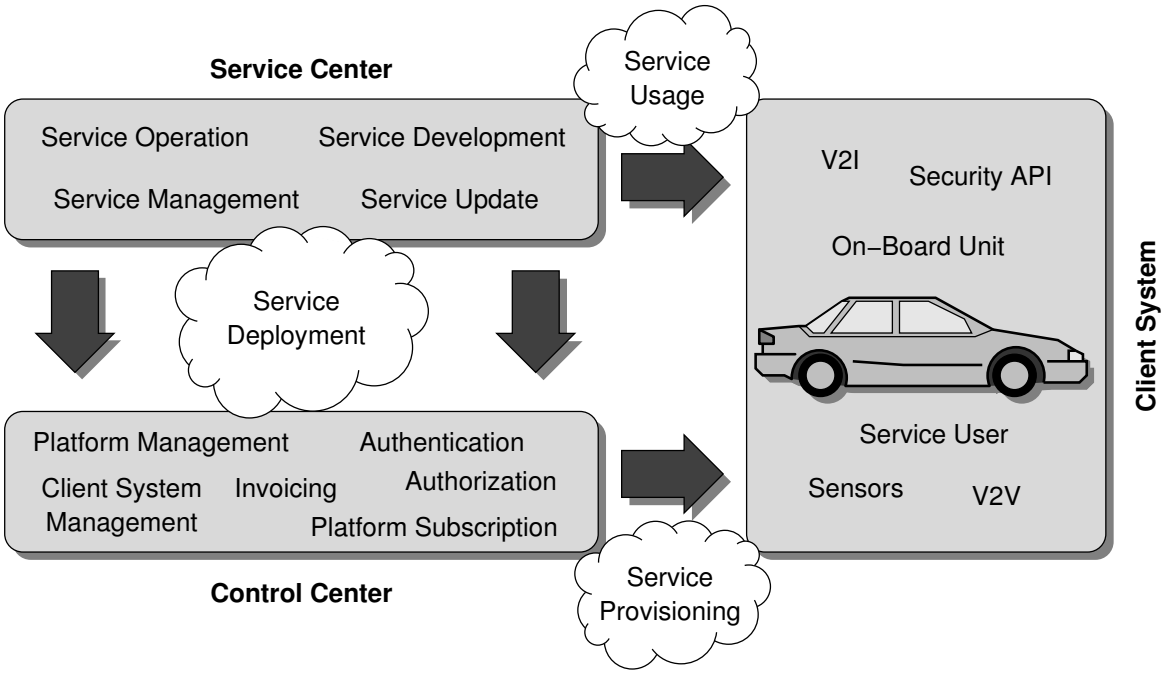


Figure (5.2) Overview on the GST high level architecture and its component interactions

to the CSs. The CS is any compatible mobile terminal, which can be an OBU or a mobile terminal.

The interactions between the components (also depicted in Fig. 5.3) show their relations. Concerning the service applications, three actions are relevant: Service deployment, service provisioning, and service usage. A service application will be developed and provided by a SC. The service will be deployed, hence, made available on the platform system, from the SC to the CC. The CC registers the service and its requirements on the CS in a database. After the deployment the service can be subscribed by the users. After a service subscription the service application is provisioned to the CS from the CC, which includes the software download and its configuration. The service consumption is then directly handled between the SC and the CS, eventually using a distributed authentication mechanism like SSO in cooperation with the CC (see Sec. 5.4.1 for details).

The full GST high level architecture and the component relations are shown in Fig. 5.3. Within GST the security is also connected directly to the CC, which is the system management component. The end user interacts with a CS which is in most cases an OBU inside a vehicle, however, other mobile terminals or mobile phones with extended functionality can also be a compliant CS.

5.2.3 General Trust Architecture Setup

Based on the performance analysis of PKI revocation mechanisms presented in Sec. 4.2.3 a recommendation for the design of the trust architecture setup is given. Practically three aspects need to be taken into account when designing the architecture: Certificate sizes, CSI distribution in the VANET, and the organizational structure of the architecture.

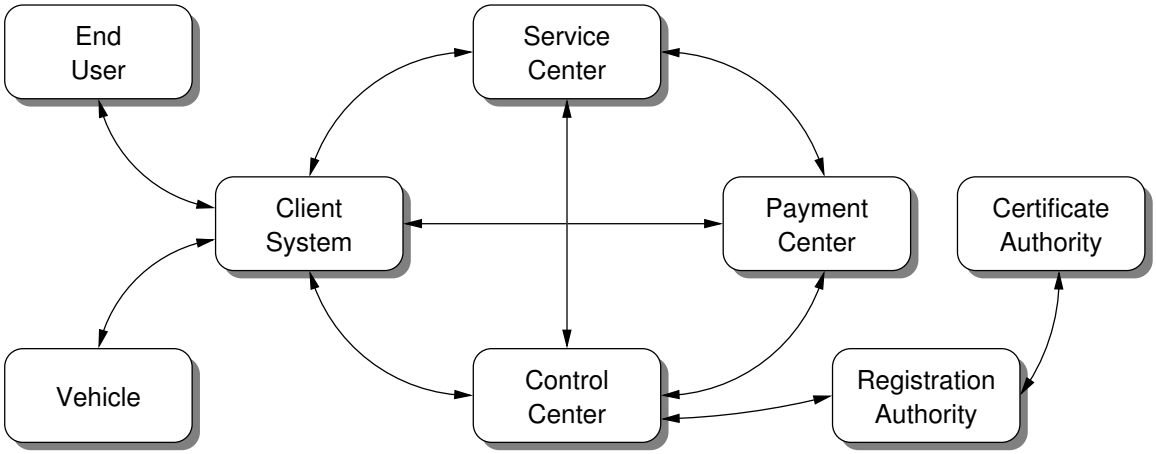


Figure (5.3) Interactions between components of the GST high level architecture

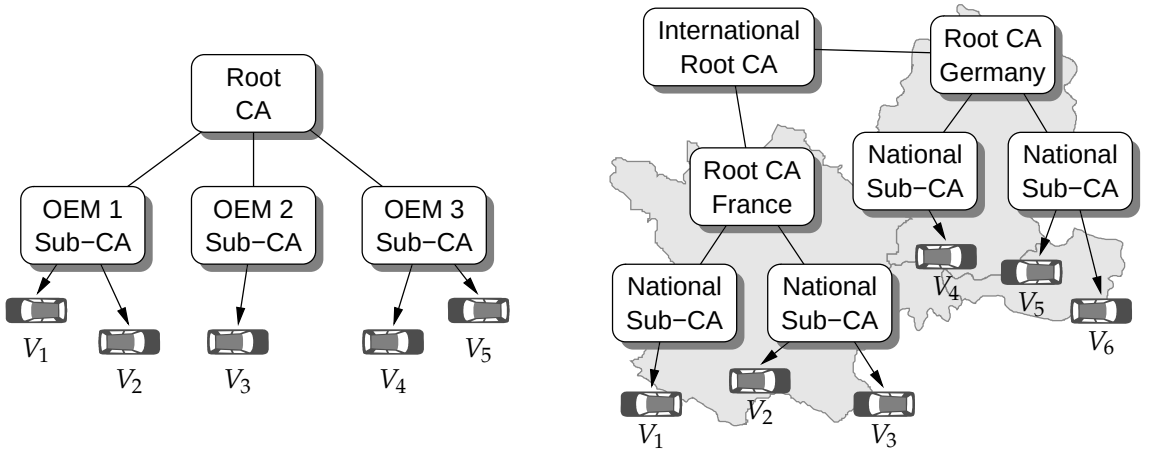
Certificates and Revocation

The results on data and processing overhead, added to messages by cryptography, presented in Sec. 4.4 clearly showed that Elliptic Curve Cryptography (ECC) is an optimal choice. It allows small certificate sizes, therefore, the data overhead is reduced to a minimum. To benefit from this overhead reduction a certificate format has to be defined which requires a minimal size while containing the required information. As a potential candidate the mCert format has been suggested in [ER06a, ER06b]. In general a certificate needs to contain the public key, a key ID, the validity period or validation targets, and the CA signature. In addition, certificates can contain additional attributes, specifying special rights or group memberships.

To realize the required revocation mechanism either a conventional Certificate Revocation List (CRL)-based approach with delta-CRLs or the NOVOMODO ticket-based validation approach is suggested. The CRL-based approach is better suited for a VANET scenario from an organizational point of view. In this case only the CSI needs to be distributed once every day, whereas the ticket-based approach requires a daily ticket distribution to all nodes individually. In an architecture where most MEs also use platform-based services the ticket-based certificate validation is to be preferred. In this case the tickets can be exchanged via the cell-based networks easily, reducing the CSI-related data transfer to a minimum.

Different PKI Organization Structures

The organization of the PKI itself plays an important role for the security, management, complexity, and performance of the whole system. An overview and discussion of various widely deployed PKI organizational structures has been presented in [Per99]. The paper helps to understand the challenges of choosing a suitable structure and rates existing approaches in terms of security and scalability. These results can be used as a basis to define a PKI structure for VNs.



(a) OEM centric PKI

(b) Country centric PKI

Figure (5.4) Two possible public key infrastructure organization concepts for vehicular networks

Depending on the range of services, the interests of OEMs and national governments, different organizational structures can be applied for the PKI used in a VN. Three different structures appear to be most suitable: A commercial solution, an OEM-centric solution, and a governmental solution on national level. A general issue for any possible solution is the financing, rather than technical issues. However, the three main approaches have to face different technical and organizational challenges.

The commercial solution is the easiest to realize, as long as the overall business case for VNs holds. In this case any of the large commercial CAs could run the PKI. However, this approach has some drawbacks for the users, the OEMs as well as the national agencies. If specific features or services shall be provided in the VN, for example, Electronic License Plates (ELPs) or automatic software updates, the CA needs to be closely incorporated into the provisioning process and needs to hold sensitive data. In case of the ELP, a third-party namely the commercial CA, will also have to know the mapping between user, vehicle, and ELP. This can lead to privacy breaches for the user and lead to organizational difficulties for the national agencies issuing license plates. In case of specific OEM-services, the OEM will have to rely on the external CA and most likely have to pay fees for the certification process. These drawbacks make a commercial solution somewhat unrealistic.

The OEM-centric PKI solution (see Fig. 5.4(a)) is a feasible approach especially for the starting years of vehicular networks. Any OEM starting to install communication technology, which uses PKI-based security functionality, into their vehicles, can set up a PKI and at least certify the self-produced vehicles. The main advantage for the OEM is the identifiability of vehicles. The certified key in the vehicle identifies the vehicle exactly and allows for logging the vehicle's history. In addition, automatic software updates as well as secured wireless access to the vehicle by a service technician become feasible.

As soon as many vehicles are equipped with communication technology, the OEMs can either cross-certify their CA-certificates or install a root-CA and expand the PKI to a full

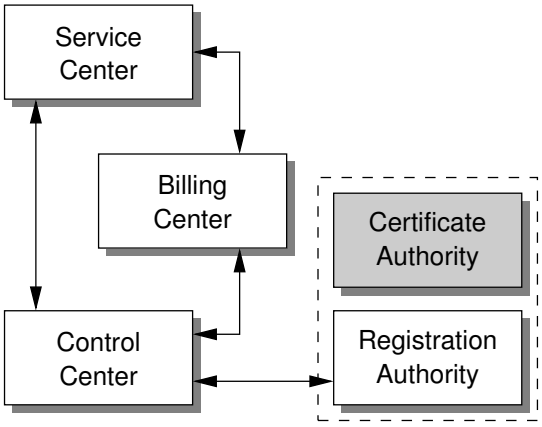


Figure (5.5) Interactions between the Backend components using the GST setup

vehicular network across OEM-borders. With this setup vehicles can identify any member of the network using the full certificate chain. Renewal of certificates can be done during the regular service intervals of the vehicles at any OEM garage. However, the introduction of the ELP encounters the same drawbacks also present in the commercial approach.

The third approach, organizing a PKI on national level, is depicted in Fig. 5.4(b). In this setup a PKI-hierarchy is introduced in every country. These national PKIs are cross-certified by an international root-CA (in Europe for example by an agency of the European Union (EU)). This approach is especially beneficial for the introduction of ELPs, since the issue process of licenses is handled by one institution only. In addition, the provisioning of trustworthy safety services is very much in the interest of national governments and the EU. Thus, by providing a PKI for VNs they could push encourage the development of VN technology, while also pursuing own goals.

5.2.4 Interactions between Backend Components

Since the Backend components are not acting on their own behalf, the interactions to other components are crucial. Hence, the most important interactions are explained in the following.

The CC is the most important component in the interactions diagram (see Fig. 5.5). Everything starts and ends with the CC. Many of its tasks can be handled without the help of other components. One exception is the management of certificates for the security mechanisms on the platform. A CA is required to issue certificates. In most cases the operator of the CC will not operate a CA as well. However, if the CC operator does provide its own CA it will still be an extra component. Hence, any existing CA will be used via a RA.

Since the CC owns and manages the end-user data, each SC needs to interact with the CC. The operator of a SC needs a contract with one or more CC operators to offer the services to the end-users. Thus, besides the technical connection, also a contractual connection needs to be established between CC and SC.

The Billing Center (BC) needs to interact with both, the CC and the SC. If an end-user subscribes to a commercial service the SC interacts with the CC to check the validity of the end-user. If the validity is confirmed the SC can subscribe the end-user and charge the respective costs using the BC. The CC provides the relevant end-user data for the invoicing process to the BC.

5.3 Mobile Entity Architecture Setup

The most important components in a VN are the MEs, which make up the core of the network. In general a ME in a VN is a vehicle equipped with the required communication and processing capability called OBU. Nevertheless, other types of MEs are conceivable, for example mobile phones or after market terminals, similar to the very popular navigation systems. Independent of the realization, a ME needs to provide a multitude of features and capabilities to fulfill the needs of the architecture concept.

5.3.1 Features of the Mobile Entity

The ME is the component closest to the user, hence, it needs to have all features the user wants to have. This includes both distributed vehicle-to-vehicle (V2V) services and centralized portal applications. To be able to provide IVC and portal services the ME needs to be equipped with several different communication technologies. At least one type of infrastructure-based communication technology (for example GPRS/UMTS) and one IVC device needs to be available. The mobile client will support three types of communication relations: vehicle-to-vehicle (V2V) communication, vehicle-to-infrastructure (V2I) communication, and vehicle-to-backend (V2B) communication.

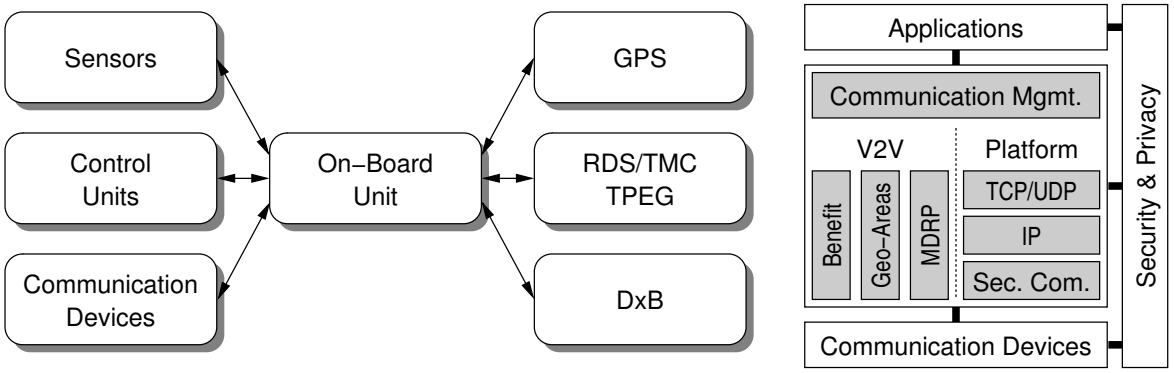
An important feature for the ME is the user interaction with a Human-Machine Interface (HMI). Especially an OBU inside a vehicle needs to provide visual as well as auditory user interaction. The OBU will practically be integrated in the on-board computer providing all types of information and entertainment features.

A ME integrated into a vehicle will have access to the on-board sensors of the vehicle. This is an important feature to collect status information especially for the IVC services. Due to the integration into the vehicle architecture, the OBU will be able to access all types of sensors, communication devices, and control units. Thus, it can aggregate information from different sources and use it for the event detection and as context information.

To be fully integrated into the architecture concept the ME needs to be able to handle security credentials. Further, the entity itself needs to be secured using soft and hardware security components. The security mechanisms need to be usable for broadcast, point-to-point (P2P), and end-to-end (E2E) communication, hence, all different communication relations that typically occur in a VN.

5.3.2 General Mobile Entity System Architecture Outline

Within the system architecture of a ME all required communication and security mechanisms are combined with the hard- and software components of the execution platform. Therefore, the concepts discussed in Chap. 3 and Chap. 4 are integrated into the ME system architecture.



(a) Components and interactions inside a Mobile Entity

(b) On-Board Unit communication stack with security integration

Figure (5.6) Mobile Entity components and architecture setup

Its main components and the protocol stack are introduced in the following. The security integration into the architecture has previously been published in [Eic07c].

Components of the In-Vehicle Architecture Setup

The components of a typical ME architecture setup are shown in Fig. 5.6(a). The central component in the architecture is the On-Board Unit (OBU), which is the execution platform for all services and related applications. The execution platform can for example be realized using Java combined with the OSGi framework [FCD06, pp. 74]. This combination allows remote servicing and service application provisioning as well as multiple networked service applications running in parallel.

The OBU is integrated into the vehicle architecture, thus, it can access and use the vehicle's sensors and control units. In addition, it can use the communication devices installed in the vehicle. Several support functions are provided by components such as Global Positioning System (GPS), which provides positioning and time information for the architecture. To extend the number of traffic information sources, broadcast services like Radio Data System (RDS), Traffic Message Channel (TMC), and the Transport Protocol Experts Group (TPEG) protocol [Tra08] can be integrated into the architecture and provide their data to the OBU. Supplementary to these broadcast services the digital broadcast technologies (Digital Audio Broadcast (DAB) and Digital Video Broadcast (DVB)) can be used in the architecture to provide both entertainment and information services.

The integration of different communication and security mechanisms is primarily influencing the design of the OBU and its system stack. In Fig. 5.6(b) the integration of security and communication mechanisms in the stack is shown. Depending on the service, application layer security mechanisms can be applied. Before the application data is sent, either V2V or platform service relevant communication protocols are used. Here, the communication mechanisms suggested in Chap. 3 are integrated into the architecture. In a final step the Secure Communication Layer is passed which is used to add the required security features.

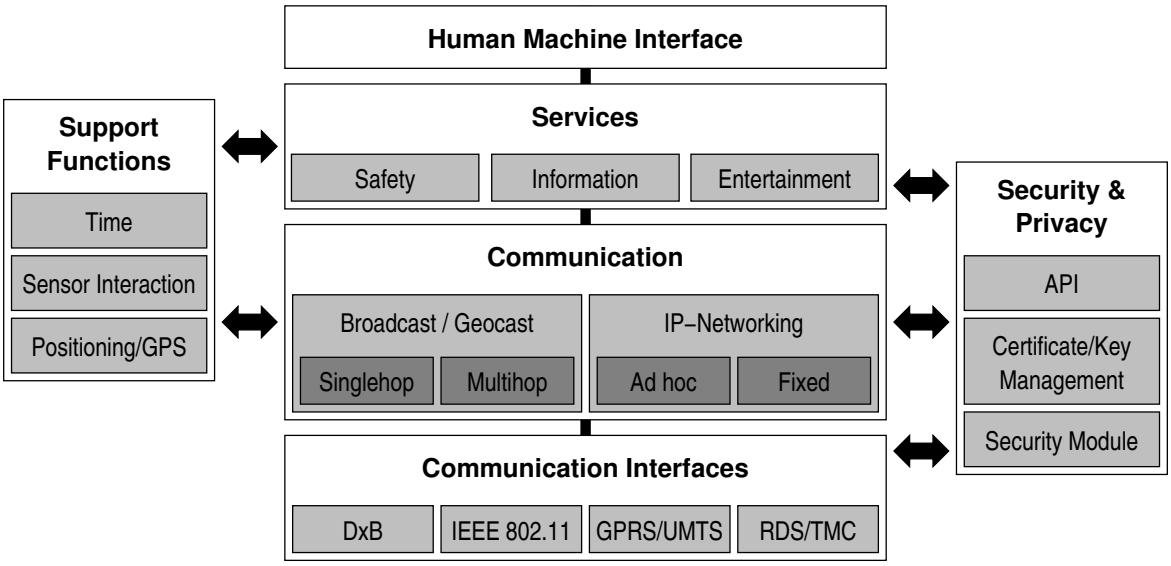


Figure (5.7) General structuring of the Mobile Entity's system protocol stack

The Architecture Protocol Stack

The architecture protocol stack has been briefly introduced above. Since it is a very important part of the in-vehicle architecture it is described in more detail. The general in-vehicle system architecture can be grouped into three component blocks: The main system stack (interfaces, communication, services, HMI), the security and privacy components, and the support components (for example in-vehicle sensors). This general architecture and the main component interactions are depicted in Fig. 5.7.

The system will rely on several communication interfaces, prominent examples are the Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 standard family, and cell-based communication systems like GPRS and UMTS. In addition, broadcast communication like DAB or DVB might complement the conventional communication means. Traffic related broadcast communication is already provided by the RDS or the TMC today, this will also be the case in future implementations. The communication layer components are primarily relevant for Dedicated Short Range Communication (DSRC) and the cell-based communication. We can differentiate between mere broadcast communication and Internet Protocol (IP) networking. Depending on the type of data and the respective service, one of the communication variants will apply. The service applications in the third component layer interact with the communication layer primarily to receive and send data content. The interaction with the user is carried out through a HMI.

In order that the primary component stack can operate, miscellaneous support functions are required. The most relevant ones are given as examples in Fig. 5.7. The interaction with sensors is crucial to generate information and detect events that the VANET services can use and disseminate. Node positioning is also an important support function needed for services as well as geocast-based message distribution. The support functions primarily interact with the services and the communication components.

The third component block contains the security and privacy modules. Since the security modules need to interact with most system components in the main thread, they are combined externally in an extra block. This setup results in several advantages. Besides the mentioned interaction diversity it is not reasonable to multiply identical security functions needed for several other layers and blocks. This would complicate the system implementation and increase the system's complexity. In fact a separate security implementation providing all required mechanisms for the whole system can be verified much better, reducing the probability of a false implementation often leading to compromised systems. In addition, the requirement for a secured overall system calls for a hardware security implementation, which is capable of sufficiently protecting the security basis. This hardware component, detailed below, shall be implemented in the system only once. Hence, a single security Application Programming Interface (API) is required.

5.3.3 Integration of Communication and Security Modules into the Mobile Entity's System Setup

The setup of the OBU and the structuring of the protocol stack has been shown in Fig. 5.6 and Fig. 5.7. Based on these results communication mechanisms and security functionalities are integrated. Similar to the extensible OSGi execution platform, the communication and security mechanisms are integrated as modules. Hence, existing modules can be updated and new modules can be added easily. This ensures the compatibility of the architecture to future system changes and new services.

Communication Features in the Protocol Stack

Many different communication features and protocols need to be integrated into the stack of an in-vehicle OBU. An example is shown in Fig. 5.6(b), depicting the vertically split communication stack. While the platform services rely on session-based communication using conventional Internet protocols the IVC applications need a variety of different communication protocols.

The application hands down its data packets to the *Communication Management* component. Thereby the application can categorize the packets and predetermine the communication modules to be used. Nevertheless, in most cases a general classification and a selection of security properties will be sufficient. An application could for example select the V2V stack and require an authenticated data transport. The Communication Management module evaluates the packet parameters and selects the needed communication modules. In addition it manages the modules and keeps them up-to-date. If new communication modules are available or required for certain applications, the Communication Management downloads and integrates the modules into the system stack.

Possible modules for the in-vehicle communication architecture are the communication mechanisms presented in Chap. 3. An example setup is shown in Fig. 5.6(b) using the message benefit evaluation, the geographical distribution areas, and the Mobile Data Request Protocol (MDRP). Another potential module is the content aggregation. Not all but many communication modules can be used for the same message, thus, combining their function-

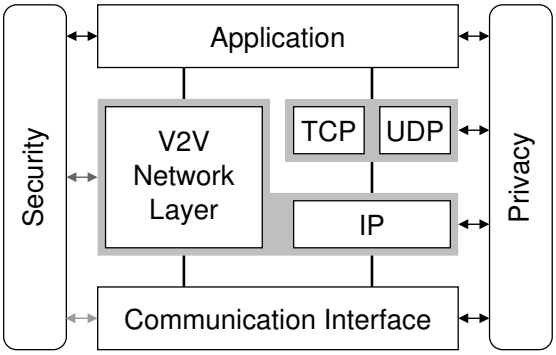


Figure (5.8) Influence and relevance of security and privacy in the communication stack

alities. This is a promising approach to integrate a multitude of different communication mechanisms into the same architecture and allows for an optimal mechanism selection.

Security and Privacy for the Mobile Entity Communication

A very crucial implementation task is the integration of security and especially privacy into the communication stack. In Fig. 5.8 the typical communication stack for a vehicular node is depicted. As introduced above, the communication stack is split into two parallel paths: The V2V and the conventional IP-based communication path.

A security integration is mainly required for the application layer. Especially for the VANET services the security can be handled solely in the application layer, as long as no secure routing mechanism shall be used in the lower layers. But for a mere broadcast application a security integration in the application layer is sufficient. However, security for the platform services also has to be integrated in the network layers. One possibility to realize this is to use a specific security layer in the communication stack (see Fig. 4.10 [p. 99]), as suggested by the GST project [EBM+05]. The functionalities shown in Fig. 4.10 can mainly be handled by the security components block depicted in Fig. 5.7, since most functions are standard mechanism also used by other layers. Specific for this security layer is the *secure communications engine*, which is managing the security sessions of the node.

Whereas the security integration into the communication stack is straight forward, the privacy integration needs more attention and care. As it is shown in Fig. 5.8 the privacy block has strong links to *all* communication stack blocks. This is crucial for the success of the privacy concept. To provide privacy effectively, the whole system architecture has to be evaluated and included in the privacy concept. In contrast to security mechanisms it is not sufficient for privacy to be included in one layer (for example the application layer). To provide a high degree of anonymity the privacy concept has to ensure the unlinkability of ”traces“ a node leaves behind. Since *all* layers in the communication stack generate characteristic labels, such as IP-addresses, Medium Access Control (MAC)-addresses or session identifiers, the privacy component has to have control over every communication layer and the generation of labels. If any label changes *all* other labels have to change as well, otherwise the messages can be linked by the unchanged label(s). Hence, a single

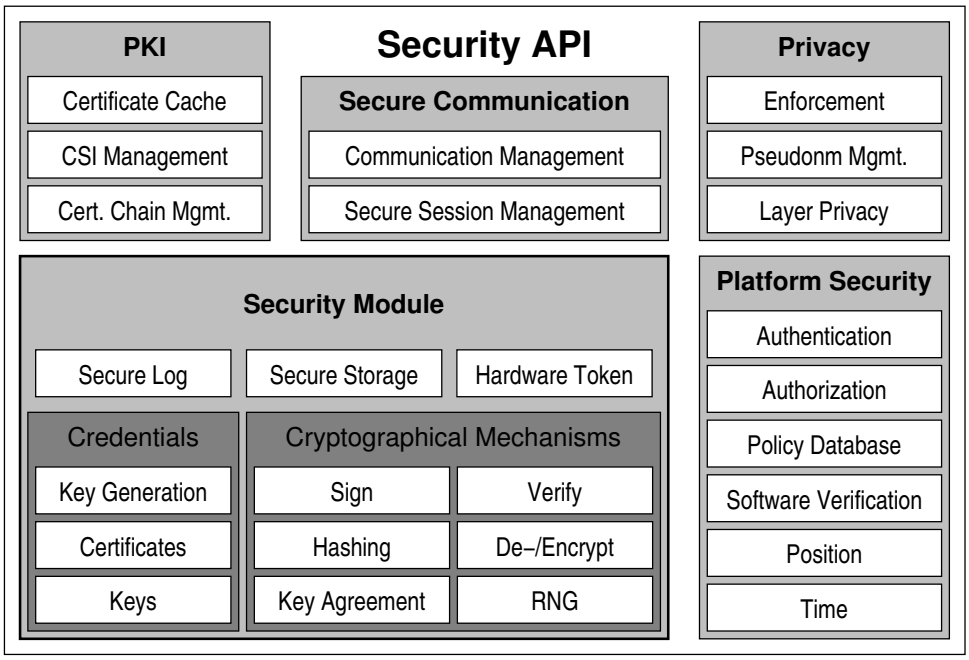


Figure (5.9) Security API and Security Module for a Mobile Entity

privacy component has to control all involved communication- and label-generating blocks in the system. Otherwise no reliable privacy concept can be provided and information leaks remain.

5.3.4 Integration of Security – Security API and Hardware Security Module

The most important component for a secure vehicular network node is the security component. The security component has to fulfill all security and privacy requirements outlined in Sec. 2.6.2 and provide the needed mechanisms. It should be implemented with two parts, a security software API and a hardware security module referred to as Security Module. As the Security Module can not function alone it can be seen somewhat integrated into the API, which provides all necessary functionality to use the hardware component. Refer to Fig. 5.9 for the full security API definition.

The API can be split into five blocks. The four software components (PKI, Secure Communication, Privacy, and Platform Security) use the security mechanisms and the secure storage functionality of the fifth component, the Security Module. The Security Module is the most crucial component of the system concerning security. It is a tamper-proof hardware security component similar to a Trusted Platform Module (TPM) [Tru08] or a smartcard. For the security level of the whole system it is crucial to implement the core security functions using a tamper-proof hardware component. This is the only way to sufficiently secure credentials, certificates, and key material on a platform being used in the field. The key materials will be generated and stored within the Security Module and can not ”leave“ it at any time without destroying the component. This feature of the system is required to maintain a high degree of security and minimize the number of overall certificate revocations

throughout the system. On the other hand, this requires a hardware component which is also capable of using the credentials and keys, generate for example secure random values, and encrypt/decrypt messages. The overall policy must be: No security credentials and keys that originate from a Security Module may leave the module in any way, all security operations have to run on the module itself.

The trust of the system and the related credentials for each node are handled in the PKI API-component. All security-related communication is handled by the Secure Communication component. The general security features (authentication, confidentiality, policy enforcement, secure software updates, and others) are handled by the Platform Security component of the API. The management of privacy and the handling of pseudonyms as well as the layer management (refer to Sec. 5.3.3 [pp. 147]) is handled by the Privacy component.

5.3.5 Interactions between Mobile Entity Components

Like in the Backend sub-scenario, the components in a ME use certain interactions. The core interactions are shown in Fig. 5.6(a).

As described above, the OBU is the central component inside a ME. All information available to the ME is managed by the OBU. It uses the sensors to acquire new information about the current vehicle status. The collected data is provided to the service applications having sufficient access rights. The link to the control units is required for applications that are allowed to influence the driving behavior of the vehicle.

Since several applications run on the system in parallel, the OBU needs to control the access to the wireless interfaces. This is primarily the case for the cell-based communication devices. A managing component, which provides access to the available wireless communication technologies, similar to the communication gateway presented in [KBS+01], could be used to coordinate the service access and optimize the connectivity. Moreover, the OBU collects information from components like the GPS or the traffic information systems and provides the information to its applications. This reduces the load on the in-vehicle communication architecture and still provides up-to-date information to all service applications.

Besides managing the access to communication devices the OBU uses a policy database to control the access to vehicle sensors, control units, and support functions like GPS. Service applications are not necessarily allowed to access all resources of the ME, therefore, a rule base is used to define the applications' access rights for the in-vehicle components. The security implementation in the execution environment of the OBU is enforcing these access rules and allows or denies resource access to applications.

5.4 Application and Management of the Vehicular Network Architecture

After having introduced the Backend and ME architecture concepts in the previous sections, the application and management of the full VN architecture is discussed. Based on a few

platform and IVC service examples the functionality of the system architecture concept is exemplified. This shall make the concept more clear and point out some of its benefits.

5.4.1 Application Examples for Platform Service Provisioning

To exemplify the use of platform services with the architecture two settings are discussed: Using Single Sign-On (SSO) for a transparent service usage and the general service provisioning when multiple service providers are involved. These two examples show the component interaction and the application of security mechanisms for platform services in VNs.

Realization of Single Sign-On for Transparent Service Usage

A very important feature for service platform architectures is the so-called SSO, where a user has to authenticate to the system architecture only once. After a successful authentication the user can transparently use all subscribed services without further authentication interaction. The main goal is to reduce the required user interaction to a minimum and provide a trust environment which includes all relevant actors of the system.

To be able to use SSO in a VN several requirements need to be fulfilled. The main prerequisite is the setup of a circle-of-trust, both on a contractual and a technical level. A commercial agreement of all participants permits a seamless interaction between different entities belonging to the circle-of-trust. The contractual basis is mapped to the technical level using security mechanisms and protocols, for example the Security Assertion Markup Language (SAML). A second requirement is the federation of information, especially identities. A user will be identified with several different identities or credentials by many different entities in the system. However, to provide a seamless interaction after one authentication these identities need to be linked with each other, which is meant by *federation*. To realize a circle-of-trust for SSO all actors in the Backend need to participate. The platform operator will be the aggregator, thus, the federating entity, managing all user data. It takes in a central role in the SSO procedures, since all authentication and verification interactions involve the aggregator.

The process of a SSO session establishment works as follows. The user authenticates with its personal identification data to the aggregator. The aggregator validates the authentication and issues an authentication credential to the user. The credential is stored in the OBU of the ME and used for each interaction within the circle-of-trust. Hence, for each interaction with a SC, for example, a service usage, the credential needs to be presented to the peer. The SC receiving the credential requests a validation from the aggregator and only if a valid credential has been presented the service is provided to the user. In course of the credential validation the aggregator will identify the user to the SC by presenting the user's identity known to the SC. Platform tasks, for example invoicing, can be handled easily using this identity management.

A more detailed description of SSO and the related federation concepts has been published in [VWE06]. In addition, further information on SSO concepts, the related security protocol (SAML), and software realizations can be found in [Lib08, OAS08, Ent08].

Service Provisioning and Usage with Multiple Service Providers

To be able to distribute service applications to MEs and configure their settings, a provisioning procedure is required. An important requirement is that multiple service providers can be used by the same ME. The service provisioning and usage follows several steps. These are explained in the following section, leaving out the security interactions in order to simplify the example. Nevertheless, the SSO example already detailed the typical security interactions for platform services.

The service-related interactions have already been pointed out in Fig. 5.2. They are again relevant for the following example. A SC needs to register its service at the platform operator, providing the application software, the requirements on the OBU, and the SC network addresses. The platform operator will add the new service to the service portfolio presented to the users. Therefore, all users having a compatible OBU will be able to subscribe to the service. All services offered to the users on the platform are listed in the service portfolio, no matter which SC is providing it.

After a service is listed in the portfolio it can be provisioned to the users. As soon as a user selects a service at the CS, the provisioning and subscription process is triggered. The CS requests the service application software and additionally required components from the platform management (the CC). The CC registers the service for the respective user and provides the necessary software components to the CS. After finalizing the provisioning process the service is available to the user at the CS. All interactions required during the service usage will take place with the respective SC.

This task distribution allows the use of several different SCs, which can be used in parallel. During the provisioning process the CS receives the access data for the SC accountable for the respective service application. The SCs do not need any specific knowledge on the CSs since the CC is taking care of the software and system configurations of the subscribing CSs. Once the services are provisioned they can be used seamlessly at the same time irrespective of the SC.

5.4.2 Application Examples for Inter-Vehicle Communication

The communication settings for V2V communication applications are much more diverse than the settings for platform services. The reason for this is that IVC services can be both, fully decentralized or with central server interaction from the Backend architecture. Hence, the protocol capabilities need to be much more multifunctional.

Inter-Vehicle Data Diffusion using Security and Distribution Strategies

The main task for IVC is the distribution of data in a VANET. As shown in Chap. 3 many different distribution and communication strategies exist to increase the effectiveness and usability of these services. In addition security mechanisms and protocols as suggested in Chap. 4 are used to increase the reliability and reputation of the distributed data. All these mechanisms have been integrated into the VN system architecture concept and lead to a specific protocol sequence.

The applications generating and processing the data can rely on different security and communication mechanisms as shown in Fig. 5.6(b). Especially many of the dissemination strategies coexist and the application as well as the default system settings influence which strategies are used. Therefore, depending on the network load and other context parameters the best matching dissemination modules are used. Moreover, the application will communicate its preferences to the Communication Management, which selects the needed modules. If the system settings require an application level security, most likely a content authentication and/or reputation, these security mechanisms are applied to the content before the data packet is passed to the communication modules.

In case an application layer security mechanism requires IVC, like it is the case for the Content Reputation System (CoRS), the Communication Management is involved. It provides the single-hop V2V packet communication and distributes the request to the neighboring nodes. All replies are directly handed back to the involved application.

The data packets for the dissemination are passed through the required communication modules before they are handed to the communication device. In addition the Communication Management adds communication security and privacy if this is required. Hence, it interacts with the security API to realize the security functionalities. The interaction with the security API is also important to control and manage the privacy settings of the communication modules. This is required to increase the unlinkability of packets to the sending node, as described in Sec. 5.3.3.

Content Distribution with Backend Server Involvement

Besides the V2V multihop communication also content distribution from the Backend to the MEs is an important application. The distribution of up-to-date CSI can be realized using this approach (see Sec. 4.2.3). In this case the Backend servers are initiating the distribution process.

In the stated example the trust environment servers generate the CSI and sign it, to provide authenticity and integrity protection. Next, the data is distributed to the gateway servers, which can distribute the CSI to the MEs. The trust servers are unloaded of the traffic as soon as the gateway nodes hold the data. Depending on the distribution mechanism the gateway nodes either send out the data on request (for example using MDRP) or they broadcast the data to the MEs within radio-range.

Once the data is distributed to the MEs they themselves are involved in the distribution process, which is similar to the IVC. However, the relevance of data, distributed by a Backend server, is in many cases not bound to any context information other than its age. Therefore, the content distribution needs to be carried out quickly to be of any use. If a ME needs a certain piece of information provided by Backend servers it can request it. However, the request will not reach the servers, since the closest gateway node will be able to reply and provide the requested information. This strategy helps to reduce the load on the main servers, similar to the content distribution used in the Internet.

Content distribution with Backend server involvement is a simple store-and-forward type application. The MEs are only validating the security information, however, other data manipulation will not occur. Further, special communication protocols such as the MDRP protocol will be used to distribute this kind of information. This is mainly due to the fact

that it is content addressable and in most cases sent on a regular basis. An alternative to this distribution method would be to use a platform-based service application, nevertheless, this excludes all MEs not participating in the platform service framework. Moreover, the platform-based distribution would put high network load on the communication networks and the server infrastructure, due to the centralized approach. Thus, the distributed V2V diffusion strategy is more scalable.

5.5 Conclusions

In this chapter the basic results of Chap. 3 and Chap. 4 have been combined to define a full architecture concept for VNs. In a first step the Backend architecture has been defined and a possible realization has been specified with the GST high level architecture. In addition, organizational aspects for the PKI have been outlined.

Besides the Backend components an architecture setup of the MEs in a VN has been suggested. Based on the Open Systems Interconnection (OSI) layer model an adapted stack with integrated security and privacy functionalities has been proposed. The different requirements for integrating security and privacy in the different layers have been taken into account. Since the tamper-proof Security Module is a key element of the architecture it has been presented in great detail. To outline the functionality of the suggested architecture several real-life scenarios have been discussed exemplifying the applicability of the concept and its advantages.

With the suggested architecture a first holistic architecture concept for VNs has been proposed, integrating both specific communication and security mechanisms. The concept is adaptable to future scenario requirements, which is an important feature for the automobile context. The architecture adaptability is provided by the modularity of the concept. At runtime new modules, for example, communication protocols or service applications, can be installed and complement the already existing features. The proposed modular architecture takes into account all features and requirements needed for a VN and can be used as an example for future real-life implementations.

6 Thesis Summary, Conclusions, and Outlook

Many challenges exist when realizing a Vehicular Network (VN). These challenges include architectural issues, as well as building blocks of such an architecture, for example, communication protocols, security, and privacy mechanisms. In this thesis several results targeting some of the existing open issues have been presented. All of the research contributions in this thesis have been placed in the context of a vision for future VN scenarios. This common vision suggests the following attributes for a VN: A system based on an open platform concept, efficiency and high performance of services and applications, scalability of the system, and integration of reasonable yet adequate security mechanisms.

Based on this common vision and other existing requirements on VN scenarios, research topics have been identified. A selection of these topics was targeted with the research activities presented in this thesis. The contributions will help to realize a VN following the vision.

6.1 Results and Contributions

The results presented in this thesis contribute to the areas of communication protocols, security mechanisms and integration, and system architecture concepts in the context of VNs. All suggested solutions have to be seen as building blocks for one holistic VN architecture. The contributions can be summarized as follows.

In Chap. 3 different communication aspects, mainly of Inter-Vehicle Communication (IVC) in VN have been addressed. The main guideline for all contributions concerning communication and message distribution methods was to realize the required features with a protocol mechanism, which is efficient enough to be used in parallel with other services. In addition, the approaches were designed to be scalable and adapted to the specific requirements and characteristics existing in Vehicular Ad Hoc Networks (VANETs).

Targeting the challenge of information diffusion in a VANET, concepts like dissemination areas and data aggregation have been suggested and evaluated. Both approaches showed very beneficial behavior in the simulator and are promising concepts to optimize the data dissemination in the distributed network part of VNs. An important aspect of IVC is the

communication technology. The Institute of Electrical and Electronics Engineers, Inc. (IEEE) is currently finalizing the IEEE 802.11p Wireless Access in Vehicular Environments (WAVE) standard, which is designed for the requirements of IVC. The standard evaluation presented in this thesis clearly shows the capabilities and limitations of the technology. Especially the high collision probability for high priority packets in dense scenarios is problematic. Nevertheless, a careful parametrization will help to use the standard successfully. The suggested packet prioritization scheme using content-utility helps to increase the utility of IVC for all nodes in a scenario and is especially suited for dense networks. The simulation results strongly support the usage of this approach in future VANETs, since it helps to identify the most important messages. Moreover, the concept is adaptable to future scenario requirements and new applications. Besides the data dissemination schemes, some information in a VN needs to be distributed on a per-request basis. This need is targeted with the Mobile Data Request Protocol (MDRP), which uses gateway nodes as initial information sources. MDRP showed to be an efficient way of distributing even large fragmented data.

For the introduction of trust the well-known concept of a Public Key Infrastructure (PKI) has been identified as the best approach in Chap. 4. In addition the simulation results evaluating different certificate revocation mechanisms for VANETs support the use of a PKI, since revocation even in a distributed scenario is feasible with acceptable delay and protocol complexity. The PKI is the best suitable concept addressing both security needs and the desired centralized control of the trust management. Based on this trust architecture different security and privacy mechanisms for VNs have been outlined. The concept for Backend supported platform services uses a tamper-proof hardware device providing the security mechanisms. It relies on a security shim layer in the system stack, enabling security even for heterogeneous communication links. For vehicle-to-vehicle (V2V) data dissemination mechanisms, security like sender authenticity and message integrity are important. Additionally, the message content trustworthiness is very important for many services. With the Content Reputation System (CoRS) a message-based content reputation concept has been suggested and evaluated. The protocol relies on threshold signatures which represent the reputation value. The analytical evaluation of the concept proved that a threshold signature system with only very few key shares relative to the number of network nodes can practically operate without share collisions.

Besides the security of VNs also privacy aspects need to be incorporated in the system concept. A very common approach to realize node anonymity in combination with a PKI as trust environment is the use of pseudonyms. The presented analysis of the node mobility influence on the use of pseudonyms helped to identify several important parameters, which need to be taken into account when configuring a pseudonym change algorithm. Node re-interaction and the communication quiet-time are crucial parameters to maximize the anonymity of Mobile Entities (MEs) in a VANET. A pseudonym management can be realized based on the given parameter evaluation in combination with one of the existing anonymity measurement approaches using information theory.

The communication and security concepts have been combined in a holistic architecture concept for VNs in Chap. 5. An architecture setup for both network parts, Backend network and VANET, has been outlined. Moreover, the two main aspects of a VN architecture, communication and security, have been integrated especially into the concept for the MEs.

The component interactions in the Backend as well as inside a ME were discussed. The functionalities of the architecture have been outlined using a few typical application examples. The suggested architecture combines all of the contributions on security and communication proposed in this thesis. It is a modular concept which can be enhanced further and adapted to future requirements and challenges of VNs, which is a very important feature due to the frequently changing needs in such a system.

6.2 Outlook on Future Work

The presented results contribute to the development of VN concepts in the future. Nevertheless, the end of the rope has not yet been reached. Many issues remain to be investigated and several challenges are still open. A few examples will be outlined in the following to show a possible path for follow-up research.

Since many different concepts for communication and security have been suggested for VNs, these ideas should be implemented in a prototypical architecture. The Java environment Open Services Gateway Initiative (OSGi) is a potential candidate for such an implementation. It has been already used in this context, especially for the conceptual implementations of the Global System for Telematics (GST) research project. A full architecture prototype can be used to do a full-scale performance evaluation estimating for example the hardware requirements inside the vehicle. Such an approach is followed with the “Sichere Intelligente Mobilität – Testfeld Deutschland (SIM-TD)” project in Germany.

A second open issue is the evaluation of IVC when multiple services run in parallel. Such a scenario is difficult to simulate, since it consumes a lot of memory and processing power. Hence, new simulation approaches for example using grid computing could be very helpful to analyze this complex setting. Moreover, a real-life implementation of the suggested message dissemination concepts in a prototype could be used to practically evaluate the approaches within the real environment. The results of these tests can be used to further improve the communication concepts, bringing them closer to reality.

Very important applications will be the safety services for collision mitigation or avoidance. Many of them have very strong delay and real-time Quality of Service (QoS) requirements, which can most likely not be met with the existing technologies. Exchanging warning messages with very low delay, meeting strong QoS requirements is very difficult to realize using a shared medium communication technology, especially in dense network scenarios. New channel model concepts with very high accuracy are required to investigate these services using a network simulator. Hence, concepts like optical ray-tracing should be taken into account to meet this requirement.

Besides the mere technical research also the acceptance of the users should be investigated. The needs and expectations of a typical user should be correlated with the existing approaches. Only if the user acceptance can be met in a very short time, VNs will be a success in the future. Therefore, a practical confrontation of potential users with a vehicle equipped with V2V services in a test environment will give very important insights.

The research community has achieved remarkable progress in the context of vehicular networking over the last years. Many contributions have been made in various areas of the research field. This thesis contributes to the areas of communication and security concepts

in VN scenarios. Additionally an architecture setup has been proposed which is capable of combining centralized and distributed services, using the suggested communication and security realizations. The architecture concept shall ease the integration of both communication and security mechanisms required in VNs. Only a VN architecture including reasonable security and scalable communication protocols will be successful in the future.

A

Simulation Environment, Supporting Models, and Parameters

An integral method for wireless network research is simulation. Many research aspects can not be directly implemented in a demonstrator, since they are very complex. In addition, such an implementation is costly and time-consuming. However, many protocols or mechanisms applied in wireless scenarios can be implemented and evaluated in a network simulator.

Besides commercial simulation libraries, such as OPNET [OPN08], many different network simulators exist in the public domain. The most important simulation environments are:

- Global Mobile Information Systems Simulation Library (GloMoSim) [ZBG08, ZBG98],
- Network Simulator 2 (NS2) [Inf08],
- Java in Simulation Time (JiST) in combination with the Scalable Wireless Ad hoc Network Simulator (SWANS) [BH08],
- Georgia Tech Network Simulator (GTNetS) [Ril03, Ril08],
- OMNeT++ in combination with the INET Framework [Var01, VH07].

For the network simulations presented in this thesis the OMNeT++ simulation framework has been used. It has been selected due to its complete code documentation and the excellent user manual, as well as its module oriented system design. The most important features and existing models of the OMNeT++ simulator and connected model libraries are introduced in the following sections. In addition, the commonly used models, the simulation parameters, and the analysis methodology are presented.

A.1 The OMNeT++ Simulator and the INET Framework

The OMNeT++ simulation framework is a discrete event-triggered simulator written in C++. It is very well suited for network simulations or parallelized/distributed systems. The

simulation framework is very well adaptable, especially due to its object-oriented nature and the fact that it is almost completely written in C++.

A.1.1 Simulations with OMNeT++: Usage and Model Composition

The basis for simulations with OMNeT++ are components, so-called modules. Hence, every simulation model has to be broken down to several components which interact. This requirement is especially well suited for communication network models, since they're usually designed in a layered system design. This International Standards Organization (ISO) Open Systems Interconnection (OSI) layered system design can easily be transferred into an OMNeT++ simulation model. The main system design steps for the OMNeT++ simulator are:

1. Design the module interface and parameter descriptions and the module interconnections with the Network Description (NED) language. This can be done with the Graphical Network Editor (GNED) or a conventional editor. The resulting NED modules are later converted into C++ modules and integrated into the simulation application.

2. Program the active modules with C++. The interface and parameter definitions of the NED modules need to be incorporated into the C++ module definitions. The actual simulation action is taking place within these C++ modules.

3. Compile the model and link it with the OMNeT++ simulation kernel. This generates the binary simulation application.

4. Configure the simulation model with an initialization file. The configuration has to configure the OMNeT++ simulation kernel as well as the simulation model. It is possible to configure several simulation runs which are simulated sequentially.

5. Run the simulation and analyze the results. The OMNeT++ simulator can generate preprocessed simulation results at the end of a simulation run as well as continuous simulation values while the simulation is running.

Steps one and two are most important for the simulation module generation. Thus, they are briefly introduced next. A detailed introduction into OMNeT++ and the programming procedures will not be given here, since a comprehensive manual exists [VH07].

Module Definition Using the NED Language

In order to be able to use a C++ model in combination with the OMNeT++ simulation kernel the interfaces and parameters of a component need to be defined first. This is done with the simple yet effective NED language. Two types of modules exist within the NED language, simple and compound modules. While a simple module is the smallest building block of an OMNeT++ model, a compound module can combine several simple modules to a new and more complex module. An example for a compound module is given in Lst. A.1.

Listing (A.1) A NED file example of a compound module (Mobile Entity)

```
1  import
       "NotificationBoard",
3      "Ieee80211NicAdhoc",
       "MobileEntityApplication",
5      "MobileEntityNetworkLayer",
       "BasicMobility";
7
   module MobileEntity
9      parameters:
           mobilityType:     string,
11         applicationType:  string,
           networkType:      string;
13     gates:
           in: radioIn;
15     submodules:
           notificationBoard: NotificationBoard;
17         applicationLayer:  applicationType like MobileEntityApplication;
           networkLayer:      MobileEntityNetworkLayer;
19         physicalLayer:     Ieee80211NicAdhoc;
           mobility:          mobilityType like BasicMobility;
21     connections nocheck:
           radioIn                    --> physicalLayer.radioIn;
23         physicalLayer.uppergateOut --> networkLayer.ifIn;
           physicalLayer.uppergateIn  <-- networkLayer.ifOut;
25         networkLayer.toUpper       --> applicationLayer.appIn;
           networkLayer.fromUpper     <-- applicationLayer.appOut;
27 endmodule
```

The example listing (Lst. A.1) is the definition of a Mobile Entity (ME) with all its submodules. Starting from line 2 other modules to be included are defined. The actual module is defined between the **module** and **endmodule** keywords. As of line 9 the module's parameters are defined. Starting with the keyword **gates** the module's interfaces are stated. The required submodules are listed starting from line 15. After line 21 the module's connections to the outside as well as the connections between the submodules are set up. No directly connected C++ module exists for the given example, since it specifies a compound module. However, the *MobileEntityApplication* is a simple module which has a C++ module realizing the functionality of the component.

Programming the Active Module with C++

Each C++ module needs three predefined methods to be able to interact with the OMNeT++ simulation kernel. These methods are:

`initialize()`: This method is called at the beginning of the simulation. It can be used to initialize the module parameters, its interactions with other components, and the statistical logging capabilities of the model.

`receiveMessage( cMessage* ):` Whenever the module receives a simulation event this method is called by the kernel. The argument is a pointer to the respective event. The event can be a self-message, which is equivalent to a timer, or a real message from a different component in the simulation.

`finish():` At the end of each simulation run the kernel calls all finish-methods of the simulation model. These methods can be used to log simulation results to the output file. In addition, memory cleanup can be done here.

Besides these three mandatory methods an OMNeT++ component can be programmed absolutely without restrictions. All C++ features can be used. This is very convenient since new modules can be derived from existing modules.

The C++ simulation library of OMNeT++ offers many additional features besides the plain discrete event handling. It includes comprehensive random number generation features, many statistical analysis methods, and topology handling features for fixed networks. These features help to develop efficient and complete simulation models that provide pre-evaluated results.

A.1.2 OMNeT++ Simulation Tools

Several tools are used for the module design process. They are part of the OMNeT++ simulation framework and ease the design and compilation process significantly.

`gned`: The tool GNED is a graphical module editor. It can be used to define simple and compound modules in the NED language. Since the top network scenario files are also defined with the NED language, GNED can be used to define the simulation scenario files.

`opp_makemake`: An OMNeT++ simulation model potentially consists of hundreds of files in different directories. The `opp_makemake` analyzes the file and directory structure and generates a makefile for the code. In order for the code to be compatible with the precompiled OMNeT++ simulation library, `opp_makemake` can use the OMNeT++ build configuration as a basis. This ensures compatible binaries, which is especially important when using shared libraries.

`seedtool`: A very crucial aspect of network simulation is the generation of sufficient randomness. OMNeT++ uses a reliable and widely accepted pseudo Random-Number Generator (RNG), the Mersenne Twister [MN98]. Nevertheless, it is important to initialize the RNG for each simulation run. The tool `seedtool` helps to set the initialization for the RNGs used in the simulation of OMNeT++ models.

The OMNeT++ simulation framework offers two execution environments to the user: A graphical Tcl/Tk user interface and a text-based user interface. The graphical interface is optimal for debugging and presentations, while the text-based interface is best used for simulations of several sequential runs.

A.1.3 INET Framework Model Overview

The OMNeT++ simulation framework is only the required event-triggered simulator, however, network models are still missing. Thus, an additional network model package is required to simulate communication networks. Several network model bundles exist for the OMNeT++ simulator. The most comprehensive model library especially for Internet Protocol (IP)-based and wireless networks is the INET Framework. Hence, this library (version 20061020) was used as the basis for the research presented in this thesis.

The INET Framework is a collection of different host types, protocols, interfaces, applications, and host mobility models. In addition, several example scenarios and test networks are included. Therefore, the framework is a complete simulation library for wireless and fixed networks, very well suited for the application in research activities. The core module types of the framework are (also see `http://www.omnetpp.org/doc/INET/neddoc/index.html` for detailed documentation):

Host types: Different network hosts exist in a network scenario. The framework provides a selection of typical hosts like router, switch, and client/server host. A host type is a compound module which combines underlying compound and simple models of different OSI layers.

Network interfaces: The hosts need a communication interface model which can be selected from several different network interfaces of the model library. Besides the Ethernet and Point-to-Point Protocol (PPP) model also several wireless interface modules based on the Institute of Electrical and Electronics Engineers, Inc. (IEEE) 802.11 standard family are contained in the framework.

Communication protocols: A wide range of communication protocols exist for the INET Framework. In addition to the IP/Transmission Control Protocol (TCP)/User Datagram Protocol (UDP) protocol family, protocols like Address Resolution Protocol (ARP) and Internet Control Message Protocol (ICMP) are implemented. The required protocol message formats are also part of the model library and can be used in combination with a wide range of application models, for example an UDP video stream server.

Routing/MPLS: The INET Framework also supports a selection of routing mechanisms and Multi-Protocol Label Switching (MPLS) models. Host addresses and static routes can be configured automatically using the so-called `FlatNetworkConfigurator` module.

Wireless support: Supplementary to the wired models the framework supports wireless communication scenarios. Besides the specific wireless host types using the wireless network interface model, a control module for the wireless channel exists. The so-called `ChannelControl` module helps to identify hosts within radio- or interference range and distributes wireless packets accordingly.

Mobility models: Since wireless communication scenarios can require host mobility, several host types support mobility. A selection of elementary mobility models, for example Random Waypoint Mobility (RWM), are implemented in the model library.

A.2 Overview on Existing Simulation Models and Approaches for VANET Simulations

A broad variety of publications on simulators, models, and simulation best practices exist. A selection of the most important ones related to this thesis is given in the following. Guidelines for network simulation and configuration of simulation models can be found in [BP96, BEF+00, HMK01]. A commonly used simulation model with reproducible settings is the best basis to generate valid results. Therefore, this was the approach to be followed for the simulations used in this thesis. One simulation environment with the same basic models has been used for the presented results.

The simulations in this thesis have been done using the OMNeT++ simulator. A simulator design for Inter-Vehicle Communication (IVC) scenarios using the NS2 simulator has been presented in [CJTD06]. The suggested settings are very similar to the ones used in this work.

Simulation Environments

Besides the more general simulation frameworks OMNeT++, NS2, GloMoSim, GTNetS, and JiST, which can be used to simulate network scenarios in general, many specialized simulators have been developed during the last years. Most of them use digital road maps as a basis for the mobility model of the nodes. The simulator GrooveSim [MWSR05] uses digital maps to provide a simulation environment for pure simulation as well as an integrated scenario with real vehicles. An improved version of this simulator has been presented in [MWR+06].

In [HGK+04] a simulator for IVC scenarios using a microscopic traffic model has been presented. In the paper the simulator concept is described in detail, allowing it to be replicated and integrated into own simulation concepts. Similar to [EOSK05] a coupling of two simulators has been presented in [LBC+05]. The authors connected the NS2 simulator with the VISSIM [PTV08] traffic simulator, providing a comprehensive movement model for vehicles and pedestrians. In this approach the well-tested NS2 models can be used for the network simulation.

A very specialized simulation approach for the evaluation of cognitive automobiles has been suggested in [VNB+07]. The simulator allows the simulation of the cognitive vehicles as well as their mobility on a road network.

Since none of the presented simulation environments had been available at the beginning of the research work for the thesis an individual solution based on the OMNeT++ simulator and its models has been used. For future evaluations the use of a more sophisticated simulator, especially in relation to the mobility model is suggested. However, the necessity of such a detailed and demanding simulation is still not proven.

Radio Propagation Models

A very important modeling part for wireless network simulations are the properties of the radio propagation. Three main concepts can be differentiated in this respect: simple

yet effective deterministic path-loss models, path-loss models in combination with probabilistic channel characteristics, and highly realistic yet very resource consuming ray-tracing modeling approaches. To keep the processing requirements for a simulator low, hence, increase its model scalability, most simulators use deterministic models such as the Friis free-space model [Fri46]. During the last couple of years new models have been suggested, however, many of them remain questionable since they have not been validated by real-life measurements.

One of the first publications discussing physical aspects of Dedicated Short Range Communication (DSRC) in context of vehicular scenarios is [TJM+04]. The authors suggest to use the Nakagami fading channel [Nak60] and present some empirical data sets.

A very sophisticated and exact radio propagation model specifically designed for IVC has been presented by Maurer et al. in [MFSW04, Mau05]. The model is based on optical ray-tracing and achieves remarkable accuracy. However, the processing requirements prohibit to use the model in large-scale network environments. Moreover, detailed information on building positions and shapes are required to use such a modeling approach.

Starting out from the IEEE 802.11a communication standard and available error models a new model for the IEEE 802.11p Wireless Access in Vehicular Environments (WAVE) communication standard has been introduced in [ZSO+05]. The presented channel model is specifically designed for highway scenarios, thus, vehicles are assumed to be aligned in a straight line. Therefore, the model is not suitable for simulations in city environments.

Based on their comparison between the two-ray ground propagation model and the Nakagami fading channel model presented in [TMJH04] the same authors showed the effects of a more realistic channel model on packet forwarding mechanisms in [TMSEFH06a]. Again a highway scenario has been used as a basis for the evaluation. However, the results also support the assumption that more realistic propagation modeling has a significant influence on simulation results. Moreover, the introduction of a probabilistic propagation model is the first step to realize a more realistic modeling with a low model complexity. Whether the Nakagami fading channel is definitely a good choice for the simulation of the vehicle-to-vehicle (V2V) radio propagation should be investigated further. Based on real-world measurements a radio propagation model for the V2V scenario has been defined in [DRS06]. This seems to be a valid approach on the way to define a well-suited propagation model for the future.

Node Mobility Models

Besides the radio propagation model and the network stack a very important part of the simulation environment is the node mobility model. When Mobile Ad Hoc Network (MANET) research started to become popular, rather simple mobility models like the RWM model have been used [JM96]. However, by refining the simulation scenarios also the models had to be made more comprehensive and suitable for the respective scenarios [Bet01, Bet03, BV05]. A survey on mobility models has been presented in [CBD02].

While for most scenarios a simple randomized movement pattern is sufficient, this is not the case for the simulation of Vehicular Networks (VNs). Hence, new approaches have been defined, for example the grid movement [OPB00] similar to the street settings in large cities in the USA. In the last years the so-called microscopic traffic models found

their way into the simulation environments. They have been a research topic for many years [Wie68, Wie74, KWG97, Kra98]. A microscopic traffic model simulates traffic patterns down to the detail of a single vehicle or even the driver behavior. They use digital map information to position the vehicles and can model real-life traffic situations very realistically. One of the first freely available mobility simulators was Simulation of Urban Mobility (SUMO) [KHRW02, Cen08]. However, since it is not combined with a network simulation environment it can not be directly used to evaluate VNs.

A similar approach is followed in [CB05a, CB05b]. The Street Random Waypoint (STRAW) simulator uses a microscopic traffic model approach in combination with digital maps. It is provided as an add-on module for the JiST/SWANS simulation framework [BH08], thus, it is already combined with a network simulator. A commercial traffic simulator is VISSIM, which is provided by the PTV AG in Karlsruhe, Germany [PTV08]. It is probably the most elaborate traffic simulator available today, which can simulate complex real-life traffic situations providing a large variety of road users.

In [SJ04] a comparison between a fully random mobility model and a street-based model is presented, pointing out the need for new mobility models to simulate Vehicular Ad Hoc Network (VANET) scenarios. The use of digital maps for traffic simulation and the resulting challenges are discussed in [KHRW05] in great detail. The authors show how map data should be handled and how the road connections at intersections can be managed. This paper provides an important information basis for the implementation of such simulation environments. Definition guidelines for new mobility models in the context of IVC scenarios are presented in [HFB05].

A.3 Newly Added Simulation Models for VANET Simulations

The existing simulation models of the INET Framework have been an excellent basis, however, several models had to be revised or complemented. The most important general changes of the models used in the thesis are presented in the following section.

A.3.1 Mobile Entity Node Type

A very important component in all simulations of the thesis is the simulation model of the ME. The component module structure of the model is shown in Fig. A.1 and Lst. A.1. Its design is based on the client/server host type of the INET Framework. Nevertheless, it is of reduced complexity since MEs communicate in the ad hoc mode only, at least in the simulations produced for this thesis. Usually, the communication is of broadcast fashion, since no discrete communication partner is addressed. Therefore, the network layer does not have to use addresses or ARP mechanisms.

The main stack of the model uses a general application layer model. It provides three main features: Separation of network and node-timer messages, generation of processing delays, and providing a basic Broadcast Storm protection.

A problem of discrete event-based simulators is absolute precision in time. Hence, by default no jitter exists. However, this does not correspond to reality at all. Therefore, the jitter of events has to be generated by the simulator. The ME simulation model provides

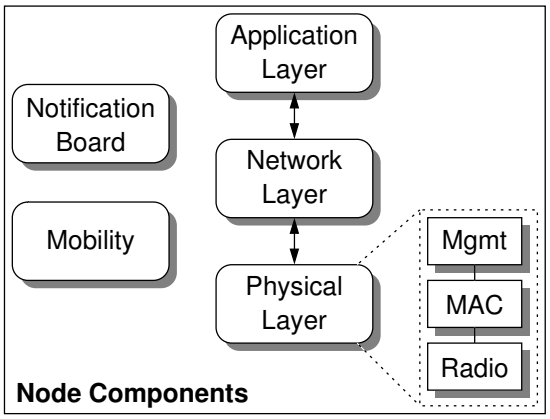


Figure (A.1) Mobile Entity component model within OMNeT++

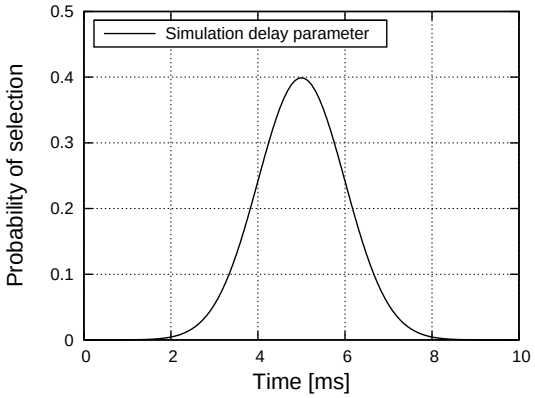


Figure (A.2) Probability density distribution for a processing delay with 5 ms mean

the possibility to generate jitter using a truncated Gaussian distribution, using only positive delay values. The model is parametrized with the mean of the jitter and its variance σ^2. In Fig. A.2 an example is given for the mean $m = 5$ ms and $\sigma^2 = 1$ ms. This processing delay can be used in the application and the network layer of the ME model.

Following the OSI layer model, the networking layer succeeds the application layer. The networking layer is primarily used as a shim layer, providing the required message class conversion from an application layer message to a compatible message format for the network interface. This conversion primarily adds the broadcast header by encapsulating the application layer message into the network layer message. In addition, the network layer model provides a processing delay generation identical to the application layer delay generation.

A very crucial mechanism of the simulation model is the Broadcast Storm protection mechanism. When mobile nodes broadcast messages and the receivers keep forwarding these messages, a Broadcast Storm will occur. This means that one message results in a storm of forwarded messages in a very short time, leading to an exponential increase of packet collisions and loss [NTCS99]. The model used in this thesis prevents the Broadcast

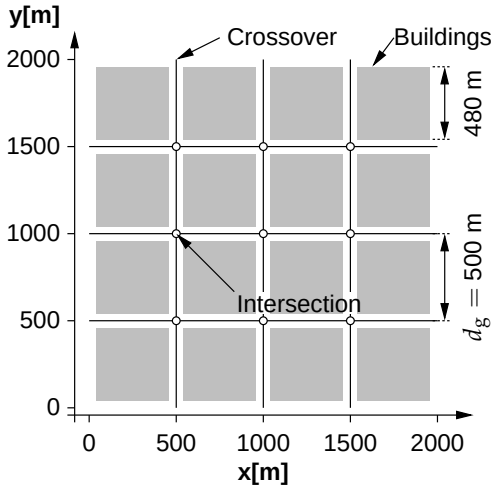


Figure (A.3) Manhattan Grid Mobility model with buildings

Storm by deleting previously seen messages. Since each message can be identified by a message ID, the model provides the possibility to delete all messages that have been received previously. The model logs the last 250 messages, which proved to be sufficient to prevent the Broadcast Storm issue.

The network layer hands messages down to the physical layer, which is a compound module composed of three simple modules: the interface management, the Medium Access Control (MAC) mechanism, and the radio module (see Fig. A.1). The management has the task to control the packet queue for outgoing messages as well as the forwarding of received messages to the upper layers. The MAC and radio modules are responsible for medium access and the physical bit layer mechanisms, including the propagation and reception model. These two components are primarily set up according to the IEEE 802.11 standard [LAN99]. Nevertheless, the components can be configured using several parameters, for example, the bit rate or the transmission power of the interface, to change the model behavior. More information on the simulation parameters can be found in Sec. A.4.

Besides the stack components, two additional modules are part of the ME simulation module. The `NotificationBoard` is used to provide cross-layer inter-module communication. For example the latest entity position is potentially needed by several layers. Thus, the mobility component publishes the node position to the `NotificationBoard`. All modules can subscribe to certain information categories provided by the `NotificationBoard`. As soon as an updated information is published the `NotificationBoard` informs all subscriber modules of the respective information. The `Mobility` module handles the node mobility. It is responsible to update and provide the latest node position to the modules of the ME.

A.3.2 Manhattan Grid Mobility Modeling

The existing mobility models of the INET Framework are elementary mobility models like the mass mobility model or the RWM. They are suitable for more general MANET

Parameter	Description	Typical Setting
`speed`	Speed of the nodes	12 m/s
`updateInterval`	Time-frequency for position updates	0.1 s
`gridLength`	Distance between roads	500 m
`numRoads`	Number of roads in each orientation	3

Table (A.1) Parameters and typical settings for the MGM model

simulations, nevertheless, for the specific mobility patterns of VNs they are too generic. Therefore, the City Section Mobility model [CBD02] has been altered and refined to the Manhattan Grid Mobility (MGM) model [NE08]. It uses a squared simulation playground with an equidistant grid. A typical simulation setting with grid length (d_g) of $d_g = 500$ m and three roads in both horizontal and vertical orientation is shown in Fig. A.3. The model distinguishes intersections and crossovers.

Even though realistic vehicle movement behavior can be realized using a microscopic traffic model [Wie74, Kra98], the required processing time makes the simulation of larger scenarios (more than 500 vehicles) unreasonable. Hence, the goal was to realize a realistic yet very efficient mobility model applicable to VN scenarios. The effects of a realistic mobility modeling on the simulation results and its influence on simulation methodology have been investigated in [EOSK05]. Thus, the integration of microscopic traffic models should be followed in future research activities, especially if exact vehicle movement is required.

Model Description and Parameters

The MGM model handles all entities equally. During the initialization phase each node is randomly placed on the grid and gets a movement direction assigned. In addition, the movement increment (d_{inc}) is determined. The mobility of nodes is not continuous, since a discrete event-triggered simulator is used. Hence, the positions of nodes are updated once every update interval t_{ui}. In combination with the node's speed (s_{node}), d_{inc} can be calculated (see Eqn. (A.1)). The model has four parameters for configuration (see Tab. A.1).

$$d_{inc} = s_{node} \cdot t_{ui} \tag{A.1}$$

The node's position and its movement direction determine the distance to the next intersection or crossover. If a node reaches an intersection its position is corrected to eliminate roundoff errors. In addition, the new movement direction is chosen randomly. In case a node reaches a crossover it leaves the scenario and reenters at a randomly chosen crossover.

Modeling of Node Replacement

A variant of the MGM model allows to remove nodes at the border of the playground and replace them with a new node. For this feature an extra model parameter has been defined (`replaceNodesAtBorder`). If this parameter is set to *true* all nodes reaching a crossover leave the scenario and are replaced by a newly initialized node. This node enters

the scenario at a randomly chosen crossover. Therefore, the node density of the scenario is maintained throughout the full simulation time.

To be able to remove and add nodes during simulation time, initiated by the mobility model, the ME simulation model shown in Fig. A.1 has to be modified slightly. The mobility component is removed. Instead a central mobility module is introduced on the global scenario level. This new mobility model component handles the mobility of *all* nodes in the scenario. Moreover, it actually generates the nodes during the initialization phase. In addition, it uses the dynamic module generation capabilities of the OMNeT++ kernel to remove and add MEs during simulation.

A.3.3 Physical Layer Radio Propagation Model

A very important part of wireless network simulations is the model for the radio wave propagation. As presented in Sec. A.2 many different proposals have been made for the simulation of radio wave propagation for MANETs and VANETs. However, so far an important issue concerning VANET scenarios has been left out, the signal shadowing caused by buildings. Thus, a new propagation model has been designed, which is interacting with the underlying mobility model and still can be simulated in with reasonable simulation speedup.

Existing Free-Space Radio Propagation Model

The included radio propagation model in the INET Framework is of simple nature. The model uses two threshold values: the reception threshold and the Signal Attenuation Threshold (SAT) (t_{sat}). The SAT, which is typically set to -110 dBm, can be used to determine the interference distance (d_{Int}) (see Eqn. (A.2)). With the given SAT, path loss $\alpha = 2$, and a transmission power of 2 mW d_{Int} equals to 4445.44 m.

$$d_{\text{Int}} = \left(\frac{\lambda^2 \cdot p_{\text{trans}}}{16 \cdot \pi^2 \cdot p_{\text{recv}}} \right)^{\frac{1}{\alpha}} \tag{A.2}$$

$$p_{\text{recv}} = 10^{\frac{t_{\text{sat}}}{10}} \tag{A.3}$$

The INET Framework uses a cumulative noise-interference calculation. Hence, a sent message is generating interference to all other nodes within d_{Int}. Only nodes within the radio-range will be theoretically capable of detecting the wireless message successfully. The radio-range is determined by the reception threshold (t_{recv}), which is typically set to -85 dBm. The reception power (p_{recv}) of an incoming packet is calculated with the simple Friis Free-Space equation (see Eqn. (A.4)) [Fri46]. If the condition in Eqn. (A.5) holds true, then the packet can be detected.

$$p_{\text{recv}} = \frac{p_{\text{trans}} \cdot \lambda^2}{16 \cdot \pi^2 d_{t \to r}^{\alpha}} \tag{A.4}$$

$$p_{\text{recv}} \geq 10^{\frac{t_{\text{recv}}}{10}} \tag{A.5}$$

In a next step the bit error probability is determined by using the Signal to Interference-plus-Noise Ratio (SINR) and the used modulation scheme. The SINR threshold (t_{SINR}) can be set as a simulation parameter and has a typical value of 4 dBm. Hence, the minimal SINR value of an incoming packet may not be smaller than t_{SINR}, otherwise the packet can not be detected and will be considered as noise. However, if t_{SINR} is exceeded the bit error probability is determined for the Message Protocol Data Unit (MPDU).

$$BER_{\text{MPDU}} = \frac{1}{2} \cdot \exp\left(-SINR_{\text{min}} \cdot \frac{B}{R}\right) \qquad \forall\, bit \leq 2\ \text{Mbit/s} \tag{A.6}$$

$$BER_{\text{MPDU}} = \frac{1}{2} \cdot \left(1 - \frac{1}{\sqrt{2^4}}\right) \cdot \text{erfc}\left(SINR_{\text{min}} \cdot \frac{B}{R}\right) \qquad \forall\, bit\, \epsilon\, 5.5\ \text{Mbit/s} \tag{A.7}$$

$$BER_{\text{MPDU}} = \frac{1}{4} \cdot \left(1 - \frac{1}{\sqrt{2^8}}\right) \cdot \text{erfc}\left(SINR_{\text{min}} \cdot \frac{B}{R}\right) \qquad \forall\, bit > 5.5\ \text{Mbit/s} \tag{A.8}$$

Depending on the modulation scheme the Bit Error Rate (BER) is determined by Eqn. (A.6) for Phase Shift Keying (PSK), (A.7) for 16 Quadrature Amplitude Modulation (QAM), and (A.8) for 256 QAM. Using the value for the BER the error probability for the header of the packet and the MPDU can be determined using Eqn. (A.9).

$$p_{\text{error}} = (1 - BER)^{length} \tag{A.9}$$

The existing radio propagation model is sufficient for simple scenarios only. Nevertheless, if a mobility model based on roads shall be used, it is beneficial to be able to take shadowing due to buildings into account. Thus, the propagation model needs to be adapted.

Radio Propagation Model with Building Obstructions

As already shown in Fig. A.3, a simple building setup has been made for the MGM model. Since the Free-Space propagation model only accounts for the signal power decay over the transmission distance, irrespective of obstructions, an adapted radio model was needed. Due to the road-based mobility modeling, the node density is artificially increased compared to a scenario using fully random node distribution (see Sec. A.4.2 and [NE08] for details). This leads to a higher connectivity in the simulator which does not correspond to the real-world behavior. In a real city scenario buildings cause signal shadowing, thus, the signal decay is much larger than the Free-Space model suggests. Therefore, a new propagation model based on the Dual-Slope model [CS92], which was proposed for similar settings, has been designed.

Scenario Geometry and Building Selection

To be able to calculate a different decay for signals obstructed by buildings, the buildings need to be introduced into the simulator. This is done with the building model. It is initialized with the buildings of the scenario, which are represented by a 2D polygonal baseline describing the obstacle's boundaries on the playground. The baselines of all obstacles are stored in a recursive Binary Space Partitioning (BSP) tree [AGMV96, AMV97], which enables an obstacle retrieval with the complexity $\mathcal{O}(n) = \log(n)$.

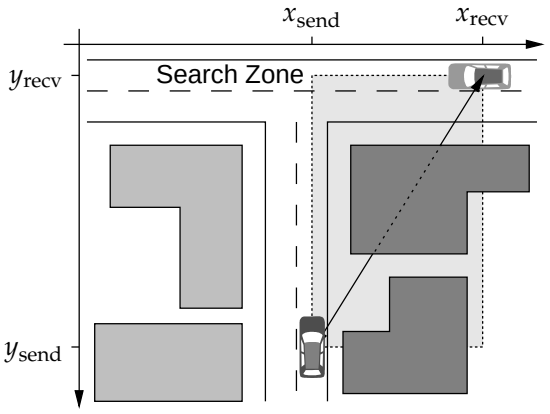


Figure (A.4) Building search zone for the BSP-algorithm used in the MGM model

Since the path loss calculation is done in each ME individually, the building information needs to be accessible to all nodes. Thus, the building model is introduced on the global scenario level. A ME can access the building model by direct method calls.

A scenario may have a large number of obstacles in the building model. But a communication link usually passes only a small selection, hence, an algorithm is needed to identify the relevant obstacles to speed up the calculation. In Fig. A.4 an example is depicted. Only the darker buildings are relevant for the given transmission. The positions of the sending node $(x_{\text{send}}, y_{\text{send}})$ and the receiving node $(x_{\text{recv}}, y_{\text{recv}})$ form the bounding rectangle of the Line of Sight (LoS) path. The BSP algorithm uses this rectangle, the so-called *search zone*, to identify the relevant obstacles which could obstruct the LoS. After identifying the obstacles, potential intersections between baselines and LoS path are searched. This results in two distances: the free-space distance d_{f} and the obstacle distance d_{o}. The two distances make up the full transmission distance and are handed to the Dual-Slope path loss model to calculate the reception power of the signal.

Dual-Slope Path Loss Calculation

The Dual-Slope path loss model uses the distances d_{f} and d_{o} in a double regression calculation, where d_{f} represents the breakpoint of the regression. The calculation directive is given by Eqn. (A.10) and Eqn. (A.11).

$$L_0(\lambda) = -20 \log_{10} \frac{\lambda}{4\pi} \tag{A.10}$$

$$L_{\text{p}} = L_0 + 10 \cdot \begin{cases} \alpha_{\text{f}} \log_{10} d & \forall\, d \leq d_{\text{f}} \\ \alpha_{\text{f}} \log_{10} d_{\text{f}} + \alpha_{\text{o}} \log_{10} \frac{d}{d_{\text{f}}} & \forall\, d_{\text{f}} < d \end{cases} \tag{A.11}$$

Eqn. (A.10) is used to calculate the reference path loss for the given wavelength λ at the reference distance of one meter. The two path loss coefficients α_{f} and α_{o} are also dependent on λ. They have been set to $\alpha_{\text{f}} = 18$ dB/decade and $\alpha_{\text{o}} = 61$ dB/decade, according to the settings of [CS92].

A.4 Typical Simulation Parameters, Model Characteristics, and Simulation Result Evaluation

Since most simulations presented in this thesis share a great deal of parameters, this section is used to introduce the typical simulation parameters. Moreover, the model characteristics are presented briefly. This helps to interpret the simulation results. All results have been statistically analyzed. The commonly used procedures are described and explained.

A.4.1 Commonly used Simulation Parameters

Number of nodes: In combination with the scenario size the number of nodes (N_n) can be used to modify the node density. For the simulations the number of nodes usually has been varied between 50 to 1000 nodes.

Simulation area: Since the MGM model has been used the simulation area was set to 2000 m $\times$ 2000 m for all simulations.

Mobility parameters: The most important mobility parameter besides the model itself is the node speed. The mobility model used throughout the thesis is the MGM model. The node speed has been set to typical speed values in city scenarios, for example 12 m/s.

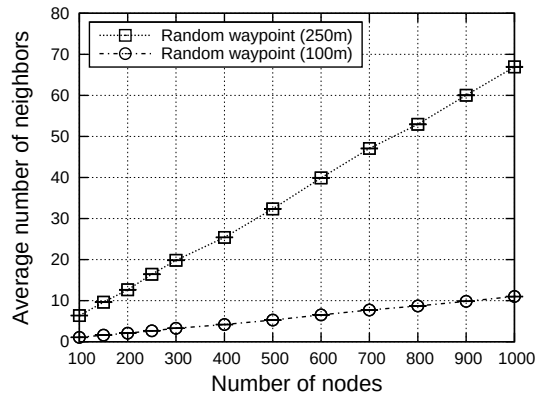
Radio parameters: To configure the wireless communication characteristics, several parameters have to be set. The transmission power p_{trans} and the reception threshold t_{recv} have an influence on the radio-range. With $p_{trans} = 2$ mW and $t_{recv} =$ -85 dBm a radio range (d_r) of 250 m is obtained for free-space propagation, when using a path-loss coefficient $\alpha_f = 2$. For the obstructed communication links the radio range has been computed using the Dual Slope path loss model specified above. Besides the mere radio range the SAT is a relevant parameter to configure the interference-range. The parameter $t_{sat} =$ -110 dBm has been used, resulting in an interference-range of 4.4 km for the free-space propagation.

A.4.2 Simulation Model Characteristics

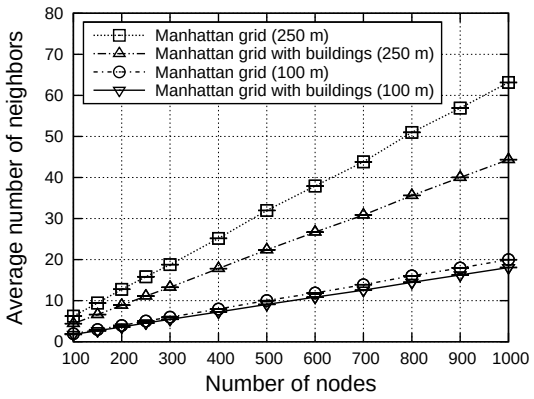
To be able to interpret simulation results based on the models described above, the most important characteristics of the basic models are described in the following. This includes the mobility model and the radio propagation model especially.

The combination of the number of nodes (N_n), the movement patterns of the mobility model, and the radio-range cause a characteristic interdependence between the N_n and the number of neighboring nodes (N_{ne}). Since many mobility-based simulations use the RWM model a comparison to the MGM model is shown in Fig. A.5(a) and Fig. A.5(b), to be able to compare the model characteristics.

The number of nodes (N_n) for the radio-ranges 100 m and 250 m is shown in Fig. A.5 for two mobility models. Considering a radio-range of 250 m the RWM and the MGM behave very similar, N_n is practically identical. However, adding buildings to the propagation model reduces N_n significantly (see Fig. A.5(b)). For shorter radio-ranges, for example 100 m, the

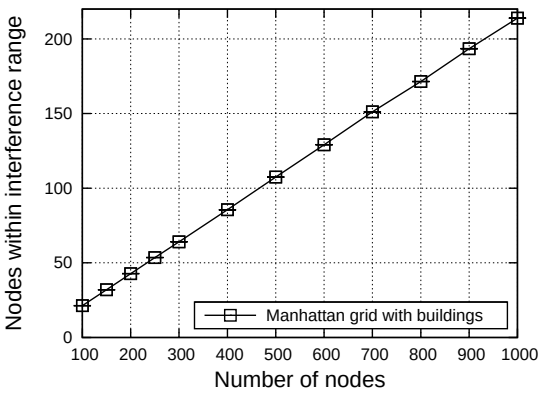


(a) N_n for the Random Waypoint Mobility model

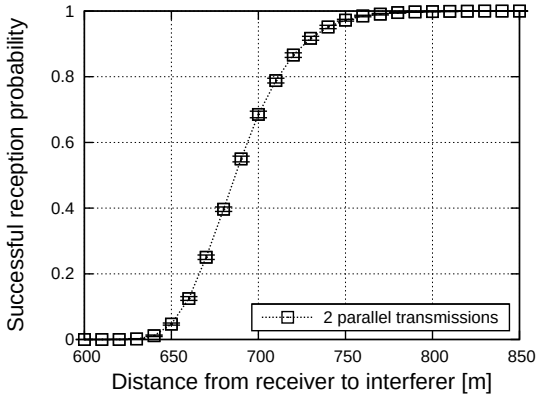


(b) N_n for the Manhattan Grid Mobility model

Figure (A.5) Comparison of neighbor densities (N_n) for the RWM and MGM models



(a) Node interference density



(b) Parallel reception probability

Figure (A.6) Interference characteristics of the radio channel model

conventional propagation models without shadowing effects show a different N_n already. This effect is intensified by adding the shadowing effects caused by obstacles.

Since the INET Framework considers noise and cumulative interference when determining the BER, the behavior of the model in relation to the SAT is important. Without taking shadowing into account the interference range amounts to just over 4000 m, thus, in the typical MGM scenario *all* nodes are within interference range of each other. This can be reduced significantly by adding the buildings to the path-loss calculation model (see Fig. A.6(a)), therefore, making the model much more realistic.

The radio model allows parallel transmissions. However, the distance between receiver and interferer has to exceed a certain range, otherwise no packet can be received. For the typical model parameters the parallel reception behavior is shown in Fig. A.6(b). In this case no obstacles obstruct the path between receiver and interferer, hence, the given result improves when obstacles shadow the interferer's signal to the receiver.

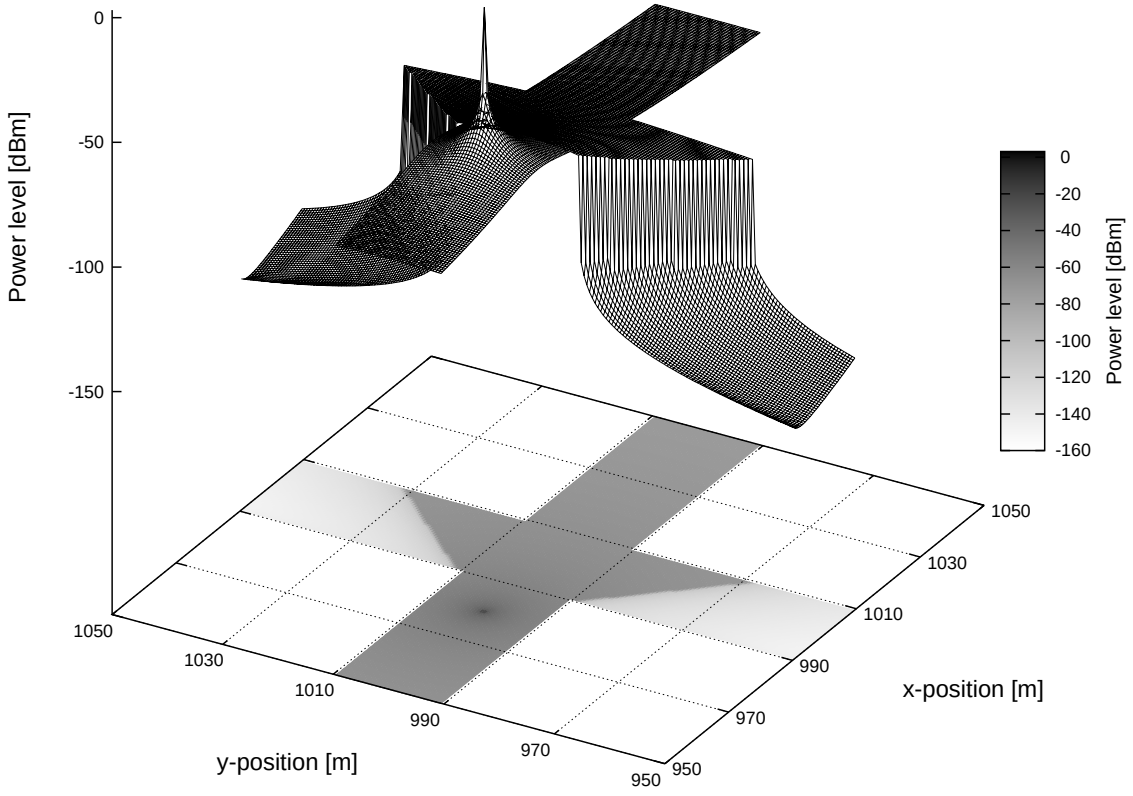


Figure (A.7) Signal strength diagram of a node approaching an intersection within the MGM model regarding building obstructions

The special characteristics of the propagation model considering buildings has been mentioned several times already. To make these attributes more clear an example is shown in Fig. A.7. A vehicle is approaching the center intersection in the MGM model. 10 m before entering the intersection a message is broadcast. Fig. A.7 shows the reception power level over the position in the scenario.

In the street of the sender a free-space signal decay is applied, since no obstacles shadow the signal. However, the signal in the crossing street shows a strong signal decay as soon as the signal has to pass the buildings next to the road. Therefore, receivers around the corner will only receive the message broadcast if they are close enough to the intersection.

A.4.3 Simulation Result Evaluation

Besides the models, a very important part of network simulation is the result evaluation. Since the results of any network simulator are a random sample statistical methods need to be used to evaluate the results. Usually one random sample is not enough to confidentially evaluate a new protocol using a network simulator. Thus, multiple simulation runs have to

be used, each initialized with different seed values for the RNG. The results of these runs are then combined by calculating the average results.

The so-called confidence interval is a valid method to be able to show the precision of results. Suppose $x_1, x_2, \ldots, x_n$ are independent random samples of a normally distributed process with mean μ and variance σ^2. Thus, the mean (Eqn. (A.12)), the standard deviation (Eqn. (A.14)), and the Student t-distribution quartiles ($t(1-\frac{\alpha}{2})$) [Gos08] can be used to calculate the confidence interval $1-\alpha = 0.95$, which theoretically contains 95% of all values (Eqn. (A.16)). The confidence intervals have been calculated using the GNU Scientific Library (GSL) and the respective cumulative distribution function to determine $t(1-\alpha/2)$ with $n-1$ as the degree of freedom. This mechanism helps to interprete the significance of simulation results.

All simulation results in this thesis have been evaluated according to this methodology. Thus, all plots containing confidence intervals show the interval $1-\alpha = 0.95$, theoretically containing 95% of all possible values.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{A.12}$$

$$x_{\mathrm{sq}} = \sum_{i=1}^{n} x_i^2 \tag{A.13}$$

$$\sigma = \sqrt{\frac{x_{sq} - \frac{x^2}{n}}{n-1}} \tag{A.14}$$

$$\mathrm{t} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{x}-\mu}{0.5} \tag{A.15}$$

$$1-\alpha = P\left(\bar{x} - \mathrm{t}\left(1-\frac{\alpha}{2}\right)\cdot\frac{\sigma}{\sqrt{n}} \leq \xi \leq \bar{x} + \mathrm{t}\left(1-\frac{\alpha}{2}\right)\cdot\frac{\sigma}{\sqrt{n}}\right) \tag{A.16}$$

Abbreviations

ABS	Anti Blocking System
AC	Access Class
AIFS	Arbitration Inter-Frame Space
AODV	Ad hoc On-Demand Distance Vector Routing
API	Application Programming Interface
ARP	Address Resolution Protocol
BC	Billing Center
BER	Bit Error Rate
BSP	Binary Space Partitioning
CA	Certificate Authority
CBF	Contention-Based Forwarding
CC	Control Center
CCH	Control Channel
CERT	Computer Emergency Response Team
CLC	Cross Layer Communication
CoRS	Content Reputation System
CRL	Certificate Revocation List
CRS	Certificate Revocation System
CRT	Certificate Revocation Tree

CS	Client System
CSI	Certificate Status Information
CTS	Clear-to-Send
CVIS	Cooperative Vehicle-Infrastructure Systems
CW	Contention Window
DAB	Digital Audio Broadcast
DAHNI	Driver Ad Hoc Networking Infrastructure
DCF	Distributed Coordination Function
DH	Diffie-Hellman Cryptosystem
DIFS	Distributed Coordination Function Inter-frame Space
DoS	Denial of Service
DSA	Digital Signature Algorithm
DSR	Dynamic Source Routing
DSRC	Dedicated Short Range Communication
DVB	Digital Video Broadcast
DWOP	Distributed Wireless Ordering Protocol
EC	European Commission
ECC	Elliptic Curve Cryptography
EDCA	Enhanced Distributed Channel Access
ELP	Electronic License Plate
ESP	Electronic Stability Program
ESRI	Environmental Systems Research Institute
EU	European Union
E2E	end-to-end
FCD	Floating Car Data
FIFO	First In, First Out
GDP	Gross Domestic Product
GNED	Graphical Network Editor

GPRS General Packet Radio Service

GPSR Greedy Perimeter Stateless Routing

GPS Global Positioning System

GSL GNU Scientific Library

GSM Global System for Mobile Communications

GST Global System for Telematics

HMI Human-Machine Interface

HTTP Hypertext Transfer Protocol

ICMP Internet Control Message Protocol

ICT Information and Communication Technology

IEEE Institute of Electrical and Electronics Engineers, Inc.

IP Internet Protocol

ISO International Standards Organization

ITS Intelligent Transportation System

IVC Inter-Vehicle Communication

LoS Line of Sight

MAC Medium Access Control

MANET Mobile Ad Hoc Network

MDRP Mobile Data Request Protocol

ME Mobile Entity

MGM Manhattan Grid Mobility

MPDU Message Protocol Data Unit

MPLS Multi-Protocol Label Switching

NED Network Description

NoW Network on Wheels

NS2 Network Simulator 2

OBD On-Board Diagnostics

OBU On-Board Unit

OEM	Original Equipment Manufacturer
OSGi	Open Services Gateway Initiative
OSI	Open Systems Interconnection
PDA	Personal Digital Assistant
PDF	Probability Density Function
PGP	Pretty Good Privacy
PKI	Public Key Infrastructure
PPP	Point-to-Point Protocol
PReVENT	PReVENTive and Active Safety Applications
PSK	Phase Shift Keying
P2P	point-to-point
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RA	Registration Authority
RDS	Radio Data System
RNG	Random-Number Generator
ROCQ	Reputation, Opinion, Credibility and Quality
RSA	Rivest, Shamir, & Adleman Public Key Encryption
RWM	Random Waypoint Mobility
RTS	Ready-to-Send
SAML	Security Assertion Markup Language
SAT	Signal Attenuation Threshold
SC	Service Center
SCH	Service Channel
SDL	Specification and Description Language
SeVeCom	Secure Vehicular Communication
SIM-TD	Sichere Intelligente Mobilität – Testfeld Deutschland
SINR	Signal to Interference-plus-Noise Ratio

SIP	Session Initiation Protocol
SP	Service Provider
SSL	Secure Socket Layer
SSO	Single Sign-On
STRAW	Street Random Waypoint
SUMO	Simulation of Urban Mobility
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TMC	Traffic Message Channel
TPEG	Transport Protocol Experts Group
TPM	Trusted Platform Module
UDP	User Datagram Protocol
UML	Universal Markup Language
UMTS	Universal Mobile Telecommunications Standard
VANET	Vehicular Ad Hoc Network
VARS	Vehicle Ad-Hoc Network Reputation System
VCP	Virtual City Portal
VN	Vehicular Network
V2B	vehicle-to-backend
V2V	vehicle-to-vehicle
V2I	vehicle-to-infrastructure
WAVE	Wireless Access in Vehicular Environments
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Network
ZRP	Zone Routing Protocol
3GT	3rd Generation Telematics

Mathematical Notations

t_{a}	AIFS period
$\mathcal{B}$	benefit function
t_{p}	channel period
t_{CRL}	certificate revocation list distribution interval
t_{CSI}	certificate status information period
t_{C}	certificate validity period
p_{coll}	collision probability
t_{c}	contention period
n_{cs}	number of contention window slots
$\mathcal{H}$	cryptographic hash function
$\mathcal{U}_{\mathrm{g}}$	global utility
d_{g}	grid length
c_{i}	information context
d_{Int}	interference distance
t_{i}	interval time
s	key share
t_{age}	MDRP aging interval
t_{wait}	MDRP wait interval
c_{m}	message context

$\mathcal{U}_\mathrm{n}$	network utility
t_int	node interaction time
t_ri	node re-interaction time
v_n	node speed
$\overline{\mathcal{U}_\mathrm{n}}$	normalized network utility
t_n	notification interval
N_a	number of attackers
N_c	number of context categories
N_m	number of messages
N_ne	number of neighbors
N_n	number of nodes
N_s	number of shares
K_priv	private key
t_pc	pseudonym change interval
K_pub	public key
t_q	quiet-time
d_r	radio range
R_r	revocation rate
t_R	revocation target
t_sim	simulation time
t_s	slot time
T	threshold
$K_\mathrm{ts,pub}$	threshold public key
Sig_ts	threshold signature
$\mathcal{U}$	utility function
t_V	validation target
c_v	vehicle context
t_w	wait time
w_j	weighing factors for benefit values

Bibliography

The references used in this dissertation have been grouped into three categories: publications by the author of the dissertation, publications from books, journals, and conference proceedings, and Internet references.

Publications by the Author

The references listed in this sub-bibliography have been previously published by the author according to the "Promotionsordnung der Technischen Universität München § 6 Abs. 1".

[AEK+06] Christian Adler, Stephan Eichler, Timo Kosch, Christoph Schroth, and Markus Strassberger. Self-organized and context-adaptive information diffusion in vehicular ad hoc networks. In *Proceedings of the International Symposium on Wireless Communication Systems (ISWCS)*, September 2006.

[EBM+05] Stephan Eichler, Jérôme Billion, Robert Maier, Hans-Jörg Vögel, and Rainer Kroh. On providing security for an open telematics platform. In *Proceedings of the 5th International Conference on ITS Telecommunications*, June 2005.

[EDS+04] Stephan Eichler, Florian Dötzer, Christian Schwingenschlögl, Jörg Eberspächer, and Francisco Javier Fabra Caro. Secure routing in a vehicular ad hoc network. In *Proceedings of the 2004 IEEE 60th Vehicular Technology Conference*, September 2004.

[EEH+07] Jörg Eberspächer, Stephan Eichler, Christian Hartmann, Silke Meister, Robert Nagel, Robert Vilzmann, and Hans-Martin Zimmermann. Wireless multi-hop networks: Classification, paradigms and constraints. Technical Report LKN-TR-4, Institute of Communication Networks, Technische Universität München, October 2007.

[Eic04] Stephan Eichler. Security challenges in MANET-based telematics environments. In *Proceedings of the 10th Open European Summer School and IFIP WG 6.3 Workshop (EUNICE)*, June 2004.

[Eic06] Stephan Eichler. Anonymous and authenticated data provisioning for floating car data systems. In *Proceedings of the 10th IEEE International Conference on Communication Systems (ICCS)*, October 2006.

[Eic07a] Stephan Eichler. MDRP: A content-aware data exchange protocol for mobile ad hoc networks. In *Proceedings of the IEEE International Symposium on Wireless Communication Systems (ISWCS)*, pages 742–746, October 2007.

[Eic07b] Stephan Eichler. Performance evaluation of the IEEE 802.11p WAVE communication standard. In *Proceedings of the 1st IEEE International Symposium on Wireless Vehicular Communications (WiVeC)*, September 2007.

[Eic07c] Stephan Eichler. A security architecture concept for vehicular network nodes. In *Proceedings of the 6th International Conference on Information, Communications and Signal Processing (ICICS)*, December 2007.

[Eic07d] Stephan Eichler. Strategies for pseudonym changes in vehicular ad hoc networks depending on node mobility. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 541–546, June 2007.

[EMR05] Stephan Eichler and Bernd Müller-Rathgeber. Performance analysis of scalable certificate revocation schemes for ad hoc networks. In *Proceedings of the 30th Conference on Local Computer Networks (LCN)*, November 2005.

[EMS06] Stephan Eichler, Christian Merkle, and Markus Strassberger. Data aggregation system for distributing inter-vehicle warning messages. In *Proceedings of the 31st Conference on Local Computer Networks (LCN)*, November 2006.

[EOSK05] Stephan Eichler, Benedikt Ostermaier, Christoph Schroth, and Timo Kosch. Simulation of car-to-car messaging: Analyzing the impact on road traffic. In *Proceedings of the 13th Annual Meeting of the IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, September 2005.

[ER06a] Stephan Eichler and Christian Roman. Challenges of secure routing in MANETs: A simulative approach using AODV-SEC. Technical Report LKN-TR-2, Institute of Communication Networks, Technische Universität München, August 2006.

[ER06b] Stephan Eichler and Christian Roman. Challenges of secure routing in MANETs: A simulative approach using AODV-SEC. In *Proceedings of the 3rd IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, October 2006.

[ES07] Stephan Eichler and Christoph Schroth. A multi-layer approach for improving scalability of vehicular ad-hoc networks. In *Proceedings of the 4th Workshop on Mobile Ad-Hoc Networks (WMAN)*, pages 503–514, March 2007.

[ESE06] Stephan Eichler, Christoph Schroth, and Jörg Eberspächer. Car-to-car communication. In *Proceedings of the VDE-Kongress - Innovations for Europe*, volume 1, pages 35–40, October 2006.

[ESKS06] Stephan Eichler, Christoph Schroth, Timo Kosch, and Markus Strassberger. Strategies for context-adaptive message dissemination in vehicular ad hoc networks. In *Proceedings of the Second International Workshop on Vehicle-to-Vehicle Communications (V2VCOM)*, July 2006.

[GE05] Ingo Gruber and Stephan Eichler. Path lifetime distributions of single- and multipath ad hoc routing strategies. In *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, July 2005.

[KAE+06] Timo Kosch, Christian J. Adler, Stephan Eichler, Christoph Schroth, and Markus Strassberger. The scalability problem of vehicular ad hoc networks and how to solve it. *IEEE Wireless Communications*, 13(5):22–28, October 2006.

[NE08] Robert Nagel and Stephan Eichler. Efficient and realistic mobility and channel modeling for VANET scenarios using OMNeT++ and INET-Framework. In *Proceedings of the 1st International Workshop on OMNeT++*. ACM Press, March 2008.

[NEE07] Robert Nagel, Stephan Eichler, and Jörg Eberspächer. Intelligent wireless communication for future autonomous and cognitive automobiles. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, June 2007.

[SE04] Christian Schwingenschlögl and Stephan Eichler. Certificate-based key management for secure communications in ad hoc networks. In *Proceedings of the 5th European Wireless Conference: Mobile and Wireless Systems beyond 3G*, pages 498–504, February 2004.

[SEMR06] Christian Schwingenschlögl, Stephan Eichler, and Bernd Müller-Rathgeber. Performance of PKI-based security mechanisms in mobile ad hoc networks. *AEÜ - Journal of Electronics and Communications*, 60(1):20–24, January 2006.

[SSEE06a] Christoph Schroth, Markus Strassberger, Robert Eigner, and Stephan Eichler. A framework for network utility maximization in VANETs. Technical Report LKN-TR-1, Institute of Communication Networks, Technische Universität München, August 2006.

[SSEE06b] Christoph Schroth, Markus Strassberger, Robert Eigner, and Stephan Eichler. A framework for network utility maximization in VANETs. In *Proceedings of the 3rd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, September 2006.

[VWE06] Hans-Jörg Vögel, Benjamin Weyl, and Stephan Eichler. Federation solutions for inter- and intradomain security in next-generation mobile service platforms. *AEÜ – Journal of Electronics and Communications*, 60(1):13–19, January 2006.

General Bibliography

[ABD+06] Amer Aijaz, Bernd Bochow, Florian Dötzer, Andreas Festag, Matthias Gerlach, Rainer Kroh, and Tim Leinmüller. Attacks on inter-vehicle communication systems – An analysis. In *Proceedings of the 3rd International Workshop on Intelligent Transportation (WIT)*, March 2006.

[ACG+07] F. Anjum, S. Choi, V. D. Gligor, R. G. Herrtwich, J.-P. Hubaux, P. R. Kumar, R. Shorey, and Lea Chin-Tau. Guest editorial vehicular networks. *IEEE Journal on Selected Areas in Communications*, 25(8):1497–1500, October 2007.

[AGMV96] P. K. Agarwal, E. F. Grove, T. M. Murali, and J. S. Vitter. Binary space partitions for fat rectangles. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 482–491, Washington, DC, USA, 1996. IEEE Computer Society.

[AK96] Ross Anderson and Markus Kuhn. Tamper resistance – a cautionary note. In *Proceedings of the 2nd USENIX Workshop on Electronic Commerce*, pages 1–11, November 1996.

[AMV97] Pankaj K. Agarwal, T. M. Murali, and Jeffrey Scott Vitter. Practical techniques for constructing binary space partitions for orthogonal rectangles. In *Proceedings of the Thirteenth Annual Symposium on Computational Geometry (SCG)*, pages 382–384, New York, NY, USA, 1997. ACM.

[BB02b] Sonja Buchegger and Jean-Yves Le Boudec. Performance analysis of the CONFIDANT protocol. In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc)*, pages 226–236, New York, NY, USA, June 2002. ACM Press.

[BB03] Sonja Buchegger and Jean-Yves Le Boudec. A robust reputation system for mobile ad-hoc networks. Technical Report IC/2003/50, Ecole Polytechnique Fédérale de Lausanne, EPFL-IC-LCA, CH-1015 Lausanne, Switzerland, 2003.

[BBC+02] Chatschik Bisdikian, Isaac Boamah, Paul Castro, Archan Misra, Jim Rubas, Nicolas Villoutreix, Danny Yeh, Vladimir Rasin, Henry Huang, and Craig Simonds. Intelligent pervasive middleware for context-based and localized telematics services. In *WMC '02: Proceedings of the 2nd International Workshop on Mobile Commerce*, pages 15–24. ACM Press, 2002.

[BEF+00] Lee Breslau, Deborah Estrin, Kevin Fall, Sally Floyd, John Heidemann, Ahmed Helmy, Polly Huang, Steven McCanne, Kannan Varadhan, Ya Xu, and Haobo Yu. Advances in network simulation. *IEEE Computer*, 33(5):59–67, May 2000. Expanded version available as USC TR 99-702b at `http://www.isi.edu/~johnh/PAPERS/Bajaj99a.html`.

[Bet01] Christian Bettstetter. Smooth is better than sharp: A random mobility model for simulation of wireless networks. In *Proceedings of the 4th ACM International*

Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWIM), pages 19–27, New York, NY, USA, 2001. ACM.

[Bet03] Christian Bettstetter. Topology properties of ad hoc networks with random waypoint mobility. In *Proceedings of the Fourth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, volume 7, pages 50–52. ACM Press, 2003.

[BFL96] Matt Blaze, Joan Feigenbaum, and Jack Lacy. Decentralized trust management. In *Proceedings of the IEEE Symposium on Security and Privacy*, 1996.

[BFW03] Marc Bechler, Walter J. Franz, and Lars Wolf. Mobile internet access in fleetnet. In *Proceedings of the Fachtagung Kommunikation in verteilten Systemen (KivS)*, 2003.

[BKK+03] Richard Bogenberger, Wolfgang Kellerer, Timo Kosch, Thomas Reicher, Christian Schwingenschlögl, Peter Sties, and Matthias Wagner. Virtual city portal – a multi-network personal information system for automobile users. In *Proceedings of Workshop on Multiradio Multimedia Communications, Communication Technology for Vehicles*, February 2003.

[BLB05] S. Buchegger and J.-Y. Le Boudec. Self-policing mobile ad hoc networks by reputation systems. *IEEE Communications Magazine*, 43(7):101–107, July 2005.

[BMJ+98] Josh Broch, David A. Maltz, David B. Johnson, Yih-Chun Hu, and Jorjeta Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proceedings of MobiCom*, 1998.

[Bou04] Azzedine Boukerche. Performance evaluation of routing protocols for ad hoc wireless networks. *Mobile Networks and Applications*, 9(4):333–342, August 2004.

[BP96] Lawrence S. Brakmo and Larry L. Peterson. Experiences with network simulation. In *Proceedings of the International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '96)*, pages 80–90. ACM Press, 1996.

[BS03] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *IEEE Pervasive Computing*, 2(1):46–55, January 2003.

[BSMM99] I. N. Bronstein, K. A. Semendjajew, G. Musiol, and H. Mühlig. *Handbook of Mathematics*. Harri Deutsch, 4th edition, 1999.

[BV05] Jean-Yves Le Boudec and Milan Vojnovic. Perfect simulation and stationarity of a class of mobility models. In *Proceedings of the IEEE Infocom*, March 2005.

[BV06] Albert Banchsa and Luca Vollero. Throughput analysis and optimal configuration of 802.11e EDCA. *Computer Networks*, 50(11):1749–1768, August 2006.

[BWSF03] Marc Bechler, Lars Wolf, Oliver Storz, and Walter Franz. Efficient discovery of internet gateways in future vehicular communication systems. In *Proceedings of the 57th IEEE Semiannual Vehicular Tecnology Conference (VTC 2003 Spring)*, April 2003.

[CB05a] David Choffnes and Fabián E. Bustamante. STRAW – an integrated mobility and traffic model for vanets. In *Proceedings of the 10th International Command and Control Research and Technology Symposium (CCRTS)*, June 2005.

[CB05b] David R. Choffnes and Fabián E. Bustamante. An integrated mobility and traffic model for vehicular wireless networks. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad-Hoc Networks (VANET)*, August 2005.

[CBD02] Tracey Camp, Jeff Boleng, and Vanessa Davies. A survey of mobility models for ad hoc network research. *Wireless Communication & Mobile Computing (WCMC)*, 2(5):483–502, September 2002.

[CBH03b] Srdjan Capkun, Levente Buttyan, and Jean-Pierre Hubaux. Self-organized public-key management for mobile ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(1):52–64, January 2003.

[CD03] Claude Crepeau and Carlton R. Davis. A certificate revocation scheme for wireless ad hoc networks. In *Proceedings of the 1st ACM Workshop Security of Ad Hoc and Sensor Networks*, pages 54–61. ACM Press, 2003.

[Cha81] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–90, February 1981.

[Cha88] David Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1(1):65–75, January 1988.

[CHD05] E. Chang, F.K. Hussain, and T.S. Dillon. Towards defining an ontology for reputation. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2601–2607, October 2005.

[CJTD06] Qi Chen, Daniel Jiang, Vikas Taliwal, and Luca Delgrossi. Ieee 802.11 based vehicular communication simulation design for ns-2. In *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 50–56, September 2006.

[CKV01] Zong Da Chen, H.T. Kung, and Dario Vlah. Ad hoc relay wireless networks over moving vehicles on highways. In *Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing*, pages 247–250, New York, NY, USA, 2001. ACM Press.

[Com03] Commission of the European Communities. Information and communications technologies for safe and intelligent vehicles. Communication from the Commision to the Council and the European Parliament COM(2003) 542, Commission of the European Communities, September 2003.

[Com06] Committee SCC32. *IEEE P1609.4 Standard for Wireless Access in Vehicular Environments (WAVE) – Multi-Channel Operation*. IEEE Intelligent Transportation Systems Council, draft standard edition, 2006.

[Cox95] D.C. Cox. Wireless personal communications: What is it? *IEEE Personal Communications*, 2(2):20–35, April 1995.

[CRW04] Antonio Carzaniga, Matthew J. Rutherford, and Alexander L. Wolf. A routing scheme for content-based networking. In *Proceedings of IEEE INFOCOM*, March 2004.

[CS92] S. T. S. Chia and P. Snow. Characterising radio-wave propagation behaviour at 1700 MHz for urban and highway microcells. In *IEE Colloquium on Micro-Cellular Propagation Modelling*, November 1992.

[CS02] Ioan Chisalita and Nahid Shahmehri. A novel architecture for supporting vehicular communication. In *In Proceedings of the 56th Vehicular Technology Conference*, volume 2, pages 1002–1006, 2002.

[CW94] W.C. Collier and R.J. Weiland. Smart cars, smart highways. *IEEE Spectrum*, 31(4):27–33, April 1994.

[CW01] Antonio Carzaniga and Alexander L. Wolf. Content-based networking: A new communication infrastructure. In *Proceedings of the NSF Workshop on an Infrastructure for Mobile and Wireless Systems*, number 2538 in Lecture Notes in Computer Science. Springer-Verlag, October 2001.

[CWKS97] Brian P. Crow, Indra Widjaja, Jeong Geun Kim, and Prescott T. Sakai. IEEE 802.11 wireless local area networks. *IEEE Communications Magazine*, 35(9):116–126, September 1997.

[Deu58] Morton Deutsch. Trust and suspicion. *Journal of Conflict Resolution*, 2(4):265–279, 1958.

[DF89] Yvo G. Desmedt and Yair Frankel. Threshold cryptosystems. In *Proceedings on Advances in cryptology (CRYPTO)*, pages 307–315. Springer-Verlag, 1989.

[DFM05] Florian Dötzer, Lars Fischer, and Przemyslaw Magiera. Vars: A vehicle ad-hoc network reputation system. In *Proceedings of the 6th International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM)*, 2005.

[DGL+02] Sastry Duri, Marco Gruteser, Xuan Liu, Paul Moskowitz, Ronald Perez, Moninder Singh, and Jung-Mu Tang. Framework for security and privacy in automotive telematics. In *Proceedings of the 2nd International Workshop on Mobile Commerce*, pages 25–32. ACM Press, 2002.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, IT-22(6):644–654, November 1976.

[DH79] W. Diffie and M.E. Hellman. Privacy and authentication: An introduction to cryptography. *Proceedings of the IEEE*, 67(3):397–427, March 1979.

[DRS06] Dominique Dhoutaut, Anthony Régis, and François Spies. Impact of radio propagation models in vehicular ad hoc networks simulations. In *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 40–49, New York, NY, USA, 2006. ACM Press.

[DSCP02] Claudia Díaz, Stefaan Seys, Joris Claessens, and Bart Preneel. Towards measuring anonymity. In Hannes Federath, editor, *Proceedings of Workshop on Privacy Enhancing Technologies (PET)*, Lecture Notes in Computer Science, April 2002.

[Döt05b] Florian Dötzer. Privacy issues in vehicular ad hoc networks. In *Proceedings of the Conference on Privacy Enhancing Technologies*, pages 197–209, 2005.

[EB05] Stephan Eichler and Jérôme Billion. Architecture and interface specifications. Project Deliverable (Security) 3.1, European IP – Global System for Telematics, September 2005. Available at: `http://www.gstforum.org/en/downloads/deliverables/`.

[Eck06] Claudia Eckert. *IT-Sicherheit*. Oldenbourg Verlag, 4th edition, 2006.

[EGR04] F. Elwailly, C. Gentry, and Z. Ramzan. Quasimodo: Efficient certificate validation and revocation. In *Proceedings of Public-Key Cryptography*, 2004.

[ES00] C. Ellison and B. Schneier. Ten risks of PKI: What you're not being told about public key infrastructure. *Computer Security Journal*, 16(1):1–7, 2000.

[EU02] Thierry Ernst and Keisuke Uehara. Connecting automobiles to the internet. In *Proceedings of the 3rd International Workshop on ITS Telecommunications*, 2002.

[EV99] Jörg Eberspächer and Hans-Jörg Vögel. *GSM Global System for Mobile Communication*. B. G. Teubner, Stuttgart - Leipzig, 2nd edition, 1999.

[FC05] M. Fazel and Mung Chiang. Network utility maximization with nonconcave utilities using sum-of-squares method. In *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, pages 1867–1874, December 2005.

[FCD06] Fabio Ferro, Cristina Chesta, and Enrico D'Acquisto. High level architecture. Project Deliverable 3.1, European IP – Global System for Telematics, October 2006. Available at: `http://www.gstforum.org/en/downloads/deliverables/`.

[Fed07] Federal Statistical Office. *Statistical Yearbook 2007 – For the Federal Republic of Germany*. Federal Statistical Office, September 2007.

[FEL01] W. Franz, R. Eberhardt, and T. Luckenbach. FleetNet – Internet on the road. In *Proceedings of the 8th World Congress on Intelligent Transportation Systems (ITS)*, October 2001.

[FMH+02] Holger Füßler, Martin Mauve, Hannes Hartenstein, Michael Käsemann, and Dieter Vollmer. A comparison of routing strategies for vehicular ad hoc networks. Technical Report TR-02-003, Department of Computer Science, University of Mannheim, July 2002.

[Fra04] Dr. Walter Franz. Car-to-Car Communication – Anwendungen und aktuelle Forschungsprogramme in Europa, USA und Japan. In *Kongressband zum VDE-Kongress 2004 – Innovationen für Menschen*. VDE, October 2004.

[Fri46] H. T. Friis. A note on a simple transmission formula. *Proceedings of the Institute of Radio Engineers (IRE)*, 34(5):254–256, May 1946.

[FSTE07] H. Füßler, S. Schnaufer, M. Transier, and W. Effelsberg. Vehicular ad-hoc networks: From vision to reality and back. In *Proceedings of the 4th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS)*, January 2007.

[FTTM+05] H. Füßler, M. Transier, M. Torrent-Moreno, H. Hartenstein, and A. Festag. Thoughts on a protocol architecture for vehicular ad-hoc networks. In *Proceedings of the 2nd International Workshop on Intelligent Transportation (WIT)*, March 2005.

[FWK+03] Holger Füßler, Jörg Widmer, Michael Käsemann, Martin Mauve, and Hannes Hartenstein. Contention-based forwarding for mobile ad hoc networks. *Ad Hoc Networks*, 1(4):351–369, November 2003.

[FWMH04] W. Franz, C. Wagner, C. Maihöfer, and H. Hartenstein. FleetNet: Platform for inter-vehicle communications. In *Procedings of 1st International Workshop on Intelligent Transportation*, 2004.

[FZZ06] Elena Fasolo, Andrea Zanella, and Michele Zorzi. An effective broadcast scheme for alert message propagation in vehicular ad hoc networks. In *Proceedings of the International Conference on Communications (ICC)*, pages 3960–3965, June 2006.

[GAB+08] Matthias Goebl, Matthias Althoff, Martin Buss, Georg Färber, Falk Hecker, Bernd Heißing, Sven Kraus, Robert Nagel, Fernando Puente León, Florian Rattei, Martin Russ, Michael Schweitzer, Michael Thuy, Cheng Wang, and Hans Joachim Wuensche. Design and capabilities of the munich cognitive automobile. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 1101–1107. IEEE Press, June 2008.

[GC05] A. Garg and R. Cascella. Reputation management for collaborative content distribution. In *Proceedings of the 6th IEEE International Symposium on a World*

of Wireless Mobile and Multimedia Networks (WoWMoM), pages 547–552, June 2005.

[Ger07] M. Gerlach. Trust for vehicular applications. In *Proceedings of the 8th International Symposium on Autonomous Decentralized Systems (ISADS)*, March 2007.

[GFL+07] M. Gerlach, A. Festag, T. Leinmüller, G. Goldacker, and C. Harsch. Security architecture for vehicular communication. In *Proceedings of the 4th International Workshop on Intelligent Transportation (WIT)*, March 2007.

[GG07] M. Gerlach and F. Güttler. Privacy in VANETs using changing pseudonyms – Ideal and real. In *Proceedings of the 65th IEEE Vehicular Technology Conference (VTC)*, April 2007.

[GK00] Piyush Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[GKL04] Ingo Gruber, Oliver Knauf, and Hui Li. Performance of ad hoc routing protocols in urban environments. In *Proceedings of the 5th European Wireless Conference (EW)*, pages 331–337, February 2004.

[GM02] Dr Lutz Gollan and Prof. Dr. Christoph Meinel. Digital signatures in automobiles?! In *Proceedings of the SCI*, Orlando, Florida, USA, 2002.

[Gos08] William Sealy Gosset. The probable error of a mean. *Biometrica*, 6(1):1–25, March 1908. Author also known as Student.

[GPZ04] Tao Gu, Hung Keng Pung, and Da Quing Zhang. Toward an OSGi-based infrastructure for context-aware applications. *IEEE Pervasive Computing*, 3(4):66–74, December 2004.

[GPZW04] Tao Gu, Hung Keng Pung, Da Qing Zhang, and Xiao Hang Wang. A middleware for context aware mobile services. In *Proceedings of the IEEE Vehicular Technology Conference (VTC Spring 2004)*, Milan, Italy, May 2004.

[GST05] GST Security Subproject. The GST security white paper. White paper, EU IP – Global System for Telematics, July 2005. Available at: `http://www.gstforum.org/en/downloads/white_papers/`.

[GWP06] T. Giuli, D. Watson, and K.V. Prasad. The last inch at 70 miles per hour. *IEEE Pervasive Computing*, 5(4):20–27, October 2006.

[Haa01] Zygmunt J. Haas. ZRP: A hybrid framework for routing in ad hoc networks. In Charles E. Perkins, editor, *Ad Hoc Networking*, chapter 7. Addison–Wesley, 2001.

[HBC01] J. Hubaux, L. Buttyan, and S. Capkun. The quest for security in mobile ad hoc networks. In *Proceeding of the ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*. ACM Press, 2001.

[HBE+01a] Hannes Hartenstein, Bernd Bochow, André Ebner, Matthias Lott, Markus Radimirsch, and Dieter Vollmer. Position-aware ad hoc wireless networks for inter-vehicle communications: the fleetnet project. In *Proceedings of the 2nd International Symposium on Mobile Ad Hoc Networking & Computing*, pages 259–262, 2001.

[HCL04] Jean Pierre Hubaux, Srdjan Capkun, and Jun Luo. The security and privacy of smart vehicles. *IEEE Security & Privacy*, 4(3):49–55, May 2004.

[Her04] Ralf G. Herrtwich. Automotive telematics – road safety versus IT security? In *Proceedings of the 23rd International Conference on Computer Safety, Reliability, and Security (SAFECOMP)*. Springer Berlin / Heidelberg, 2004.

[HFB05] Jérôme Härri, Fethi Filali, and Christian Bonnet. A framework for mobility models generation and its application to inter-vehicular networks. In *Proceedings of the 3rd IEEE International Workshop on Mobility Management and Wireless Access*, June 2005.

[HGK+04] Y. Hörmann, H.P. Großmann, W.H. Khalifa, M. Salah, and O.H. Karam. Simulator for inter-vehicle communication based on traffic modeling. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, June 2004.

[HK02a] Albert Held and Rainer Kroh. IT-security and privacy for telematics-services. In *Proceedings of the PAMPAS 02 — Workshop on Requirements for Mobile Privacy & Security*, July 2002.

[HL02] Lifei Huang and Ten-Hwang Lai. On the scalability of IEEE 802.11 ad hoc networks. In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc)*, pages 173–182, New York, NY, USA, June 2002. ACM Press.

[HLO99] Werner Huber, Michael Lädke, and Rainer Ogger. Extended floating-car data for the acquisition of traffic information. In *Proceedings of the 6th World Congress on Intelligent Transport Systems*, 1999.

[HMK01] John Heidemann, Kevin Mills, and Sri Kumar. Expanding confidence in network simulation. *IEEE Network Magazine*, 15(5):58–63, September 2001.

[HMYS05] Leping Huang, Kanta Matsuura, Hiroshi Yamane, and Kaoru Sezaki. Enhancing wireless location privacy using silent period. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, March 2005.

[Hua06] Dijiang Huang. On measuring anonymity for wireless mobile ad-hoc networks. In *Proceedings of the 2nd International Workshop on Performance and Management of Wireless and Mobile Networks (P2MNet)*, pages 779–786, November 2006.

[IEE05] IEEE P1609.2 Working Group. *IEEE P1609.2 Standard for Wireless Access in Vehicular Environments (WAVE) – Security Services for Applications and Management Messages*. IEEE Intelligent Transportation Systems Council, draft standard, 3rd edition, November 2005.

[JM96] David B Johnson and David A Maltz. *Mobile Computing*, volume 353, chapter Dynamic Source Routing in Ad Hoc Wireless Networks, pages 153–181. Kluwer Academic Publishers, 1996.

[JTM+06] D. Jiang, V. Taliwal, A. Meier, W. Holfelder, and R. Herrtwich. Design of 5.9 GHz DSRC-based vehicular safety communication. *IEEE Wireless Communications*, 13(5):36–43, October 2006.

[Juh03] John Juhasz. 3GT white paper. Project white paper, European IP — Third Generation Telematics (3GT), June 2003.

[KAN99] H. Kikuchi, K. Abe, and S. Nakanishi. Performance evaluation of public-key certificate revocation systemwith balanced hash tree. In *Proceedings of the International Workshops on Parallel Processing*, pages 204–209, September 1999.

[Kar03] Frank Kargl. *Sicherheit in Mobilen Ad hoc Netzwerken*. Phd thesis, Universität Ulm, October 2003.

[KBS+01] Wolfgang Kellerer, Christian Bettstetter, Christian Schwingenschlögl, Peter Sties, Karl-Ernst-Steinberg, and Hans-Jörg Vögel. (Auto)mobile communication in a heterogeneous and converged world. *IEEE Personal Communications Magazine*, 8(6):41–47, December 2001.

[KC98] Anil Kini and Joobin Choobineh. Trust in electronic commerce: Definition and theoretical considerations. In *Proceedings of the 31st Annual Hawaii International Conference on System Sciences*, volume 4. IEEE Computer Society, 1998.

[KHRW02] Daniel Krajzewicz, Georg Hertkorn, Christian Rössel, and Peter Wagner. Sumo (simulation of urban mobility); an open-source traffic simulation. In *Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM2002)*, September 2002.

[KHRW05] Daniel Krajzewicz, Georg Hertkorn, Julia Ringel, and Peter Wagner. Preparation of digital maps for traffic simulation; part 1: Approach and algorithms. In *Proceedings of the 3rd Industrial Simulation Conference*, June 2005.

[KK00] Brad Karp and H. T. Kung. GPSR: Greedy perimeter stateless routing for wireless networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*. ACM Press, August 2000.

[KLS+02] V. Kanodia, C. Li, A. Sabharwal, B. Sadeghi, and E. Knightly. Ordered packet scheduling in wireless ad hoc networks: Mechanisms and performance analysis. In *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking & Computing*. ACM Press, June 2002.

[Koc98] Paul C. Kocher. On certificate revocation and validation. In *Proceedings of the Second International Conference on Financial Cryptography*, pages 172–177. Springer Verlag, 1998.

[Kos05] Timo Kosch. *Situationsadaptive Kommunikation in Automobilen Ad-hoc Netzen*. Phd thesis, Technische Universität München, March 2005.

[Kra98] Stefan Krauß. *Microscopic Modelling of Traffic Flow: Investigation of Collision Free Vehicle Dynamics*. Phd thesis, Universität Köln, April 1998.

[KRM07] Wolfgang Kiess, Jedrzej Rybicki, and Martin Mauve. On the nature of inter-vehicle communication. In *Proceedings of the 4th Workshop on Mobile Ad-Hoc Networks (WMAN)*, pages 493–502, March 2007.

[KSB02] Timo Kosch, Christian Schwingenschlögl, and Christian Bettstetter. Situative IP-basierte Fahrerinformationssysteme: Szenarien, Routing und prototypische Realisierung. In *Proceedings of the VDE Kongress 2002 – Networlds*. VDE, October 2002.

[KT75] Leonard Kleinrock and Fouad A. Tobagi. Packet switching in radio channels: Part I – carrier sense multiple-access modes and their throughput-delay characteristics. *IEEE Transactions On Communications*, 23(12):1400–1416, December 1975.

[KV99] Young-Bae Ko and Nitin H. Vaidya. Geocasting in mobile ad hoc networks: Location-based multicast algorithms. In *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, 1999.

[KV00] Young-Bae Ko and Nitin H. Vaidya. GeoTORA: A protocol for geocasting in mobile ad hoc networks. In *Proceedings of the 8th International Conference on Networking Protocols (ICNP)*, November 2000.

[KVS02] Wolfgang Kellerer, Hans-Jörg Vögel, and Karl-Ernst Steinberg. A communication gateway for infrastructure-independent 4G wireless access. *IEEE Communications Magazine*, 40(3):126–131, March 2002.

[KWG97] Stefan Krauß, Peter Wagner, and Christian Gawron. Metastable states in a microscopic model of traffic flow. *Physical Review E*, 55(304):55–97, May 1997.

[LAN99] LAN/MAN Standards Committee. *Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*. IEEE Computer Society, September 1999.

[LBC+01] Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee, and Robert Morris. Capacity of ad hoc wireless networks. In *Proceedings of the 7th ACM International Conference on Mobile Computing and Networking (MobiCom '01)*, Rome, Italy, July 2001.

[LBC+05] Christian Lochert, Andreas Barthels, Alfonso Cervantes, Martin Mauve, and Murat Caliskan. Multiple simulator interlinking environment for IVC. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 87–88. ACM Press, 2005.

[LH04] Jun Luo and Jean-Pierre Hubaux. A survey of inter-vehicle communication. Technical Report IC/2004/24, School of Computer and Communication Sciences EPFL, CH-1015 Lausanne, Switzerland, 2004.

[LI04] Jinshan Liu and Valérie Issarny. Enhanced reputation mechanism for mobile ad hoc networks. In *Proceedings of the 2nd International Conference on Trust Management*, pages 48–62. Springer Berlin / Heidelberg, March 2004.

[LL99] Chunhung Richard Lin and Jain-Shing Liu. QoS routing in ad hoc wireless networks. *Journal on Selected Areas in Communications*, 17(8):1426–1438, August 1999.

[LN98] Ilari Lehti and Pekka Nikander. Certifying trust. In *Proceedings of the Practice and Theory in Public Key Cryptography (PKC) '98*. Springer-Verlag, February 1998.

[Lüb04] Andreas Lübke. Car-to-Car Communication – Technologische Herausforderungen. In *Kongressband zum VDE-Kongress 2004 – Innovationen für Menschen*. VDE, October 2004.

[Mau05] Jürgen Maurer. *Strahlenoptisches Kanalmodell für die Fahrzeug-Fahrzeug Kommunikation*. Phd thesis, Universität Karlsruhe (TH), May 2005.

[MB05] Marcin Michalak and Torsten Braun. Common gateway architecture for mobile ad-hoc networks. In *Proceedings of the Second Annual Conference on Wireless On-demand Network Systems and Services*, pages 70–75, January 2005.

[MC04] Christian Maihöfer and Luca Coletti. Cartalk 2000: Results of communication technology and applications for automotive. In *Proceedings of the IEEE Vehicular Technology Conference*, May 2004.

[Mer79] Ralph Charles Merkle. *Secrecy, Authentication, and Public Key Systems*. Phd thesis, Department of Electrical Engineering, Stanford University, 1979.

[Mer80] R. C. Merkle. Protocols for public key cryptography. In *Proceedings of the IEEE Symposium on Security and Privacy*, pages 122–134, 1980.

[MFSW04] Jürgen Maurer, Thomas Fügen, Thomas Schäfer, and Werner Wiesbeck. A new inter-vehicle communications (IVC) channel model. In *Proceedings of the 61st Vehicular Technology Conference (VTC)*, September 2004.

[MGLB00] Sergio Marti, T. J. Giuli, Kevin Lai, and Mary Baker. Mitigating routing misbehavior in mobile ad hoc networks. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pages 255–265, New York, NY, USA, 2000. ACM.

[Mic95] Silvio Micali. Enhanced certificate revocation. Technical Report MIT/LCS/TM-542, Massachusetts Institute of Technology, 1995.

[Mic96] Silvio Micali. Efficient certificate revocation. Technical Report MIT/LCS/TM 542b, Massachusetts Institute of Technology, Cambridge, MA, USA, 1996.

[Mic97] Silvio Micali. Efficient certificate revocation. In *Proceedings of the RSA Data Security Conference*, 1997.

[Mic02] Silvio Micali. NOVOMODO: Scalable certificate validation and simplified PKI management. In *Proceedings of the First Annual PKI Research Workshop*, April 2002.

[MJK+00] Robert Morris, John Jannotti, Frans Kaashoek, Jinyang Li, and Douglas Decouto. Carnet: A scalable ad hoc wireless network system. In *Proceedings of the 9th ACM SIGOPS European workshop: Beyond the PC: New Challenges for the Operating System*. ACM Press, September 2000.

[MK01] Dave Marples and Peter Kriens. The open services gateway initiative: An introductory overview. *Communications Magazine*, 39(12):110–114, December 2001.

[MLS05] Tony K. Mak, Kenneth P. Laberteaux, and Raja Sengupta. A multi-channel VANET providing concurrent safety and commercial services. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 1–9, New York, NY, USA, 2005. ACM Press.

[MM02] Pietro Michiardi and Refik Molva. CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks. In *Proceedings of the Sixth Joint Working Conference on Communications and Multimedia Security*, pages 107–121, Deventer, The Netherlands, The Netherlands, 2002. Kluwer, B.V.

[MML04] K. Matheus, R. Morich, and A. Lübke. Economic background of car-to-car communications. In *Proceedings of Informationssysteme für mobile Anwendungen (IMA)*, October 2004.

[MN98] M. Matsumoto and T. Nishimura. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulations*, 8(1):3–30, January 1998.

[MR04] Ajay Mahimkar and Theodore S. Rappaport. SecureDAV: A secure data aggregation and verification protocol for sensor networks. In *Proceedings of the 47th IEEE Global Telecomunications Conference (GLOBECOM)*, November 2004.

[MvOV96] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, October 1996. `http://www.cacr.math.uwaterloo.ca/hac/`.

[MWR+06] Rahul Mangharam, Daniel Weller, Raj Rajkumar, Priyantha Mudalige, and Fan Bai. Groovenet: A hybrid simulator for vehicle-to-vehicle networks. In *Proceedings of the 2nd International Workshop on Vehicle-to-Vehicle Communications (V2VCOM)*, July 2006.

[MWSR05] Rahul Mangharam, Daniel S. Weller, Daniel D. Stancil, and Ragunathan Rajkumar. GrooveSim: A topology-accurate simulator for geographic routing in vehicular networks. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks*, pages 59–68, September 2005.

[Nak60] M. Nakagami. The m-distribution, a general formula of intensity distribution of the rapid fading. In W. G. Hoffman, editor, *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium held at the University of California*, pages 3–36. Permagon Press, 1960.

[Nam05] H. Namin, A.S.; Ruizhong Wei; Weiming Shen; Ghenniwa. Applying secret sharing schemes to service reputation. In *Proceedings of the 9th International Conference on Computer Supported Cooperative Work in Design*, volume 2, pages 696–703, May 2005.

[NDLI04] Tamer Nadeem, Sasan Dashtinezhad, Chunyuan Liao, and Liviu Iftode. TrafficView: A scalable traffic monitoring system. In *Proceedings of 2004 IEEE International Conference on Mobile Data Management (MDM)*, pages 13–26, January 2004.

[NI97] Julio C. Navas and Tomasz Imielinski. Geocast – geographic addressing and routing. In *Proceedings of the 3rd Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 66–76, New York, NY, USA, 1997. ACM Press.

[NN00] M. Naor and K. Nissim. Certificate revocation and certificate update. *IEEE Journal on Selected Areas in Communications*, 18(4):561–570, April 2000.

[NTCS99] Sze-Yao Ni, Yu-Chee Tseng, Yuh-Shyan Chen, and Jang-Ping Sheu. The broadcast storm problem in a mobile ad hoc network. In *Proceedings of the 5th International Conference on Mobile Computing and Networking*, pages 151–162, 1999.

[OPB00] Karine Olmos, Samuel Pierre, and Yves Boudreault. A traffic simulator for urban mobile networks. In *Proceedings of the Canadian Conference on Electrical and Computer Engineering*, volume 2, pages 1042–1046, 2000.

[PCS07] Jun Peng, Liang Cheng, and Biplab Sikdar. A wireless MAC protocol with collision detection. *IEEE Transactions on Mobile Computing*, 6(12):1357–1369, December 2007.

[Per99] Radia Perlman. An overview of PKI trust models. *IEEE Network*, 13(6):38–43, November 1999.

[Per01] Charles E. Perkins. Summary and future work. In Charles E. Perkins, editor, *Ad Hoc Networking*, chapter 11. Addison–Wesley, 2001.

[PH07] Andreas Pfitzmann and Marit Hansen. *Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology*. TU Dresden and ULD Kiel, 0.31 edition, February 2007.

[PJFY06] Anand Patwardhan, Anupam Joshi, Tim Finin, and Yelena Yesha. A data intensive reputation management scheme for vehicular ad hoc networks. In *Proceedings of the 2nd International Workshop on Vehicle-to-Vehicle Communications (V2VCOM)*, pages 1–8, July 2006.

[PKHK06] P. Papadimitratos, A. Kung, J.-P. Hubaux, and F. Kargl. Privacy and identity management for vehicular communication systems: A position paper. Technical report, EU Project SeVeCom, 2006.

[PNM06] Klaus Plößl, Thomas Nowey, and Christian Mletzko. Towards a security architecture for vehicular ad hoc networks. In *Proceedings of The 1st International Conference on Availability, Reliability and Security (ARES)*, pages 374–381, 2006.

[PP05] Bryan Parno and Adrian Perrig. Challenges in securing vehicular networks. In *Proceedings of the 4th Workshop on Hot Topics in Networks (NotNets)*, November 2005.

[Pra03] Sojen Pradhan. Mobile commerce in the automobile industry. In *Proceedings of the International Conference on Information Technology: Computers and Communications (ITCC)*, pages 276–280, April 2003.

[Pre99] Bart Preneel. *Lectures on Data Security: Modern Cryptology in Theory and Practice*, volume 1561, chapter The State of Cryptographic Hash Functions, pages 158–182. Springer-Verlag, Heidelberg, 1999.

[PSP03] Bartosz Przydatek, Dawn Song, and Adrian Perrig. SIA: Secure information aggregation in sensor networks. In *Proceedings of the 1st ACM Conference on Embedded Networked Sensor Systems (SenSys)*, November 2003.

[Rei05] G.L. Rein. Reputation information systems: A reference model. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS)*, January 2005.

[RH05] Maxim Raya and Jean-Pierre Hubaux. The security of vehicular ad hoc networks. In *Proceedings of the 3rd ACM Workshop on Security of Ad Hoc and Sensor Networks*, pages 11–21. ACM Press, 2005.

[Ril03] George F. Riley. The Georgia Tech network simulator. In *Proceedings of the ACM SIGCOMM Workshop on Models, Methods and Tools for Reproducible Network Research (MoMeTools)*, pages 5–12. ACM Press, August 2003.

[RKEC04] Konrad Rossrucker, Antonio Kung, Stephan Eichler, and Danny De Cock. GST security use case and system requirements. Project Deliverable (Security) 2.1, EU IP – Global System for Telematics, August 2004. Available at: `http://www.gstforum.org/en/downloads/deliverables/`.

[RKZF00] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Communications of the ACM*, 43(12):45–48, 2000.

[RMM+02] D. Reichardt, M. Miglietta, L. Moretti, P. Morsink, and W. Schulz. CarTALK 2000: Safe and comfortable driving based upon inter-vehicle-communication. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 545–550, June 2002.

[RMVS05] Y. Rebahi, V.E. Mujica-V, and D. Sisalem. A reputation-based trust mechanism for ad hoc networks. In *Proceedings of the 10th IEEE Symposium on Computers and Communications (ISCC)*, pages 37–42, June 2005.

[Rom86] Raphael Rom. Collision detection in radio channels. In *Local area and multiple access networks*, pages 235–249. Computer Science Press, Inc., New York, NY, USA, 1986.

[RPH06] Maxim Raya, Panos Papadimitratos, and Jean-Pierre Hubaux. Securing vehicular communications. *IEEE Wireless Communications*, 13(5):8–15, October 2006.

[RR98] Michael Reiter and Aviel Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, February 1978.

[Sch96] Bruce Schneier. *Applied Cryptography*. John Wiley & Sons, 2nd edition, 1996.

[Sch05b] Christian Schwingenschlögl. *A Framework for Secure and Efficient Communication in Mobile Ad Hoc Networks*. Phd thesis, Technische Universität München, June 2005.

[SD02] A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In *Proceedings of the Workshop on Privacy Enhancing Technologies (PET)*, pages 41–53. Springer Berlin / Heidelberg, 2002.

[SFK07] C. Stiller, G. Färber, and S. Kammel. Cooperative cognitive automobiles. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, pages 215–220. IEEE Press, June 2007.

[SH02] Christian Schwingenschlögl and Marc-Philipp Horn. Building blocks for secure communication in ad-hoc networks. In *Proceedings European Wireless 2002. Next Generation Wireless Networks: Technologies, Protocols, Services and Applications*, February 2002.

[Sha48] Claude Elwood Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423; 623–656, 1948.

[Sha79] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, November 1979.

[SHL+05] Krishna Sampigethaya, Leping Huang, Mingyan Li, Radha Poovendran, Kanta Matsuura, and Kaoru Sezaki. CARAVAN: Providing location privacy for VANET. In *Proceedings of the Workshop on Embedded Security in Cars (ESCAR)*, 2005.

[Sho00] Victor Shoup. Practical threshold signatures. In Bart Preneel, editor, *Advances in Cryptology - Proceedings of Eurocrypt*, volume 1807 of *Lecture Notes in Computer Science*, pages 207–221, May 2000.

[SJ04] Amit Kumar Saha and David B. Johnson. Modeling mobility for vehicular ad hoc networks. In *Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, October 2004.

[SKL+06] Elmar Schoch, Frank Kargl, Tim Leinmüller, Stefan Schlott, and Panos Papadimitratos. Impact of pseudonym changes on geographic routing in VANETs. In *Proceedings of the 3rd European Workshop on Security and Privacy in Ad hoc and Sensor Networks (ESAS)*, September 2006.

[SLHP07] K. Sampigethaya, Mingyan Li, Leping Huang, and R. Poovendran. AMOEBA: Robust location privacy scheme for VANET. *IEEE Journal on Selected Areas in Communications*, 25(8):1569–1589, October 2007.

[SMSR02] C. Santivanez, B. McDonald, I. Stavrakakis, and R. Ramanathan. The scalability of ad hoc routing protocols. In *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom)*, June 2002.

[SOC04] H. Ozgur Sanli, Suat Ozdemir, and Hasan Cam. SRDA: Secure reference-based data aggregation protocol for wireless sensor networks. In *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC)*, September, 2004.

[SS01] Jordi Sabater and Carles Sierra. REGRET: Reputation in gregarious societies. In *Proceedings of the 5th International Conference on Autonomous Agents*, pages 194–195, New York, NY, USA, 2001. ACM.

[Sta05] Standards Committee. *Wireless LAN Medium Access Control (MAC) and Physical layer (PHY) specifications: Amendment 8: Medium Access Control (MAC) Quality of Service Enhancements*. IEEE Computer Society, November 2005.

[Str07] Markus Straßberger. *Kontextbereitstellung in Automobilen Ad-hoc Netzen*. Phd thesis, Technische Universität München, June 2007.

[STW02] R.P. Schäfer, K.U. Thiessenhusen, and P. Wagner. A traffic information system by means of real-time floating-car data. In *Proceedings of the ITS World Congress*, October 2002.

[SW89] J. A. Simpson and E. S. C. Weiner. *The Oxford English Dictionary*. Clarendon Press, 2nd edition, March 1989. `http://dictionary.oed.com/`.

[SWP03] Bo Sun, Kui Wu, and Udo W. Pooch. Alert aggregation in mobile ad hoc networks. In *Proceedings of the ACM Workshop on Wireless Security*, pages 69–78. ACM Press, 2003.

[Tas06] Task Group p. *IEEE P802.11p: Wireless Access in Vehicular Environments (WAVE)*. IEEE Computer Society, draft standard edition, 2006.

[TJM+04] Vikas Taliwal, Daniel Jiang, Heiko Mangold, Chi Chen, and Raja Sengupta. Empirical determination of channel characteristics for dsrc vehicle-to-vehicle communication. In *Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*. ACM Press, 2004.

[TK75] Fouad A. Tobagi and Leonard Kleinrock. Packet switching in radio channels: Part II – the hidden terminal problem in carrier sense multiple-access and the busy-tone solution. *IEEE Transactions On Communications*, 23(12):1417–1433, December 1975.

[TM07b] Marc Torrent-Moreno. *Inter-Vehicle Communications: Achieving Safety in a Distributed Wireless Environment*. Phd thesis, Universität Karlsruhe, July 2007.

[TMFH06] M. Torrent-Moreno, A. Festag, and H. Hartenstein. System design for information dissemination in VANETs. In *Proceedings of the Workshop on Intelligent Transportation (WIT)*, March 2006.

[TMJH04] Marc Torrent-Moreno, Daniel Jiang, and Hannes Hartenstein. Broadcast reception rates and effects of priority access in 802.11-based vehicular ad-hoc networks. In *Proceedings of the First ACM Workshop on Vehicular Ad Hoc Networks (VANET)*. ACM, ACM Press, October 2004.

[TMSEFH06a] M. Torrent-Moreno, F. Schmidt-Eisenlohr, H. Füßler, and H. Hartenstein. Effects of a realistic channel model on packet forwarding in vehicular ad hoc networks. In *Proceedings of the IEEE Wireless Communication and Networking Conference (WCNC)*, pages 385–391, April 2006.

[TMSH05] Marc Torrent-Moreno, Paolo Santi, and Hannes Hartenstein. Fair sharing of bandwidth in VANETs. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 49–58. ACM Press, 2005.

[TP06] Zhifeng Tao and S. Panwar. Throughput and delay analysis for the IEEE 802.11e enhanced distributed channel access. *IEEE Transactions on Communications*, 54(4):596–603, April 2006.

[TTIF89] K. Takada, Y. Tanaka, A. Igarashi, and D. Fujita. Road/automobile communication system (RACS) and its economic effect. In *Proceedings of the Vehicle Navigation and Information Systems Conference*, September 1989.

[Var01] András Varga. The OMNeT++ discrete event simulation system. In *Proceedings of the European Simulation Multiconference (ESM)*, June 2001.

[VNB+07] Stefan Vacek, Robert Nagel, Thomas Batz, Frank Moosmann, and Rüdiger Dillmann. An integrated simulation framework for cognitive automobiles. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)*, pages 221–226, June 2007.

[VVM+04] Dimitar Valtchev, Erwin Vermassen, Hans-Ulrich Michel, Hans-Joerg Voegel, Maarten Verhoeven, Oene Kerstjens, Paul van Koningsbruggen, and Sofia Doncheva. Operational concept description (OCD). Project Deliverable 2.1, European IP — Global System for Telematics, June 2004. Available at: `http://www.gstforum.org/en/downloads/deliverables/`.

[WER+03] Lars Wischhof, Andre Ebner, Hermann Rohling, Matthias Lott, and Rüdiger Halfmann. SOTIS – A self-organizing traffic information system. In *Proceedings of the 57th IEEE Vehicular Technology Conference (VTC)*, April 2003.

[WFGH04] Hao Wu, Richard M. Fujimoto, Randall Guensler, and Michael Hunter. MDDV: A mobility-centric data dissemination algorithm for vehicular networks. In *Proceedings of the 1st International Workshop on Vehicular Ad Hoc Networks (VANET)*, pages 47–56, October 2004.

[WFR04] Hao Wu, Richard Fujimoto, and Georg Riley. Analytical models for information propagation in vehicle-to-vehicle networks. In *Proceedings of the 61st IEEE Vehicular Technology Conference (VTC)*, September 2004.

[Wie68] Rainer Wiedemann. *Verkehrsablauf hinter Lichtsignalanlagen*. Phd thesis, Universität Karlsruhe, February 1968.

[Wie74] Rainer Wiedemann. Simulation des Straßenverkehrsflusses. Technical Report 8, Universität Karlsruhe, 1974.

[Woh00] Petra Wohlmacher. Digital certificates: A survey of revocation methods. In *Proceedings of the Multimedia Workshop*, pages 111–114, Marina Del Rey, 2000. ACM Press.

[XB02] Jin Xi and Christian Bettstetter. Wireless multihop internet access: Gateway discovery, routing, and addressing. In *In Proceedings of the International Conference on Third Generation Wireless and Beyond*, May 2002.

[XKG02] Hong Xiaoyan, Xu Kaixin, and M. Gerla. Scalable routing protocols for mobile ad hoc networks. *IEEE Network*, 16(4):11–21, July 2002.

[XMKS04] Qing Xu, Tony Mak, Jeff Ko, and Raja Sengupta. Vehicle-to-vehicle safety messaging in DSRC. In *Proceedings of the 1st ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, October 2004.

[YM03] Po-Wah Yau and C.J. Mitchell. Reputation methods for routing security for mobile ad hoc networks. In *Proceedings of the 1st Workshop on Mobile Future and Symposium on Trends in Communications*, pages 130–137, October 2003.

[ZBG98] Xiang Zeng, Rajive Bagrodia, and Mario Gerla. GloMoSim: A library for parallel simulation of large-scale wireless networks. In *Proceedings of the 12th Workshop on Parallel and Distributed Simulation (PADS)*, pages 154–161, May 1998.

[ZC03] Hua Zhu and Imrich Chlamtac. An analytical model for IEEE 802.11e EDCF differential services. In *Proceedings of the 12th International Conference on Computer Communications and Networks (ICCCN)*, pages 163–168, October 2003.

[ZH99] Lidong Zhou and Zygmunt J. Haas. Securing ad hoc networks. *IEEE Network*, 13(6):24–30, November 1999.

[Zha02] Yilin Zhao. Telematics: Safe and fun driving. *IEEE Intelligent Systems*, 17:10–14, January 2002.

[Zhe03] Peifang Zheng. Tradeoffs in certificate revocation schemes. *ACM SIGCOMM Computer Communications Review*, 33(2):103–112, April 2003.

[ZMTV02] Magda El Zarki, Sharad Mehrotra, Gene Tsudik, and Nalini Venkatasubramanian. Security issues in a future vehicular network. In *Proceedings of European Wireless, Next Generation Wireless Networks*, volume 1, pages 270–274, February 2002.

[ZR06] Xin Zhang and George F. Riley. Scalability of an ad hoc on-demand routing protocol in very large-scale mobile wireless networks. *Simulation*, 82(2):131–142, February 2006.

[ZSO+05] Yunpeng Zang, Lothar Stibor, Georgios Orfanos, Shumin Guo, and Hans-Juergen Reumerman. An error model for inter-vehicle communications in highway scenarios at 5.9GHz. In *Proceedings of the 2nd ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks*, pages 49–56, New York, NY, USA, 2005. ACM Press.

[Zug03] Alf Zugenmaier. The Freiburg privacy diamond. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, volume 3, pages 1501–1505, December 2003.

[ZWH04] Daqing Zhang, Xiaohang Wang, and Kai Hackbarth. OSGi based service infrastructure for context aware automotive telematics. In *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, May 2004.

Internet References

The date information given in the references stands for the last access and not for the latest revision of the respective page.

[BH08] Rimon Barr and Zygmunt J. Haas. Java in simulation time / scalable wireless ad hoc network simulator (jist/swans). Website, February 2008. http://jist.ece.cornell.edu/.

[Cen08] Centre for Applied Informatics (ZAIK). SUMO – Simulation of urban mobility. Website, May 2008. http://sumo.sourceforge.net/.

[CER08] CERT. Vulnerability remediation statistics. Website, May 2008. http://www.cert.org/stats/fullstats.html.

[Ent08] Entróuvert. LASSO – liberty alliance single sign-on. Website, April 2008. http://lasso.entrouvert.org/.

[Eur08] Europe's Information Society. eSafety Website. Website, February 2008. http://ec.europa.eu/information_society/activities/esafety/.

[Fle08] FleetNet. FleetNet project website. Website, May 2008. http://www.et2.tu-harburg.de/fleetnet/.

[Gir07] Damien Giry. Keylength.com – cryptographic key length recommendation. Website, 2007. http://www.keylength.com/.

[Glo04] Global System for Telematics. GST forum. Website, 2004. http://www.gstforum.org/.

[HFPS99] R. Housley, W. Ford, W. Polk, and D. Solo. Internet X.509 Public Key Infrastructure Certificate and CRL Profile. RFC 2459 (Proposed Standard), January 1999. Obsoleted by RFC 3280, http://www.ietf.org/rfc/rfc2459.txt.

[HPFS02] R. Housley, W. Polk, W. Ford, and D. Solo. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. RFC 3280 (Proposed Standard), April 2002. Updated by RFCs 4325, 4630.

[Inf08] Information Sciences Institute. The network simulator - ns-2. Website, February 2008. http://www.isi.edu/nsnam/ns/.

[KB08] Kraftfahrt-Bundesamt. Statistiken. Website, February 2008. http://www.kba.de/.

[Lib08] Liberty Alliance Project. Liberty alliance project homepage. Website, April 2008. http://www.projectliberty.org/.

[OAS08] OASIS. Oasis – Advancing open standards for the information society. Website, April 2008. http://www.oasis-open.org/.

[Ope08] OpenSSL Project. OpenSSL project homepage. Website, May 2008. http://www.openssl.org/.

[OPN08] OPNET Technologies, Inc. OPNET Technologies – Making networks and applications perform. Website, February 2008. http://www.opnet.com/.

[PTV08] PTV AG. Vissim simulator. Website, May 2008. http://www.ptv.de/cgi-bin/traffic/traf_vissim.pl.

[Ril08] George F. Riley. Georgia Tech network simulator (GTNetS). Website, February 2008. http://www.ece.gatech.edu/research/labs/MANIACS/GTNetS/.

[Tra08] Transport Protocol Experts Group. TPEG Forum. Website, April 2008. http://www.tpeg.org/.

[Tru08] Trusted Computing Group. Trusted platform module (TPM) specifications. Website, June 2008. https://www.trustedcomputinggroup.org/specs/TPM/.

[VH07] Andras Varga and Rudolf Hornig. OMNeT++ community site. Website, 2007. http://www.omnetpp.org/.

[ZBG08] Xiang Zeng, Rajive Bagrodia, and Mario Gerla. Global mobile information systems simulation library (GloMoSim). Website, February 2008. http://pcl.cs.ucla.edu/projects/glomosim/.

Index

W