

Thomas Strang Andreas Festag
Alexey Vinel Rashid Mehmood
Cristina Rico Garcia Matthias Röckl (Eds.)

LNCS 6596

Communication Technologies for Vehicles

Third International Workshop, Nets4Cars/Nets4Trains 2011
Oberpfaffenhofen, Germany, March 2011
Proceedings



Springer

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Thomas Strang Andreas Festag
Alexey Vinel Rashid Mehmood
Cristina Rico Garcia Matthias Röckl (Eds.)

Communication Technologies for Vehicles

Third International Workshop, Nets4Cars/Nets4Trains 2011
Oberpfaffenhofen, Germany, March 23-24, 2011
Proceedings

Volume Editors

Thomas Strang
Cristina Rico Garcia
German Aerospace Center (DLR)
82234 Wessling/Oberpfaffenhofen, Germany
E-mail: {thomas.strang, cristina.ricogarcia}@dlr.de

Andreas Festag
NEC Laboratories Europe
69115 Heidelberg, Germany
E-mail: andreas.festag@neclab.eu

Alexey Vinel
Russian Academy of Sciences (SPIIRAS)
199178 St. Petersburg, Russia
E-mail: vinel@ieee.org

Rashid Mehmood
Swansea University, College of Engineering
Swansea SA2 8PP, UK
E-mail: r.mehmood@swansea.ac.uk

Matthias Röckl
In2Soft GmbH/KPIT Cummins
80797 München, Germany
E-mail: matthias.roeckl@in2soft.de

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-19785-7 e-ISBN 978-3-642-19786-4
DOI 10.1007/978-3-642-19786-4
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011922381

CR Subject Classification (1998): C.2, C.3, K.6.5, K.8.1

LNCS Sublibrary: SL 5 – Computer Communication Networks and Telecommunications

© Springer-Verlag Berlin Heidelberg 2011

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Communication Technologies for Vehicles workshop series provides an international forum on latest technologies and research in the field of intra- and inter-vehicle communications in which to present original research results in all areas relating to communication protocols and standards, mobility and traffic models, experimental and field operational testing, and performance analysis.

Previous Nets4Cars workshops were held in Saint Petersburg, Russia (2009) and in Newcastle, UK (2010). These proceedings contain the papers presented at the Third International Workshop on Communication Technologies for Vehicles (Nets4Cars and Nets4Trains 2011), which for the first time had dedicated tracks for road- and rail-based approaches and took place in Oberpfaffenhofen near Munich, Germany, in March 2011.

Our call for papers resulted in 34 submissions, 13 for the rail track and 21 for the road track. Each of them was assigned to at least four members of our outstanding Technical Program Committee with specific expertise in the field. After a double-blind review process in just 2 weeks and some online discussion on boundary cases, the Program Committee Co-chairs selected 19 full papers for publication in these proceedings and presentation at the workshop, 7 of them for the rail track and 12 for the road track. In addition, one invited paper was accepted from a strong industrial stakeholder in the rail track who also gave the keynote. The order of the papers in these proceedings was aligned with the workshop program.

We extend a sincere “thank you” to all the authors who submitted papers of their most recent work, to all the members of our hard-working comprehensive Technical Program Committee, as well as the thoughtful external reviewers.

March 2011

Thomas Strang and Andreas Festag,
TPC Co-Chairs

Alexey Vinel and Rashid Mehmood,
General Co-Chairs

Cristina Rico Garcia,
Rail Track Chair

Matthias Röckl,
Road Track Chair

Organization

Conference Organizers

General Co-chairs	Alexey Vinel (SPIIRAS, Russia)
Technical Program Co-chairs	Rashid Mehmood (Swansea University, UK) Thomas Strang (DLR, Germany)
Rail Track Chair	Andreas Festag (NEC, Germany)
Road Track Chair	Cristina Rico Garcia (DLR, Germany) Matthias Röckl (In2Soft, Germany)

Steering Committee

Axel Sikora	Duale Hochschule Baden-Wurttemberg, Germany
Tsutomu Tsuboi	Renesas Corp., Japan
Fei Liu	University of Twente, The Netherlands
Xu Li	State University of New York at Buffalo, USA
Yan Zhang	Simula Research Laboratory, Norway
Antonella Molinaro	University Mediterranea of Reggio Calabria, Italy
Marion Berbineau	INRETS, France
Juan de Dios Sanz Bobi	CITEF, Spain

Technical Program Committee

Marina Aguado	University of the Basque Country (Spain)
Onur Altintas	Toyota InfoTechnology Center (Japan)
Atif Alvi	LUMS (Pakistan)
Petros Belimpasakis	Nokia Research Center (Finland)
Marion Berbineau	INRETS (France)
Mohamed Boucadair	France Telecom (France)
Torsten Braun	University of Bern (Switzerland)
Marcello Caleffi	University of Naples Federico II (Italy)
Eduardo Cerqueira	University of Coimbra (Portugal)
Soumaya Cherkaoui	University of Sherbrooke (Canada)
Marilia Curado	University of Coimbra (Portugal)
Robil Daher	University of Rostock (Germany)
Thierry Ernst	INRIA (France)

VIII Organization

Andreas Festag	NEC Laboratories Europe (Germany)
Fethi Filali	Qatar University Wireless Innovations Center (Qatar)
Francisco Garcia	Agilent Technologies (UK)
Benoit Geller	ENSTA (France)
Javier Goikoetxea	Construcciones y Auxiliar de Ferrocarriles (Spain)
Javier Gozalvez	Universidad Miguel Hernandez de Elche (Spain)
Christophe Gransart	INRETS (France)
Oleg Gusikhin	Ford (USA)
Jerome Harri	EURECOM (France)
Geert Heijenk	University of Twente (The Netherlands)
Muhammad Ali Imran	University of Surrey (UK)
Sithamparanathan Kandeepan	CREATE-NET (Italy)
Yevgeni Koucheryavy	Tampere University of Technology (Finland)
Uwe Kucharzyk	Bombardier Transportation (Germany)
Long Le	NEC Laboratories Europe (Germany)
Andreas Lehner	German Aerospace Center (DLR)(Germany)
Tim Leinmüller	DENSO AUTOMOTIVE Deutschland GmbH (Germany)
Fei Liu	University of Twente (The Netherlands)
Katrin Lüdecke	German Aerospace Center (DLR)(Germany)
Juliette Marais	INRETS-LEOST (France)
Rashid Mehmood	Swansea University (UK)
Markus Miche	SAP Research (Germany)
David Mottier	Mitsubishi Electric R&D Centre Europe (France)
John Murphy	University College Dublin (Ireland)
Augusto Neto	Universidade Federal de Goias (Brazil)
Brian Park	University of Virginia (USA)
Cristina Rico-Garcia	German Aerospace Center (DLR)(Germany)
Matthias Röckl	In2Soft / KPIT Cummins (Germany)
Paolo Santi	IIT-CNR (Italy)
Divitha Seetharamdoo	INRETS (France)
Thomas Strang	German Aerospace Center (DLR)(Germany)
Markus Strassberger	BMW Group Research and Technology (Germany)
Jouni Tervonen	University of Oulu (Finland)
Ozan Tonguz	Carnegie Mellon University (USA)
Tsutomu Tsuboi	Renesas Technology Corp (Japan)
Bart van Arem	TU Delft (The Netherlands)
Alexey Vinel	SPIIRAS (Russia)
Martine Wahl	INRETS (France)
Michelle Wetterwald	EURECOM (France)

Christian Wewetzer
Nawaporn Wisitpongphan
Yunpeng Zang
Yang Zhang

Volkswagen Group (Germany)
KM Univ. of Techn. North Bangkok (Thailand)
RWTH Aachen (Germany)
Pennsylvania State University (USA)

Additional Reviewers

Robert Schmidt

DENSO AUTOMOTIVE Deutschland GmbH
(Germany)

Osianoh Aliu

University of Surrey (UK)

Amin Amich

University of Surrey (UK)

Herv Bonneville

Mitsubishi Electric R&D Centre Europe

(France)

Miguel Sepulcre

University Miguel Hernandez of Elche (Spain)

Hosting Institution

Nets4Cars & Nets4Trains 2011 was hosted by the Institute of Communications and Navigation at the German Aerospace Center (DLR)

Sponsoring Institutions

German Aerospace Center (DLR), Germany

SPIIRAS, Russia

Swansea University, UK

Tampere University of Technology, Finland

Table of Contents

Keynote

Requirements for Wireless Technology on Rolling Stock	1
<i>Uwe Kucharzyk</i>	

Rail Track

An Experimental Study of Multi-radio Platform Coexistence in the 5 GHz Band for Railway Applications	11
<i>Jorge Higuera, Elli Kartsakli, Carlos Collado, José M. González-Arbesú, Luis Alonso, José Luis Valenzuela, Andrés Laya, Enrique Flores, Isabel Navarro, Raquel Martínez, Jesús González, José Hierro, and Adrián Vlad</i>	
Train Tracking and Shadowing Estimation Based on Received Signal Strength	23
<i>Hadi Noureddine, Damien Castelain, and Ramesh Pyndiah</i>	
Delivering Broadband Internet Access for High Speed Trains Passengers Using an Innovative Network Mobility Solution	34
<i>Bernadette Villeforceix</i>	
Measurement and Analysis of the Direct Train to Train Propagation Channel in the 70 cm UHF-Band	45
<i>Andreas Lehner, Cristina Rico García, Thomas Strang, and Oliver Heirich</i>	
WiMax'ble Pervasive Cloud – Empowering Next Generation Intelligent Railway Infrastructure	58
<i>Subrahmanyam Venkata Radha Krishna Rao and Vivek Diwanji</i>	
The MIH (Media Independent Handover) Contribution to Mobility Management in a Heterogeneous Railway Communication Context: A IEEE802.11/802.16 Case Study	69
<i>Marina Aguado, Jasone Astorga, Jon Matias, and Maider Huarte</i>	
Multiple Description Coding and Scalable Video Coding Combined with Multiple Input Multiple Output Techniques: Two Strategies to Enhance Train to Wayside Video Transmissions in Tunnels	83
<i>Imade Fahd Eddine Fatani, Yann Cocheril, Crépin Nsiala, Marion Berbineau, François-Xavier Coudoux, Marie Zwingenstein-Colin, and Patrick Corlay</i>	

Road Track

VANET Architectures and Protocol Stacks: A Survey	95
<i>Sajjad Akbar Mohammad, Asim Rasheed, and Amir Qayyum</i>	
Behavior Specification of a Red-Light Violation Warning Application – An Approach for Specifying Reactive Vehicle-2-X Communication Applications	106
<i>Sebastian Röglinger and Christian Facchi</i>	
Wireless Protocol Design for a Cooperative Pedestrian Protection System	119
<i>Dirk Lill, Manuel Schappacher, Shahidul Islam, and Axel Sikora</i>	
A Vehicular Mobility Model Based on Real Traffic Counting Data	131
<i>Yoann Pigné, Grégoire Danoy, and Pascal Bouvry</i>	
Driver-Centric VANET Simulation	143
<i>Pedro Gomes, Cristina Olaverri-Monreal, Michel Ferreira, and Luís Damas</i>	
Simulative Evaluation of the Potential of Car2X-Communication in Terms of Efficiency	155
<i>Benno Schweiger, Philipp Ehnert, and Johann Schlichter</i>	
Performance Study of an In-Car Switched Ethernet Network without Prioritization	165
<i>Hyung-Taek Lim, Kay Weckemann, and Daniel Herrscher</i>	
Degradation of Communication Range in VANETs Caused by Interference 2.0 - Real-World Experiment	176
<i>Robert K. Schmidt, Bernhard Kloiber, Florian Schüttler, and Thomas Strang</i>	
Real-World Measurements of Non-Line-Of-Sight Reception Quality for 5.9GHz IEEE 802.11p at Intersections	189
<i>Thomas Mangel, Matthias Michl, Oliver Klemp, and Hannes Hartenstein</i>	
Interoperability Testing Suite for C2X Communication Components	203
<i>Fabian de Ponte Müller, Juan María Reveriego Sierra, Bernhard Kloiber, Matthias Röckl, and Thomas Strang</i>	
Towards Standardization of In-Car Sensors	216
<i>Zubair Nabi, Atif Alvi, and Rashid Mehmood</i>	

Secure Automotive On-Board Protocols: A Case of Over-the-Air Firmware Updates	224
<i>Muhammad Sabir Idrees, Hendrik Schwepppe, Yves Roudier, Marko Wolf, Dirk Scheuermann, and Olaf Henniger</i>	
Author Index	239

Requirements for Wireless Technology on Rolling Stock

Uwe Kucharzyk

Bombardier Transportation GmbH, Schöneberger Ufer 1,
D-10761 Berlin, Germany
uwe.kucharzyk@de.transport.bombardier.com

Abstract. This paper gives an overview of railway specific requirements on wireless technologies and discusses the influence of different product life cycles in railway industry and communication technology. The main challenges are the heterogeneous communication infrastructure at the wayside and the specific requirements on vehicle speed and geographical characteristics. The current state of European research and international standardization activities for on-board communication equipment is summarized and the need for a standardized application interface on the railway vehicle and at the wayside is explained. Finally, challenges for future research and development in railway applications are identified.

Keywords: Wireless communication, railway requirements, train operation, maintenance, passenger information, railway application interface, product life cycle.

1 Introduction

Wireless technology is nowadays fully integrated into our everyday life, our social communication as well as effectively supporting large distributed automation systems. As travellers and railway vehicles are usually moving, you would expect that wireless communication is used everywhere and to a great extent on rolling stock to improve the quality of the overall railway system. But, after more than 20 years of experience with wireless technologies in the railway sector we have to admit that the introduction in a broad scale seems to be much slower than in other areas.

First applications of wireless communication on rolling stock go back to the 70ies and 80ies of the last century. Examples are i.e. analogue cab radio for the audio communication between driver and control center and data communication for remote diagnostic of the ICE1-trains.

With upcoming digital wireless technology – mobile phones and wireless LANs – the application of wireless technologies on railways has developed in signalling systems, like GSM-R for ETCS, diagnostic and telemetric applications for remote upload of vehicle diagnostic data or monitoring data (i.e. energy consumption) and also passenger information with dynamic seat reservation. Later, also private mobile phone communication on board a train (repeater) and Internet access for travelers has was implemented in a lot of project, mostly on special routes or for some parts of the fleet.

Designing a new railway vehicle and its communication infrastructure you have to consider the interests of different stakeholders in the railway system (fig. 1). Although these roles are somehow generic, they are assigned differently to business organizations depending on the market segment, as i.e. public transport, long distance

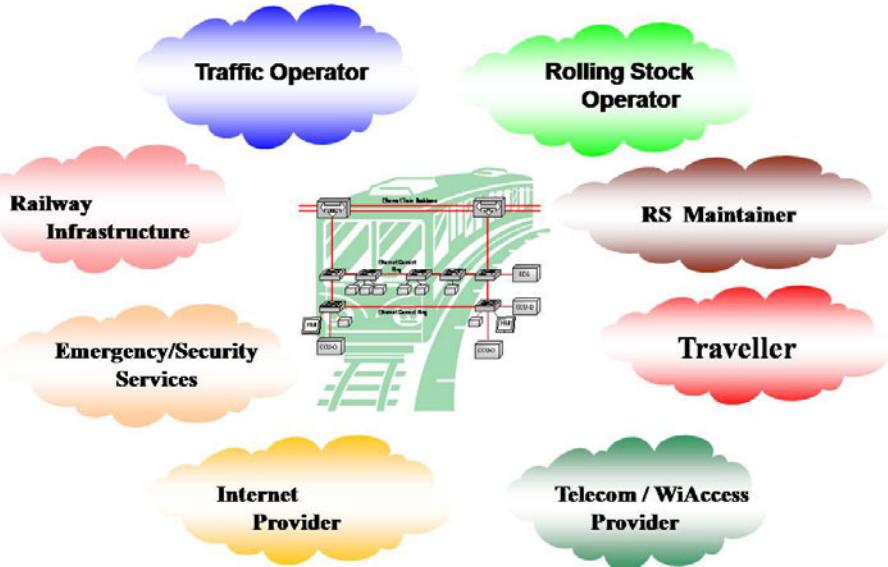


Fig. 1. Stakeholders in a Railway System

travel and freight and also on the country and the legal situation regarding liberalization of the railway business.

2 Technical and Systems Requirements

2.1 Railway Infrastructure

The railway system is - as other transportation systems - very complex and consists of

- Tracks
- Vehicle Fleets from different vendors and age
- Stations
- Depots
- Signaling system
- Passenger information system
- Train operation control

According to the main transportation goal we have also a variety of systems which set different main focus for the communication needs and environmental conditions as:

- Closed transportation systems on airports and in metropolitan areas
- Public transport systems in cities as light rail vehicles and metros, usually on dedicated railway infrastructure
- Regional traffic on national railway infrastructure but operated by local operator or private operator
- Long-distance traffic including high speed traffic on national and international infrastructure

The vehicle speed in those systems varies also and can be divided into four scenarios [1]:

- 0 – 7km/h: Stationary
- 7 – 80km/h Low Speed
- 80 – 160 km/h Medium Speed
- 160 – 350 km/h High Speed

While it is possible to have a homogeneous communication system on closed transportation systems and also on public transport systems, the solutions for regional and long-distance traffic have always to consider a very heterogeneous infrastructure, a wide variety of landscape conditions, like city areas, mountains, tunnels, rural areas with less equipment of standard mobile technologies and also different generations of communication equipment along the track.

To have clear planning guidelines for railway communication systems, the impact of vehicle speed on the data transmission and the impact of landscape conditions on the radio signal propagation are important questions for the industry.

Figure 2 illustrates a generic scenario of a train moving on its journey through differently covered wireless networks [2]. The networks are differentiated not only by the wireless technology itself, as i.e. GSM-R, GPRS, UMTS, Wi-Fi, WiMax and others, the access points belong also to different networks and network providers.

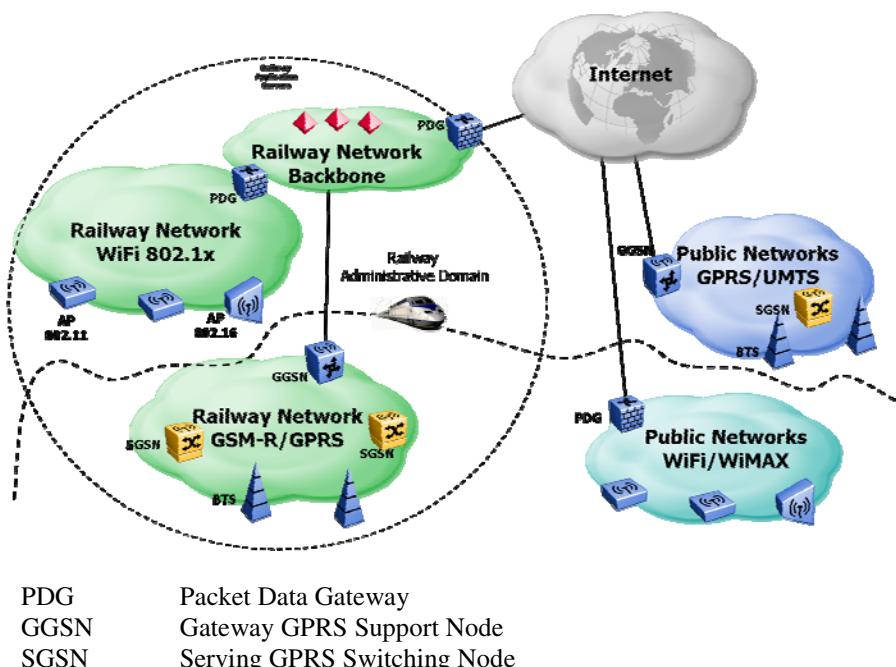


Fig. 2. Data Communication Network for Railway Application

A general approach for a continuous “always-on” communication between railway vehicle and wayside has to consider handover between cells of same technology and same provider or different provider and even, as the general case, cells of different technology, if a train leaves for instance a metropolitan area or station with WiMax connectivity and runs through rural areas with only mobile technology or satellite connection.

During the handover there is impact on the data transmission from short reduction in throughput up to interruption of the connection, which is critical for applications like VPN.

Behavior of communication during the handover and improved handover procedures for the different speeds and technologies are therefore of great interest for the railway vehicle industry.

2.2 Wireless Railway Applications

From application point of view the generic wireless communication requirements can be classified into 4 main groups:

- A) Operational safety relevant data with extreme high availability requirements, as for signaling system.
- B) Operational relevant real-time data, i.e. train position, actual passenger load, actual schedule, alarm messages, security information, ...
- C) Operational relevant best effort data, i.e. update of passenger information system, diagnostic information for the depot, ...
- D) Communication for personal use of the traveler.

Class A communication is today implemented according to requirements for closed safety relevant communication systems as defined in EN50159-1 [3]:

- Communication system is closed.
- Number of connectable end devices has to be known and fixed.
- Physical parameters and qualities of the transmission system need to be known (including worst case environmental conditions).

The new standard proposal [4] covers both - closed and open communication systems – and defines 3 categories of data transmission system, while category 1 system is similar to the closed system of the actual standard with more strict requirements regarding unauthorized access.

In principle it is allowed also to transmit data from applications without specific safety requirements but due to the high (re-)homologation effort if changing applications are added to this communication channel the use will practically be restricted in the mid-term future to the signaling system and probably to some predefined class B data.

Class B and C communication is used for different stakeholders to optimize their business processes from passenger information via train operation to maintenance.

For class B the communication channel needs to be implemented along the whole train route to allow real-time access. The necessary bandwidth varies very much by the supported application from some kbit/s to Mbit/s. Class C communication can be implemented in hot-spots in stations and depots. Usually higher amount of data will be transferred which requires higher transmission rate to reach acceptable upload / download reaction times.

Class D communication includes all personal traveler requirements, starting from mobile phone use, via e-mail, Internet- and VPN-access. The requirements in this area are developing with changing social behavior and depend strongly on mean travel time. In general there will be a tendency that the traveler of tomorrow expects aboard a train the same or similar connectivity as at home or in a nomad environment (i.e. cafe, railway station, office, etc.).

2.3 Architecture Train Control System

Looking at the architecture of a train control system we distinguish separate locations for wireless communication:

T2W train to wayside communication could be classified as T2Infrastructure and T2Internet. T2Infrastructure includes signaling system (class A data) and also special hot-spots in stations and depots (class C data), or special/public infrastructure at the wayside (class B data). T2Internet establishes a communication between the traveler and the Internet (class D communication). This can also be used for class B and C applications if the availability and real-time behaviour of the Internet communication is sufficient for the operational purpose. Railway specific research is necessary for the signal propagation of the different technologies in tunnels or, metropolitan areas with high buildings and also mountains.

T2T train to train, this will not be considered here as it is usually not necessary, because the relations between trains will be handled by the signaling system and the wayside control centers --> both based on T2W communication.

V2V vehicle-to-vehicle, is today implemented as part of train control and management system by cable based data busses.

Wireless connectivity could save cost for automatic couplers and could even reduce cabling effort for refurbishment and modernization projects. Even reliability improvements for the V2V communication could be reached because mechanical contacts would be replaced by electronic.

Although this is an interesting approach, there are no practical solutions in place. The main requirements to such a solution could be defined in comparison to the existing cable solutions which are based on MVB, CAN, WTB or Ethernet [5]. Specific technical requirements could be found in the respective standard documents. A proper wireless solution would have to provide comparable availability via redundancy management. The influence of radio disturbances by trains on neighbor tracks and also EMC conditions need to be investigated to get a stable solution.

Inside car we have the wireless connectivity of travelers to Internet GW or mobile phone repeater which can be implemented by standard technology.

Wireless sensor or computer connections to the on-board vehicle control system, i.e. between bogie and car body, could enable new monitoring applications. While data rates would not be so extreme, the main requirements are defined by the harsh environmental conditions and by EMC questions, like mutual disturbance of different wireless technologies or susceptibility of wireless technology against propulsion influences EN 50121[6].

3 Business Aspects

3.1 Product Life Cycle

Railway vehicle have a very long life cycle of more than 30 years, which was not a problem with mechanical and simple electric control systems on the vehicles. With the introduction of microprocessor control equipment on railway vehicles the discrepancy of the vehicle life cycle versus the electronic subsystems' life cycle became obvious. This trend is getting worse with the use of commercial off the shelf (COTS) products, especially for control and for communication modules.

The life cycle of industrial/railway control units is in the area of 10 years compared to communication modules for wireless communication which are lower than 5 years.

This does not automatically mean that we have to replace the electronic equipment after 10 – 15 years as the economic life-time of such equipment can be much longer. One practical example is the ICE1 which was developed in the 80ies and started revenue service 1991. After 15 years a major modernization of the interiors took place together with an upgrade of the passenger information system. All electronic equipment for the control, monitoring and diagnostic of the vehicles remained in original condition because the electronic is still reliable and expected remaining service life is sufficient. Also, the requirements for communication to subsystems on the train and for remote diagnostic support have not been changed. On the other hand, replacing the electronic would have caused extreme re-engineering and re-homologation cost.

In contrary to this the passenger information system has been upgraded to fulfill new requirements of a modern operational service and traveler expectations. Especially an electronic seat reservation system – also with new wireless communication technology - has been included.

Generally we can take the conclusion that the overall design of a railway vehicle has to be prepared for the later integration of new equipment due to new operational, signalling or end customer requirements. Main focus has to be put on long-living electric and informational interfaces. Also the mechanical interfaces, especially space for antennas outside the vehicle, should not be changed.

3.2 Business Case

The main stakeholder (refer also to fig. 1) in improved wireless connectivity between train and wayside are listed according to their increasing wireless communication system needs:

- Rolling stock manufacturer (during warranty phase)
- Rolling stock maintainer
- Rolling stock operator with security services and advanced passenger information services
- Telecom/WiAccess provider
- Internet provider
- Traveler.

Manufacturer and vehicle maintenance organization can already improve their business substantially with quite simple “island” solutions that are for instance based on a wireless hot-spot in the depot or the terminal stations and probably improved by a narrow-band data transmission with best effort access.

The rolling stock operator needs more complex wireless system solution which starts from a network of hot-spots in each station and a high available narrow-band “always-on” connection to the train up to broadband connectivity between train and wayside for high-end passenger information and security services (i.e. live video transmission from cabin to security center).

The traveler requirements for personal communication can only be fully satisfied by the possibility to get at any time an “always-on” connectivity between his personal equipment (i.e. phone, PDA, iPod, computer) and the world’s communication infrastructure as he has it at home.

In general, with a high available broadband connectivity the requirements of all stakeholders could be fulfilled but the investment to reach this step for a railway vehicle fleet in a certain region is huge. Only joint efforts of train operator, wireless and phone provider as well as Internet (content) provider could solve this.

First experiences with Internet-on-train projects have shown that refinancing of the investment by traveler user fees is not in general a successful approach. New business models have to be developed. The return on invest needs to come from other business areas, like for instance increasing passenger numbers and/or more 1st class tickets for the railway operator, cross financing via commercials or other business by Internet service provider or increased customer binding for the communication network provider.

With the liberalization of the railways in the European countries a new challenge has been faced, that the vehicles have to be adopted after the service contract period of 10 – 12 years to the infrastructural requirements of new regions and the traveler and operator needs for the next decade. This leads possibly to refurbishment and upgrade of wireless on-board technology with every new service provider.

4 Standardisation / EU-Projects

4.1 Train Communication Network

The on-board communication network for trains has been standardized over several years and is defined in IEC 61375 [5]. The train communication network consists of cable based bus systems for railway vehicles (i.e. a fixed coupled unit of 1-n railway vehicles, which will not be separated during normal service) and a dynamically reconfigurable cable network for the automatic train coupling. So far possible consist networks have been MVB and CAN and the train network was the WTB wired train bus. The working group WG43 of IEC TC9 – and a parallel working group in CENELEC TC 9X - defines a new Ethernet Consist Network ECN and a new Ethernet Train Backbone ETB. Both provide much more performance than the older ones and are based on UDP and TCP/IP telegrams.

It has been early identified that the T2W communication needs also to be standardized and a Mobile Communication Gateway MCG has been defined in Part 1 of the standard: Train Communication Network General Architecture [5].

4.2 EU-Projects

In parallel the T2W communication aspect has been part of several EU research projects, like TrainCom (2001 – 2004) and IntegRail (2005 – 2008).

IntegRail investigated the methods for horizontal handover on layer 2 of the OSI reference model between wireless cells of the same technology (as example the classical handover within a GSM network) and vertical handover between wireless cells of different technologies, i.e. between Wi-Fi and GSM (refer to fig. 2). As vertical handover includes activities in layer 2 and 3 and involves two different networks it has in general higher switching delays. To reduce the possible switchover delay methods of “soft” handover are described which are based on parallel operation of the 2 adjacent cells for the switchover time.

As a 100% accessibility of the data transmission is important to operate applications with class B and class D (i.e. VPN) requirements, we see a need of developing optimal algorithms and solutions for this switchover.

4.3 Future Standardization

An *advanced communication system within the IntegRail project* (ICOM) has been worked out as proposal for future standardization work [7]. It defines a development, distribution and operation framework for railways in four parts (see also fig. 3):

- Development Environment
- Application Environment Layer
- Information and Transfer Layer
- Deployment.

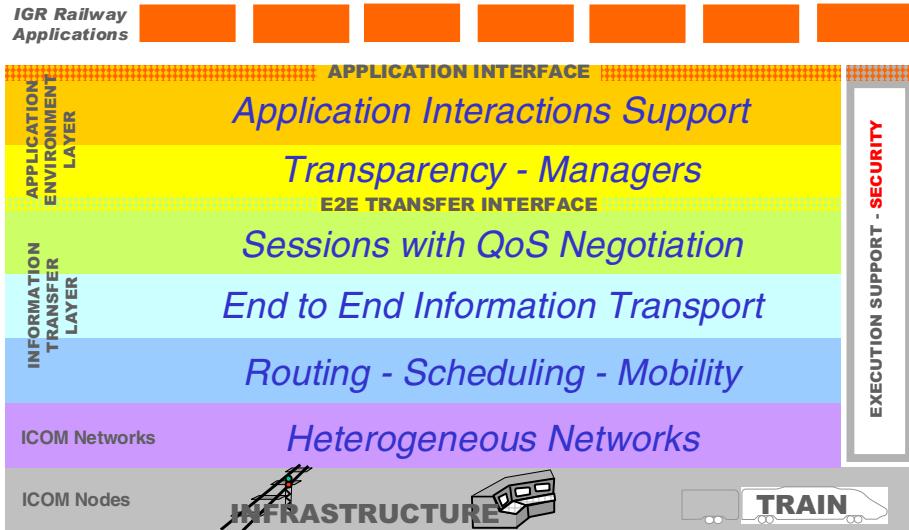


Fig. 3. ICOM Functional Layers according to [7]

The standardization proposal contains the ICOM architecture and the interfaces between the components of this architecture and two specific components - Session Manager and Network Selector. ICOM layers are based on standard OSI layers to offer a clear separation of application and network domains.

Meanwhile IEC TC WG43 has taken over this document from IntegRail and plans further work on the standard proposal for the railway mobile communication architecture within 2011. A first draft shall be available as *IEC61375-2-6 TRain to Ground Communication* in Q3 2011.

The architecture shall allow the integration of different existing and upcoming wireless networks, like for instance the next generation mobile network *Long Term Evolution LTE*. LTE is currently also under standardization and railway specific requirements, as priority, pre-emption and group call functionality have not been considered so far [8]. But even, if LTE could in long term replace GSM-R one and be the broadband standard for the railway, the necessity of vertical handover, i.e. switching between different wireless technologies, will remain, as the product cycles of railway infrastructure and communication technology differ so much.

IEC TC9 WG43 has also a liaison with the European Telecommunications Standards Institute (ETSI), a not-for-profit organization with more than 700 member organizations coming from 62 countries, and will coordinate this work with responsible ETSI Technical Committees.

5 Summary and Outlook

Heterogeneous nature of the railway system is the basic reasons for the use of various wireless technologies along the wayside of a running train. Suitable wireless

technologies have to be carefully selected according to train speed conditions, geographical characteristics and radio wave propagation.

The main problems to be solved are to provide sufficient bandwidth for the defined application profile and to stay connected with the wayside when wireless technologies or network are changing along the train route. A detailed knowledge of horizontal and vertical handover technologies is necessary and the performance of vertical handovers between different technologies needs to be improved to fulfill the real-time requirements of new applications. In addition we have to consider technologies that bundle different wireless technologies for bandwidth improvement.

The long product life cycle of railway applications and equipment require a standardized application interface on both the vehicle and the wayside [8]. Only this approach can save the investments in remote railway application over the next technology change, as for instance the migration from IPv4 networks into IPv6 networks or the introduction of new wireless technologies, like LTE. European research projects and ongoing standardization of IEC (TC9 WG43) and CENELEC (TC 9X) concentrate on this topic.

Nevertheless, in mid-term perspective the use of parallel and separated wireless networks for signaling system, operational relevant data services and traveler connectivity will remain because of the differences in network availability and bandwidth.

Further applications for wireless technologies in railway system are inter- and intra-car connection to save cabling effort or to implement new monitoring and diagnostic systems.

References

1. IntegRail, GSM-R Alternatives Study, Deliverable reference no: IGR-D-NOR-005-14, February 1, pp. 91–92 (2008)
2. IntegRail, ICOM Infrastructure-to-Train Communication Subsystem Specification, Deliverable reference no: IGR-D-NOR-022-08, February 14, p. 13 (2008)
3. EN 50159-1: Railway applications – Communication, signaling and processing systems – Part1: Safety related communication in closed transmission systems, European Standard (2001)
4. prEN 50159:2009, Railway applications – Communication, signaling and processing systems – Safety-related communication in transmission systems, European Standard Proposal (2009)
5. IEC 61375: Electronic Railway Equipment – Train Communication Network – Part 1, 2, 3
6. IEC 50121: Railway Applications – Electromagnetic Compatibility, Part 1, 2, 3, 4 (2006)
7. Gouëllo, P.: IntegRail, ICOM Standard Proposal, Document reference no: IGR-D-ALS-366-01, ver. 2, December 15, p. 12 (2008)
8. Andre, O.: LTE and its Applications in Railways, IEC TC9 WG46 Meetings, Firenze, November 25, p. 24 (2010)
9. Weiss, A.-H., Kucharzyk, U.: Generic ground side interface for remote train access. In: 9th International Conference on Intelligent Transport System Telecommunication, Lille, pp. 320–324 (2009)

An Experimental Study of Multi-radio Platform Coexistence in the 5 GHz Band for Railway Applications

Jorge Higuera¹, Elli Kartsakli¹, Carlos Collado¹, José M. González-Arbesú¹, Luis Alonso¹, José Luis Valenzuela¹, Andres Laya¹, Enrique Flores², Isabel Navarro², Raquel Martínez³, Jesús González⁴, José Hierro⁴, and Adrian Vlad⁴

¹ Technical University of Catalonia (UPC),
Department of Signal Theory and Communications (TSC)

² SENER Ingeniería y Sistemas

³ Spanish Railway Infrastructure Administrator (ADIF)

⁴ Applied Magnetism Institute (IMA),
Complutense University of Madrid

Abstract. This paper studies the practical challenges that arise due to the coexistence of two wireless technologies, both operating in the license-exempt 5 GHz band. In particular, WiFi and WiMAX equipment have been used in the experiments. The mutual interference caused by the two technologies operating in different but narrowly separated frequency channels has a negative impact on the performance of both systems. Further challenges are introduced when the two systems are in close physical proximity of each other or, in a more extreme scenario, share the same antenna as could be required in railway applications. This paper investigates these issues through a series of experimental tests based on a multi-radio platform testbed. The conclusions drawn from this study will be used as a base for the implementation of a multi-radio platform to provide communications between train and land in both directions in the context of the Spanish high-speed railway system.

Keywords: mutual interference, wireless multi-radio platform, multi-radio coexistence, high-speed railway applications.

1 Introduction

In recent years, Spain has made a significant and sustained investment in the national railway system, focusing especially on high speed rail services. According to data provided by the Spanish railway infrastructure administrator, ADIF, by 2011 the country is expected to have more than 2200 km of high-speed railway in use, thus exceeding countries with a long tradition on railway transportation, such as Japan and France [1].

This continuous growth in railway infrastructure demands new innovative services to develop efficient wireless communication systems to provide control monitoring and enhance security. The challenge is even greater in scenarios involving a wireless link between land infrastructure and high-speed trains that can travel over 300 km/h.

Another decisive factor in the selection of the most appropriate wireless technology is the type of services that should be supported. These vary from low rate telemetry applications for the collection and transmission of sensor data on infrastructure-related or environmental conditions of the railway tracks (e.g., obstacle detection or wind speed and wind direction measurements) to high rate on-board video surveillance systems operating in real-time.

Especially for the latter scenario that involves the transmission of video between high-speed trains and the land infrastructure, the available wireless solutions are limited to few proprietary technologies as in [2]. Hence, an interesting question arises on whether conventional wireless technologies for video applications can be employed, in spite of not being explicitly designed for high-speed railway networks.

Some wireless standards that could potentially be adopted for video and data transmission in dynamic applications are the IEEE 802.11a/n specification [3], widely known as WiFi, other recent specifications as IEEE 802.11p for Wireless Access in Vehicular Environments (WAVE) [4], and the IEEE 802.16d/e Worldwide Interoperability for Microwave Access (WiMAX) standard [5]. However, establishing simultaneous multi-radio links in the same scenario is not a straightforward matter, due to mutual interference between the two technologies operating in the same frequency band.

This paper will investigate the feasibility of the coexistence of WiMAX and WiFi technologies, both operating in the license-exempt band of 5 GHz. The first technology is used for video surveillance applications in railways, whereas the second is used to monitor the railway infrastructure or as a Passenger Information System (PIS). The main novelty of this work is that it examines the coexistence from a practical point of view, based on a series of experimental tests carried out on a real multi-radio platform that encompasses the two wireless technologies in a railway environment. These problems have increased in the last decade due the development of different wireless standards in the 5 GHz band for use in railway applications. Also, there are not enough extensive studies and models that address these issues of coexistence in railway applications. The presented results refer to two different scenarios. First the best-case system performance will be measured, obtained when the two technologies operate in an independent and isolated manner. Then, we will examine the worst-case scenario in which the two systems share a single multi-band antenna and work simultaneously. The conclusions drawn from this study will be used to determine the most efficient and feasible configuration for a multi-radio platform, that will eventually be employed to establish a link between high-speed trains and land infrastructure for real time video surveillance applications.

The remaining of this paper is organized into four different sections. Section 2 provides a brief overview of the WiFi and the WiMAX specification, which are the technologies being used in the measurements. Section 3 describes the experimental setup and methodology that have been considered in this work. Results on the performance evaluation are presented and discussed in Section 4. Finally, the last section is dedicated to conclusions and lessons learned for the deployment of a multi-radio platform with standard WiFi and WiMAX wireless communication technologies.

2 WiFi and WiMAX Coexistence

This section will provide a brief overview of the two wireless technologies that will be evaluated in this paper, namely WiFi and WiMAX. Even though the standards are defined in more than one frequency band, the main focus of this paper will be laid on the 5 GHz band. This band has been selected due the availability of license-exempt frequency sub-bands and equipment to consider in high-speed railway environment.

Table 1. Overview of the WiFi and WiMAX standards in the 5 GHz band

Standard	Frequency bands (GHz)	Modulation	Multiple access	Max Data rate
IEEE 802.11a	5.15 – 5.825	OFDM with BPSK, QPSK, 16-QAM, 64-QAM	CSMA/CA	54 Mbps
IEEE 802.11n	2.4 – 5.8	OFDM with BPSK, QPSK, 16-QAM, 64-QAM	CSMA/CA	300 Mbps (2 antennas)
IEEE 802.11p	5.9	OFDM with BPSK, QPSK, 16-QAM, 64-QAM	CSMA/CA	54 Mbps
WiMAX IEEE 802.16	10 - 66 (single carrier), 2.495 - 2.686, 3.3 - 3.8, 5.15-5.35, 5.75-5.825	BPSK, QPSK, 16-QAM, 64-QAM, 256-QAM	TDMA, OFDMA	75 Mbps, 134 Mbps (28 MHz channel bandwidth) TDD/FDD
Mobile WiMAX IEEE 802.16e	5.15 – 5.825	BPSK, QPSK, 16-QAM, 64-QAM	TDMA, S-OFDMA (128, 256, 512, 1024 2048 subcarriers)	128 Mbps downlink, 56 Mbps uplink. TDD / FDD

The main features of the two technologies are summarized in Table 1. The IEEE 802.11 specification includes a family of standards based on the Carrier Sensing Multiple Access with Collision Avoidance (CSMA/CA) mechanism, according to which WiFi terminals must listen to the channel and wait until it is sensed idle before attempting transmission [3]. The various amendments of the standard define different Modulation and Coding Schemes (MCS) and operate in the unlicensed bands of either 2.4 or 5 GHz. In this paper, the 802.11a amendment will be considered, which employs Orthogonal Frequency Division Multiplexing (OFDM) in the 5 GHz band. The OFDM signal is comprised by a group of closely-spaced frequency subcarriers (48 data subcarriers and 4 pilots with a 0.3125 MHz separation, in the case of 802.11a). The input data flow is divided in low rate streams and each stream is mapped onto a sub-carrier using one of four possible modulations (BPSK, QPSK, 16-QAM and 64-QAM), yielding data transmission rates that range from 6 to 54 Mbps. Additionally, IEEE 802.11n

defines MIMO support with one to four antennas that can achieve a throughput up to 600 Mbps.

WiMAX is mainly intended for point-to-multipoint scenarios in which an Access Unit (AU) serves a number of subscriber units (SU). Even though the IEEE 802.16 specification defines several physical layer implementations, the main channel access method of WiMAX is Time Division Multiplexing (TDM) / Time Division Multiple Access (TDMA). Both Time and Frequency Division Duplex schemes (TDD and FDD, respectively) are supported for the uplink and downlink channels, although the first method is more widely used to reduce the hardware implementation. As in 802.11a, OFDM is employed and the number of subcarriers varies depending on the implementation. In addition, the IEEE 802.16e specification for mobile WiMAX adopts a Scalable Orthogonal Frequency Division Multiple Access (S-OFDMA) scheme in which the size of the Fast Fourier Transform (FFT) is adjusted to the channel bandwidth that varies from 1.25 to 20 MHz.

This paper will study the coexistence between the two wireless technologies, WiFi and WiMAX in the 5 GHz band. As indicated in [6, 7], coexistence scenarios can be generally classified in two categories: proximity and collocation. Proximity refers to the case in which a different wireless technology operates simultaneously on separate devices that, while being physically separated, are located close enough to interfere with each other. Collocation is related with simultaneous operation of multiple radios on the same hardware device. In this case, interference is not only due to the RF radiation but also can be a result of conducted interference due to physical contact in a conductive medium.

In general, mutual interference between two technologies operating in closely spaced frequency channels is mainly due to three mechanisms:

- Spurious emissions of a transmitter in adjacent bands due to imperfect filtering, which cause an increase in Adjacent Channel Power (ACP).
- Receiver is blocked under the presence of a strong interfering signal. As the result, the wanted signal is received attenuated or may not be well received.
- Third-order inter-modulation products of interfering transmissions that fall within the frequency channel of the wanted signal.

Although all these factors affect the system performance, in this paper we will focus on the measurement of the ACP due to spurious transmissions. In practice, the effects of interference can be manifested as degradation in throughput or as a limitation of the coverage range.

Apart from these physical layer impairments, the operation of upper layers may also be affected. For example, an increased interference level may be misinterpreted by the 802.11 CSMA/CA mechanism as channel activity, thus forcing the WiFi device to defer transmission for a considerable amount of time. The experimental setup and the performance metrics employed to measure these effects of coexistence interference will be discussed in the following section.

In the past a few tests have been performed to study the ability of coexistence with different wireless standards for railway applications. In many cases, WiFi standards do not support the movement at more than 80 Km/h [8] which is restrictive in

high-speed train applications. The new IEEE 802.11p standard defines WiFi enhancements for vehicular environments moving at speeds of up 250Km/h in a 5.9 GHz band[9].

3 Case Study

Two experimental setups will be considered in order to study the WiFi/WiMAX coexistence and determine how the simultaneous operation of the two technologies in close frequency bands affects their performance:

- Scenario 1 considers the performance of WiFi and WiMAX systems operating independently. The obtained metrics will serve as a reference benchmark, given that there is no mutual interference between these two technologies. The experimental conditions include the use of WiFi Hirschmann BAT-300F devices were we perform interference measurements based on the total maximum and average channel power in the WiMAX 5.7 GHz band with the use of a signal analyzer (Agilent N9020A) that was connected to a WiMAX device (Alvarion BreezeACCESS VL) through a diplexer filter. The input attenuation of the signal analyzer was adjusted to zero in all cases to reduce the noise floor of the equipment and detect the interference contributions of the WiFi bands. In this scenario interference measurements were carried out with the WiFi device in operation to study the interference over WiMAX band. Similarly, measurements of interference and throughput in the same WiFi bands due to WiMAX operation in the 5.7 GHz band were carried out to get a reference benchmark of the maximum throughput channel capacity.
- Scenario 2 considers the interference of two wireless devices within a multi-radio platform. The setup is shown in Fig. 1. On one side, a multi-radio platform is employed, which includes both a WiFi and a WiMAX transceiver sharing a single multi-band antenna. Each device is connected to the antenna through a system formed by a diplexer filter and a hybrid-combiner coupler. On the other side of the link, each WiFi and WiMAX transceiver is connected to its own antenna. In scenario 2, to obtain significant interference measurements, ten independent measurements were carried out for WiFi and WiMAX. In addition, we perform throughput measurement taking in account the total number of data bits successfully transmitted per the time unit. The maximum throughput capacity has been measured for UDP and TCP with WiFi and WiMAX technologies working at the same time.

Commercial WiFi/WiMAX equipments have been selected for the experimental setup. As illustrated in Fig. 1, one WiFi transceiver shares a multi-band antenna included in the multi-radio platform, which is a Huber+Suhner Sencity Rail SWA omnidirectional antenna with 8 dBi for high-speed railway applications. In land, the second WiFi device employs an omnidirectional dipole antenna of 12 dBi.

With respect to the WiMAX technology, an Access Point and a subscriber unit have been selected. The AP in the multi-platform uses the same multi-band antenna described before, whereas the subscriber unit in land has an integrated directional antenna with a gain of 20 dBi.

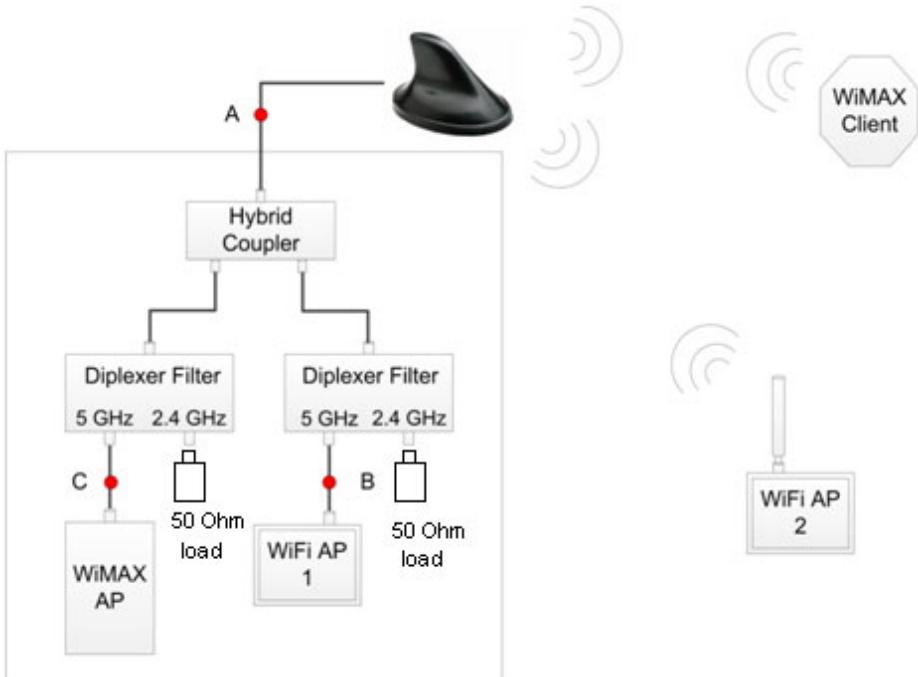


Fig. 1. Setup WiFi and WiMAX multi-radio platform for railway applications

As shown in Fig. 1, the WiFi and WiMAX transceivers share the same multiband antenna and each device is connected to a diplexer (Microlab BK-26 series) which serves as a band-pass filter for frequencies in the range of 3.3 to 5.85 GHz (with 0.2 dB typical and 0.8 dB maximum band-pass losses). Each diplexer has an additional input port for signals ranging from 80 MHz to 2.69 GHz, intended to connect ZigBee and Bluetooth transceivers that are out of scope of this paper. For the experiments, 50 Ohm loads have been put in each port. The output of the two diplexers is combined using a broadband hybrid coupler (Microlab CA-13 series) with 3 dB nominal coupling loss and 18 dB isolation.

A summary of the WiFi and WiMAX specifications is given in Table 2. Even though several modulation and coding schemes are available in both devices, the presented performance evaluation have been based on the lowest and more robust configuration that yields a nominal throughput of 6 Mbps for both technologies. As far as the frequency of operation is concerned WiMAX has been set to 5.7 GHz. On the other hand, three different frequency channels have been selected for WiFi, in 5.18 GHz, 5.5 GHz and 5.735 GHz (channels 36, 100 and 147, respectively), in order to study the mutual interference between WiFi/WiMAX as their performance in different frequency bands with shared infrastructure.

Table 2. WiFi and WiMAX specifications

	Hirschmann BAT-300F	Alvarion BreezeACCESS VL
Technology	WiFi 802.11a	WiMAX 802.16
Frequency Channel	5.18 / 5.5 / 5.735 GHz	5.7 GHz
Channel Bandwidth	20 MHz	20 MHz
Lower Modulation Scheme	OFDM 64 FFT with BPSK	OFDM 64 FFT with BPSK
Nominal Min.Throughput	6 Mbps	6 Mbps
Receiver Sensitivity	-95 dBm (@ 6Mbps)	-89 dBm (@ 6Mbps)
Max Tx Power Level	18 dBm (@ 6Mbps)	21 dBm

The evaluation of the WiFi/WiMAX coexistence performance is based on several metrics:

- Spurious transmissions of one device on the frequency band of the other. For example we measure the interference caused by the operation of WiMAX at the 5.7 GHz band in the 5.18 GHz channel assigned to WiFi.
- The system throughput, defined as the total number of data bits successfully transmitted per time unit. The maximum throughput capacity has been measured for UDP and TCP flows generated by the traffic generation tool *iperf*[10].

4 Experimental Results

The WiFi and WiMAX coexistence has been evaluated through a set of experimental measurements. First, a parameter named Adjacent Channel Power (ACP) defined as:

$$ACP (dB) = P_{adjacent_channel}(dBm) - P_{operation_frequency}(dBm) \quad (1)$$

Where $P_{operation_frequency}$ is the total transmitted power by the wireless device within the allocated frequency bandwidth and $P_{adjacent_channel}$ is the total measured unwanted transmitted power by the same device on the adjacent sub-bands.

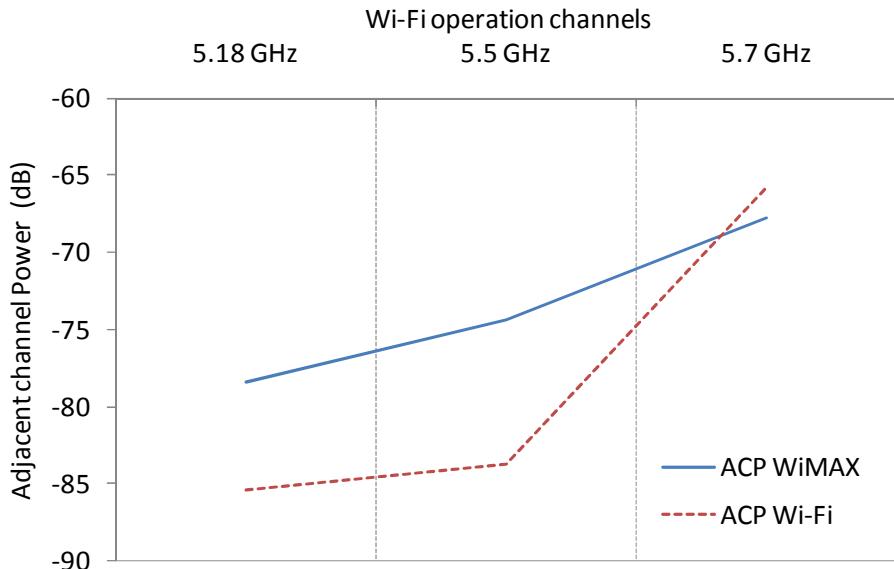
The ACP measurement was performed in two steps. Initially, we measured the interference caused by the WiFi device, located on the multi-radio platform, to the WiMAX band of operation considering three different WiFi channels, at 5.18, 5.5 and 5.735 GHz, with a bandwidth of 20 MHz in each case. In experimental tests with different WiFi sub-bands, the transmitted power was measured at the antenna port in point A of Fig. 1. Then, we measured the interference caused by each WiFi sub-band on the WiMAX channel at 5.7 GHz. These measurements took place at the input of the diplexer filter that corresponds to the WiMAX device (point C in Fig. 1). The WiFi transmission power and interference values are included in the upper part of Table 3.

In a similar way, we have measured the WiMAX interference on WiFi sub-bands. For this test, we have maintained the WiMAX operation frequency at 5.7 GHz with 20 MHz bandwidth and we have measured the total transmitted power at the antenna port A. Next, we measured the interference caused by WiMAX on the three different WiFi channels (point B in Fig. 1) and included these measurements in the bottom part of Table 3.

Table 3. WiFi and WiMAX interference power measurements

WiFi interference on WiMAX (5.7 GHz)		
WiFi operation frequency (GHz)	Average WiFi Tx power (dBm/20 MHz)	Average WiFi interference (dBm/20 MHz)
5.180	11	-74,3
5.500	11,2	-72,4
5.735	8,7	-57
WiMAX interference on WiFi (5.18, 5.25 and 5.735 GHz)		
WiFi operation frequency (GHz)	Average WiMAX Tx power (dBm/20 MHz)	Average WiMAX interference (dBm/20 MHz)
5.180	13,2	-65,2
5.500	13,2	-61,2
5.735	13,2	-54,5

The power and interference measurements have been employed for the calculation of ACP for the two systems, according to equation (1). The results are shown in Fig. 2. It can be seen that very near operating frequencies (i.e., WiFi operates at 5.735 GHz and WiMAX at 5.7 GHz) the WiFi Adjacent Channel Power is higher as shown the red dotted line in Figure 2. In addition, the WiMAX device is more likely to be affected by the interference than WiFi, due to its lower sensitivity level, as shown in Table 2. The value of ACP in the WiMAX channel (5.7 GHZ) was -76.87 dB in the case where there is no WiFi interference. On the other hand, the ACP in a WiFi channel (5.18 GHz) was -76.79 dB in the case where there is no WiMAX signal.

**Fig. 2.** Adjacent Channel Power for WiFi and WiMAX operation

Next, we measured the average throughput for TCP and UDP data flows, in both uplink and downlink (i.e., transmitting to or from the multi-radio platform). We have considered the following four cases:

- WiFi and WiMAX operating separately (no mutual interference).
- WiFi operating at the 5.18 GHz channel and WiMAX at the 5.7 GHz channel.
- WiFi operating at the 5.5 GHz channel and WiMAX at the 5.7 GHz channel.
- WiFi operating at the 5.735 GHz channel and WiMAX at the 5.7 GHz channel.

The results are shown in Fig. 3 and Fig. 4 for downlink and uplink, respectively. TCP throughout is represented by solid lines whereas dashed lines indicate UDP traffic. Several observations can be made from the figures. First, it can be clearly seen that performance is significantly affected when the two systems coexist in close frequency channels. Even with a frequency separation of approximately 520 MHz (when WiFi operates at 5.18 GHz and WiMAX at 5.7 GHz) throughput for both technologies drops significantly, especially in the case of downlink where WiFi and WiMAX share a single antenna for data transmission.

For TCP in downlink mode, the WiMAX throughput was affected by the interference from WiFi as shown in Figure 3. In this case both devices share the same antenna and both nodes try to seize the available bandwidth. In the same form, this phenomenon was observed in UDP traffic where an increment of UDP WiFi throughput decreases the WiMAX throughput.

Another interesting observation in the same sense is that throughput degradation occurs in two different ways. On one hand, in some cases both technologies are able

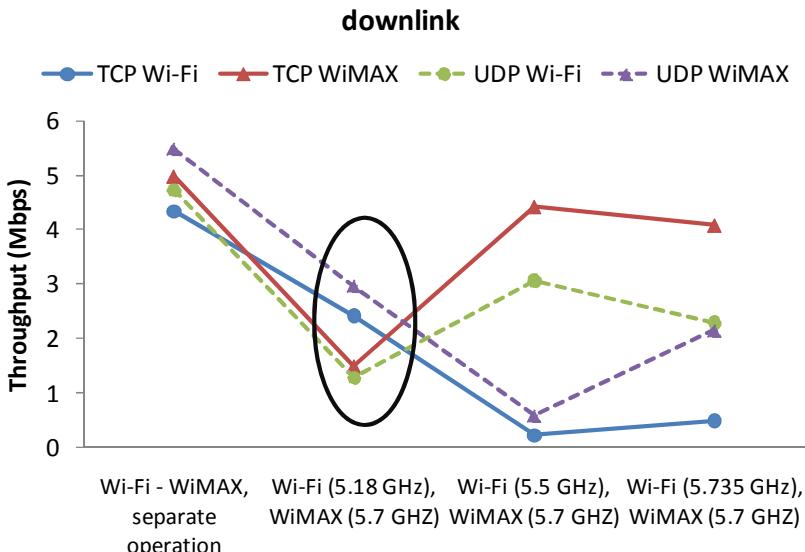


Fig. 3. WiFi and WiMAX TCP and UDP throughput in downlink

to transmit data but with a relatively low throughput. An example of such a case is marked by a circle in Fig. 3. On the other hand, there are cases where one technology (usually WiFi) seizes the channel and blocks completely the transmission of WiMAX as shown in Fig. 4.

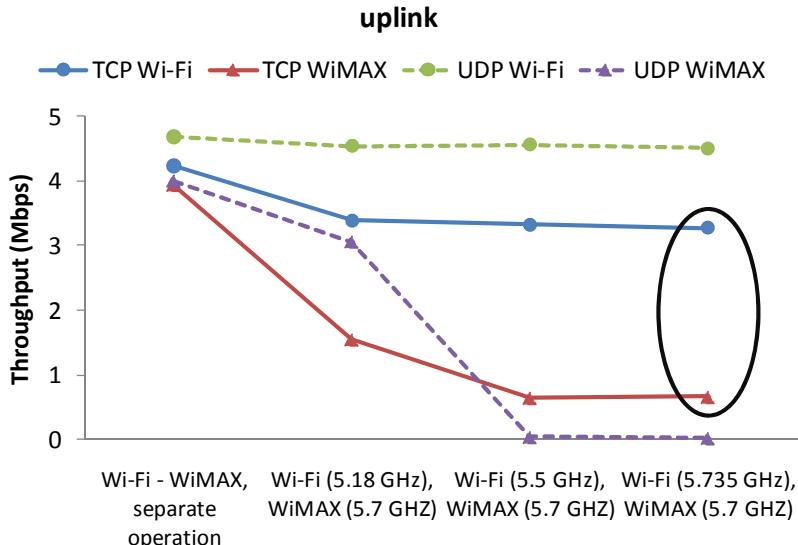


Fig. 4. WiFi and WiMAX TCP and UDP throughput in uplink

To better illustrate the downlink throughput fluctuations marked in a circle of figure 3, we have measured the instantaneous throughput for UDP downlink data transmission for WiFi operating in 5.18 GHz and WiMAX operating in 5.7 GHz , obtaining the fluctuations illustrated in Fig. 5 and consistent with the behavior observed in Figure 3.

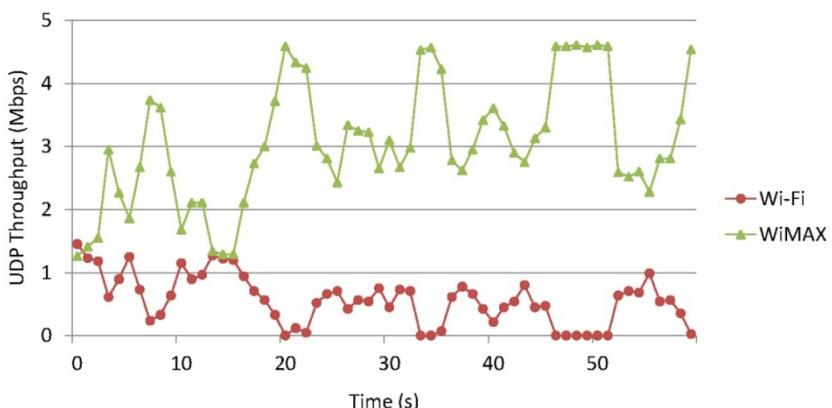


Fig. 5. Instantaneous UDP downlink throughput for WiFi and WiMAX

In this case, WiMAX performs better than WiFi in general. However, it can be observed that the instantaneous throughput for both technologies suffers from many fluctuations. In addition, when throughput drops for one technology, it increases for the other. This fact corroborates that the medium access control mechanism (MAC) is intensely affected by the interferences produced by both signals, and in some cases the channel is almost completely seized by one system while the other one sees the channel as busy and defers its transmissions.

5 Conclusions

The simultaneous operation of WiFi and WiMAX technologies with shared infrastructure using a multi-radio platform involves high interference levels that reduce the data rate performance despite the use of a system of diplexers and a hybrid combiner. A possible solution to mitigate the ACP interference could be the definition of separated (as much as possible) sub-bands within the operation bandwidth and even the use of a high selectivity pass-band filter to reduce the interference effects between sub-band. A proper hybrid combiner design with a higher isolation between its ports should be addressed to achieve this goal, which is left for a future line of research. By these means the coexistence of WiFi and WiMAX will be allowed in high dynamic environments such as high-speed railway transportation systems. Throughput measurements with WiFi and WiMAX simultaneous operation in the 5 GHz band show a significant reduction in the throughput for TCP and UDP data traffic. All tests have been carried out with transceivers that have different characteristics in terms of sensitivity and transmission power because we want to study the coexistence of two different wireless technologies in a railway environment to identify the effects of interference when the devices operating in adjacent bands. As a practical conclusion, better isolation between the two technologies in the 5 GHz band is required.

Acknowledgments. This work has been carried out within an ongoing project funded by the Spanish Government with the participation of ADIF (Spanish Railway Infrastructure Administrator), the Institute of Applied Magnetism (IMA) of the Complutense University of Madrid, SENER Engineering and Systems S.A., PLANYTEC and the Wireless Communications and Technologies Research Group (WiComTec), which belongs to the Signal Theory and Communications Department of the Technic-al University of Catalonia-BarcelonaTECH (UPC-BarcelonaTECH).

References

1. Administrator of Railway Infrastructures (ADIF): Infrastructure and Stations, High Speed Lines, http://www.adif.es/en_US/infraestructuras/lineas_de_alta_velocidad/lineas_de_alta_velocidad.shtml
2. Telefunken TRainCom® HST, http://www2.tfk-racoms.com/racoms/train_communication//highspeed_train/highspeed_train.php

3. IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems-Local and Metropolitan Area Networks-Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999), pp. C1–1184 (2007)
4. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments, July 15, pp.1–51 (2010)
5. IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, IEEE Std 802.16-2009 (Revision of IEEE Std 802.16-2004), pp. C1–2004 (2009)
6. Waltho, A.: Performance Analysis and Design Considerations for Multi-Radio Platforms. Intel Developer Forum, Taipei (2006)
7. Zhu, J., Waltho, A., Yang, X., Guo, X.: Multi-Radio Coexistence: Challenges and Opportunities. In: Proceedings of 16th International Conference on Computer Communications and Networks (ICCCN 2007), pp. 358–364 (2007)
8. Bouchez, B., de Coen, L.: Communication Systems for Railway Applications. In: Nagel, W.E., Walter, W.V., Lehner, W. (eds.) Vehicular Networking: Automotive Applications and Beyond, pp. 83–104. John Wiley & Sons, United Kingdom (2010)
9. MODURBAN Project. Homepage of the Modular Urban Guided Rail System (MODURBAN) Project, <http://www.modurban.org>
10. Iperf network testing tool. IPerf 2.0.0 released (2008),
<http://iperf.sourceforge.net/>

Train Tracking and Shadowing Estimation Based on Received Signal Strength

Hadi Noureddine^{1,2}, Damien Castelain¹, and Ramesh Pyndiah²

¹ Mitsubishi Electric R&D Centre Europe, France

² Telecom Bretagne, France

Abstract. In this work, we present an on-board solution for train position tracking that can be used in cases of GPS failures and that does not suffer from the error accumulation problem of Dead Reckoning (DR). It is based on Received Signal Strength (RSS) measured in radio communication systems by several mobile stations having antennas placed on top of different carriages of the train. As the RSS is affected by the slow fading or shadowing, both the position and the shadowing are jointly tracked. We estimate the shadowing atlas consisting of the shadowing maps along the railway of the different base stations. The proposed solution applies Bayesian filtering for efficiently processing the observations.

Keywords: Train tracking, received signal strength, shadowing, shadowing atlas, spatial correlation, Gaussian process, Bayesian filtering, particle filters.

1 Introduction

Reliable and accurate knowledge of train position and speed plays an important role in avoiding collisions and optimizing the traffic by increasing lines capacities. In classical train control systems, localization is based on track side equipments. The new trends in design of train control systems consist of integrating on-board solutions so as to reduce the need of track-side equipments with their inherent roll-out and maintenance costs and to simplify the deployment of new technologies and configuration changes.

Several methods for on-board position measurements are discussed in [1]. These methods are tachometers, Inertial Navigation Systems (INS) (e.g. accelerometers and gyroscopes), Doppler effect and GPS (and possibly other Global Navigation Satellite Systems such as GALILEO).

The GPS solution provides a good precision when sufficient non-obstructed satellite signals are available. When the GPS fails (e.g. tunnels, valleys, some urban areas), other solutions are required. In [2], dead reckoning the train position from on-board sensors (odometers and accelerometers) is performed during the GPS failures where a data fusion approach based on Kalman filtering is developed. All the above-mentioned technologies but the GPS are DR as the current position is estimated based on a previous position and an estimation of the traveled distance over elapsed time. The performance of DR deteriorates

with time due to error accumulation and a way is needed to compensate this error especially if the GPS failure lasts for a long time duration.

In this paper, we consider the use of RSS measurements in radio communication systems for on-board train positioning. A radio communication system can be either dedicated for communication between the train and the railway regulation control centers (e.g. GSM-R) or a public network offering services to passengers. This solution does not suffer from error accumulation of DR and can be integrated with on-board sensors that can give more information and improve the accuracy.

The RSS measurements are affected by the slow fading or shadowing. The shadowing can be divided into two parts: A time variant part caused by moving obstacles (e.g. a nearby train) and a time invariant one which is function of the position. In the following, we call shadowing the time invariant part. The shadowing in decibels (dB) is modeled by a Gaussian random process along the railway as it is spatially correlated [3] and log-normally distributed [4]. For one base station, we call this process a shadowing map. The shadowing of several base stations is also cross-correlated as reported in [7]. We define the shadowing ‘atlas’ being the collection of the maps of the different base stations. The knowledge of the shadowing improves the position estimation and is useful for communication purposes (e.g. handover anticipation).

The train position and the shadowing are jointly tracked. For this purpose, we apply Bayesian filtering for efficiently fusing the incoming measurements by recursively updating the posterior probability distributions of the position and the atlas. Due to the non-linearities encountered, sequential Monte Carlo methods known as particle filters [5] will be implemented for approximating the probability density functions.

This paper is organized as follows. In Section 2, the motion model describing the evolution of the train position and the RSS observation model are presented. In Section 3, we introduce the shadowing atlas and show how it can be updated based on RSS measurements when the train position is known. In Section 4, Bayesian filtering and particle filters implementation are presented. Simulation results and conclusions are presented in Sections 5 and 6.

2 System Model

In this section, we describe the motion model of the train and the RSS observation model. They allow for computing the a priori and likelihood probabilities in the Bayesian filtering processing, respectively.

2.1 Motion Model

The train is constrained to move on a known railway track and the position is defined as the Cartesian coordinate of a reference point belonging to the train.

The rail can be seen as a parametric curve of one parameter that we call rail coordinate. We assume that the length of the rail (in meters (m)) between two rail coordinates r_1 and r_2 is equal to the absolute value of $r_1 - r_2$. The train position can be described by the rail coordinate as there is a mapping to the Cartesian coordinates. The train speed is the derivative of the rail coordinate.

At time kT , where $k \in \mathbb{N}$ and T is a time step, we define the state vector \mathbf{s}_k comprising the scalar rail coordinate r_k , the scalar speed v_k and possibly other kinematic parameters (e.g. acceleration).

\mathbf{s}_k is issued from a known Markov process of initial distribution $p(\mathbf{s}_0)$ and transition probability $p(\mathbf{s}_k|\mathbf{s}_{k-1})$.

2.2 RSS Observation Model

We consider several mobile stations having their antennas placed on the top of different carriages of the train for a better sensitivity and penetration losses avoidance. The antennas are spaced enough (several meters to several tens of meters depending on the train size) so that the shadowing values are different at two antennas' positions for the same time instant. We denote N_{MS} the number of these antennas. RSS measurements are made in order to estimate the signal attenuation due to path loss and shadowing. They are usually obtained by low-pass filtering the received power in order to remove the small scale fading. The used filter must be sufficiently narrow in bandwidth to remove the multipath fluctuations, yet sufficiently wide to track the shadow fading.

A simple filtering solution is the sample-average which returns the average of a set of discrete time samples in dB . In [8], simulations in a flat Rayleigh environment with a carrier frequency of $900MHz$ show that a root mean square error of less than $1.6dB$ can be achieved by using the sample-average estimator. This result is obtained by considering a shadowing standard deviation of $\sigma_{sh} = 10dB$ and an exponential correlation function:

$$E\{\epsilon(\mathbf{x}_1)\epsilon(\mathbf{x}_2)\} = \sigma_{sh} \exp(-\ln(2)\|\mathbf{x}_1 - \mathbf{x}_2\|/d_{corr}) \quad (1)$$

where $\epsilon(\mathbf{x}_i)$ is the shadowing at position \mathbf{x}_i and $d_{corr} = 346m$ which is the value estimated in [3] for suburban areas. Under a Rician fading, the error is smaller as the variance of the Rice distribution decreases with the ratio of the specular power to the scattered power. As mentioned in [9], the error can be reduced if the channel bandwidth is sufficiently wide so that several paths can be resolved or if the mobile station has multiple nearby antennas (these antennas are spaced of the order of the wavelength and deliver one RSS measurement as they are affected by the same path loss and shadowing values). The remaining error after filtering is nearly Gaussian distributed.

Fast fading of train radio channels has been investigated in [10]. The measurements showed that only a weak multipath propagation exists and the radio

environment is mainly rural. The rural area model is also adopted for tunnels. These studies have also shown that a significant line of sight component exists especially when the base stations are on the track side.

We model the RSS measured by a mobile station having the antenna at position \mathbf{x} as

$$y = f(\mathbf{x}) + \epsilon(\mathbf{x}) + n \quad (2)$$

where f is a deterministic path loss function, ϵ is the shadowing and n is a Gaussian error gathering the time variant shadowing, average power estimation and other non-shadowing errors. Many general path loss prediction models for different deployment environments have been developed. We don't investigate here the path loss model that we assume to be known. In [10], a standard deviation of the shadowing between 1.2 and 3.7dB is reported over several railway lines.

3 Shadowing Atlas

3.1 Definition

For one base station, the shadowing is represented by a vector of real values where the rail is discretized and each vector entry corresponds to a discrete position. We call this vector a shadowing map. To obtain the value of the shadowing at an arbitrary position of the rail, we apply a piecewise constant interpolation by assigning the same value of the nearest point having an entry in the shadowing map. In fact, this shadowing map is not perfectly known and can be described by a Gaussian random vector of known mean and covariance. For several base stations, we define the shadowing atlas as being a collection of the shadowing maps obtained by a concatenation of the vectors.

We denote L_{map} the size of the shadowing map vector and $L_{atlas} = N_{BS} \times L_{map}$ the size of the atlas where N_{BS} is the number of base stations encountered along the railway. The size of the atlas covariance matrix is a large $L_{atlas} \times L_{atlas}$ sparse matrix thanks to the low shadowing spatial correlation between distant positions. We can also reduce N_{BS} and L_{map} by considering an atlas for a sector of the railway.

3.2 Atlas Update

Here, we describe how to update the atlas based on RSS measurements made at a known train position.

At time kT , we denote l_k the number of RSS measurements made by the N_{MS} train antennas. The positions of the antennas can be easily deduced from the rail coordinate r_k . The observation vector of size l_k is

$$\begin{aligned} \mathbf{y}_k &= \mathbf{f}_k(r_k) + \boldsymbol{\Sigma}_k(r_k) + \mathbf{n}_k \\ &= \mathbf{f}_k(r_k) + \mathbf{H}_k \boldsymbol{\Lambda} + \mathbf{n}_k \end{aligned} \quad (3)$$

where $\mathbf{f}_k(r_k)$ and $\boldsymbol{\Sigma}_k(r_k)$ are the path loss vector and the shadowing vector respectively, \mathbf{n}_k is the Gaussian error process, $\boldsymbol{\Lambda}$ is the shadowing atlas vector of size L_{atlas} , and \mathbf{H}_k is an $l_k \times L_{atlas}$ matrix satisfying $\mathbf{H}_k(i, j + L_{map} \times (l-1)) = 1$ if the i^{th} measurement is made with l^{th} base station and the corresponding shadowing has the j^{th} entry in the map vector, and otherwise it is equal to zero.

Denote \mathbf{M}_{k-1} and \mathbf{C}_{k-1} the mean and covariance of $\boldsymbol{\Lambda}$ at time $(k-1)T$ respectively and assume that the process \mathbf{n}_k is white with respect to time. The observation \mathbf{y}_k , which is a linear function of $\boldsymbol{\Lambda}$, is used to update \mathbf{M}_k and \mathbf{C}_k by means of a linear Minimum Mean Square Error (MMSE) estimator [6] as follows:

$$\mathbf{M}_k = \mathbf{M}_{k-1} + \mathbf{Q}_k(\mathbf{y}_k - \mathbf{f}_k(r_k) - \mathbf{H}_k\mathbf{M}_{k-1}) \quad (4)$$

and

$$\mathbf{C}_k = \mathbf{C}_{k-1} - \mathbf{Q}_k \mathbf{H}_k \mathbf{C}_{k-1} \quad (5)$$

where \mathbf{Q}_k is the gain factor given by:

$$\mathbf{Q}_k = \mathbf{C}_{k-1} \mathbf{H}_k^T (E\{\mathbf{n}_k \mathbf{n}_k^T\} + \mathbf{H}_k \mathbf{C}_{k-1} \mathbf{H}_k^T)^{-1}. \quad (6)$$

In the next section, we describe how to perform this estimation when the position is not perfectly known.

4 Bayesian Tracking Solution

In this section, we present the Bayesian filtering adapted to our tracking problem and its implementation using particle filters.

4.1 Bayesian Filtering

A Bayesian filtering consists of determining recursively in time the posterior distribution $p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k})$, where we denote by $\mathbf{z}_{0:k} = [\mathbf{z}_0^T, \dots, \mathbf{z}_k^T]^T$ for any sequence \mathbf{z}_k , \mathbf{y}_k is the RSS observation vector and $\mathbf{u}_{1:k}$ is an observation vector that might be obtained from INS sensors. The goal is to apply the MMSE estimator:

$$\hat{\mathbf{s}}_k = \int \mathbf{s}_k p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k}) d\mathbf{s}_{0:k}. \quad (7)$$

This posterior distribution allows to compute $p(\boldsymbol{\Lambda}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k})$ as follows:

$$p(\boldsymbol{\Lambda}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k}) = \int p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k}, \mathbf{y}_{1:k}) p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k}) d\mathbf{s}_{0:k} \quad (8)$$

where $p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k}, \mathbf{y}_{1:k})$ is a Gaussian distribution whose parameters can be estimated using the linear MMSE estimator described previously.

The distribution $p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k})$ is computed recursively according to

$$\begin{aligned} p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k}) &\propto \\ p(\mathbf{s}_{0:k-1}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k-1})p(\mathbf{s}_k|\mathbf{s}_{k-1})p(\mathbf{u}_k|\mathbf{s}_{k-1}, \mathbf{s}_k)p(\mathbf{y}_k|\mathbf{s}_{0:k}, \mathbf{y}_{1:k-1}). \end{aligned} \quad (9)$$

By assuming that the error process \mathbf{n}_k is white, the observations at different instants are independent given the atlas:

$$p(\mathbf{y}_i, \mathbf{y}_j|\mathbf{s}_i, \mathbf{s}_j, \boldsymbol{\Lambda}) = p(\mathbf{y}_i|\mathbf{s}_i, \boldsymbol{\Lambda})p(\mathbf{y}_j|\mathbf{s}_j, \boldsymbol{\Lambda}). \quad (10)$$

Thus, a possible computation of $p(\mathbf{y}_k|\mathbf{s}_{0:k}, \mathbf{y}_{1:k-1})$ is

$$p(\mathbf{y}_k|\mathbf{s}_{0:k}, \mathbf{y}_{1:k-1}) = \int p(\mathbf{y}_k|\boldsymbol{\Lambda}, \mathbf{s}_k)p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k-1}, \mathbf{y}_{1:k-1})d\boldsymbol{\Lambda} \quad (11)$$

Equations (7) and (8) are analytically untraceable since the observation equation (3) is non-linear with respect to r_k and \mathbf{H}_k depends on r_k . A solution based on particle filters [5], which apply sequential Monte Carlo methods for approximating numerically the posterior densities, is presented in the following.

4.2 Implementation Using Particle Filters

Assume that at time $(k-1)T$, $p(\mathbf{s}_{0:k-1}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k-1})$ is approximated by the set of M weighted trajectories $\{w_{k-1}^i, \mathbf{s}_{0:k-1}^i\}_{i=1}^M$:

$$p(\mathbf{s}_{0:k-1}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k-1}) \approx \sum_i w_{k-1}^i \delta(\mathbf{s}_{0:k-1}, \mathbf{s}_{0:k-1}^i) \quad (12)$$

where δ is the Kronecker delta function.

The atlas posterior distribution is computed according to (8):

$$p(\boldsymbol{\Lambda}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k-1}) \approx \sum_i w_{k-1}^i p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k-1}^i, \mathbf{y}_{1:k-1}). \quad (13)$$

The Gaussian distributions $p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k-1}^i, \mathbf{y}_{1:k-1}) = N(\mathbf{M}_{k-1}^i, \mathbf{C}_{k-1}^i)$ are obtained using the linear MMSE estimator by one of two methods: The first method is to update $p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k-2}^i, \mathbf{y}_{1:k-2})$ at each time step where a huge amount of memory is needed in order to store the M atlases. The second method is to compute it from scratch at each time step using all the previous observations but with a complexity that increases with time.

We propose a suboptimal solution that approximates the weighted mixture of Gaussian distributions $\{w_{k-1}^i, N(\mathbf{M}_{k-1}^i, \mathbf{C}_{k-1}^i)\}_{i=1}^M$ by a single Gaussian one $N(\hat{\mathbf{M}}_{k-1}, \hat{\mathbf{C}}_{k-1})$. This operation is repeated every q time steps by substituting $p(\boldsymbol{\Lambda}|\mathbf{s}_{0:k+q-1}^i, \mathbf{y}_{1:k+q-1})$ by $p(\boldsymbol{\Lambda}|\mathbf{s}_{k:k+q-1}^i, \mathbf{y}_{k:k+q-1})$ where the initial distribution $p(\boldsymbol{\Lambda})$ is substituted by $N(\hat{\mathbf{M}}_{k-1}, \hat{\mathbf{C}}_{k-1})$.

The posterior distribution $p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k})$ at time kT can be obtained in three steps that are summarized in Fig. 1.

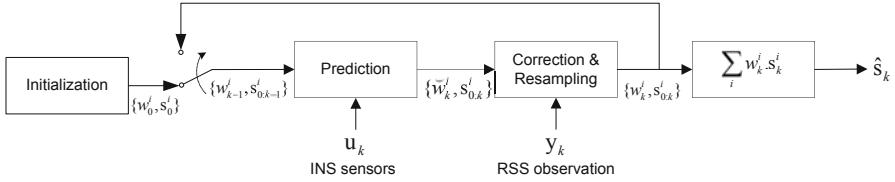


Fig. 1. Particle filter for estimating the state vector \mathbf{s}_k

Prediction Step. The predictive distribution $p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k})$ can be written as

$$\begin{aligned} p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k}) &\approx \text{constant} \times \sum_i w_{k-1}^i p(\mathbf{u}_k|\mathbf{s}_{k-1}^i) p(\mathbf{s}_k|\mathbf{s}_{k-1}^i, \mathbf{u}_k) \delta(\mathbf{s}_{0:k-1}, \mathbf{s}_{0:k-1}^i) \\ &\approx \sum_i \check{w}_k^i p(\mathbf{s}_k|\mathbf{s}_{k-1}^i, \mathbf{u}_k) \delta(\mathbf{s}_{0:k-1}, \mathbf{s}_{0:k-1}^i) \end{aligned} \quad (14)$$

where $\check{w}_k^i \propto w_{k-1}^i p(\mathbf{u}_k|\mathbf{s}_{k-1}^i)$. A Monte Carlo approximation of this distribution is obtained by drawing \mathbf{s}_k^i from $p(\mathbf{s}_k|\mathbf{s}_{k-1}^i, \mathbf{u}_k)$:

$$p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k-1}, \mathbf{u}_{1:k}) \approx \sum_i \check{w}_k^i \delta(\mathbf{s}_{0:k}, \mathbf{s}_{0:k}^i). \quad (15)$$

The measurement \mathbf{u}_k is included in this step since the INS observation models are in general linear with respect to the state and samples can be easily drawn from $p(\mathbf{s}_k|\mathbf{s}_{k-1}^i, \mathbf{u}_k)$.

Correction Step. In this step, the observation \mathbf{y}_k is used in order to compute $p(\mathbf{s}_{0:k}|\mathbf{y}_{1:k}, \mathbf{u}_{1:k})$ by updating the weights according to

$$w_k^i \propto \check{w}_k^i p(\mathbf{y}_k|\mathbf{s}_{0:k}^i, \mathbf{y}_{1:k-1}). \quad (16)$$

The value $p(\mathbf{y}_k|\mathbf{s}_{0:k}^i, \mathbf{y}_{1:k-1})$ can be estimated with two methods:

- Using (11) with the complexity or memory drawbacks mentioned above.
- Using the following equation:

$$p(\mathbf{y}_k|\mathbf{s}_{0:k}^i, \mathbf{y}_{1:k-1}) = \int p(\mathbf{y}_k|\boldsymbol{\Sigma}_k, \mathbf{s}_k^i) p(\boldsymbol{\Sigma}_k|\boldsymbol{\Sigma}_{0:k-1}, \mathbf{s}_{0:k}^i) p(\boldsymbol{\Sigma}_{0:k-1}|\mathbf{s}_{0:k-1}^i, \mathbf{y}_{1:k-1}) d\boldsymbol{\Sigma}_{0:k} \quad (17)$$

where $\boldsymbol{\Sigma}_{0:k-1} = [\boldsymbol{\Sigma}_0^T, \dots, \boldsymbol{\Sigma}_{k-1}^T]^T$ and $\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}_k(\mathbf{r}_k)$.

The distribution $p(\boldsymbol{\Sigma}_k|\boldsymbol{\Sigma}_{0:k-1}, \mathbf{s}_{0:k}^i)$ can be computed from the initial a priori atlas distribution $p(\boldsymbol{\Lambda})$, and $p(\boldsymbol{\Sigma}_{0:k-1}|\mathbf{s}_{0:k-1}^i, \mathbf{y}_{1:k-1})$ is obtained by a Kalman filter. As the trains run on nearly straight rails, the shadowing at a given time step is uncorrelated with the shadowing at much older time steps and the process $p(\boldsymbol{\Sigma}_k|\boldsymbol{\Sigma}_{0:k-1}, \mathbf{s}_{0:k}^i)$ can be replaced by a p -order process $p(\boldsymbol{\Sigma}_k|\boldsymbol{\Sigma}_{k-p:k-1}, \mathbf{s}_{k-p:k}^i)$ and thus limiting the complexity.

Resampling Step. In order to avoid the weights degeneracy due to the increase of the variance of the set $\{w_k^i\}_{i=1}^M$, a resampling step is performed. Thus, some trajectories of weak weights are eliminated and some of strong weights repeated. We execute this step when the effective number of particles $N_{eff} = 1 / \sum_i (w_k^i)^2$ falls below a threshold.

5 Simulation Results

In this section we introduce the motion and the observation model scenarios assumed for the computer simulations and show the results.

The motion of the train is modeled by a simple linear Markov process. The state vector is $\mathbf{s}_k = [r_k, v_k]^T$ where r_k is the rail coordinate and v_k is the speed. The state vector \mathbf{s}_k evolves according to the following difference equation:

$$\mathbf{s}_k = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \mathbf{s}_{k-1} + \begin{bmatrix} 0 \\ T \end{bmatrix} (a_{k-1} + e_{k-1}) \quad (18)$$

where a_{k-1} is the acceleration vector provided by an accelerometer and e_{k-1} is a Gaussian variable that accounts for the acceleration estimation errors. We take the standard deviation of e_{k-1} equal to $1m/s^2$ and the mean equal to zero (the accelerometer bias is assumed to be known, otherwise it can be added to the state vector and tracked by the particle filter). The time step is $T = 0.5s$. The train moves on a linear rail of length $4km$ and there are two base stations located at $[-1390, -1032]^T$ and $[1589, -688]^T$ (in meters) as shown in Fig. 2. The train is moving from the left to the right. The speed is linearly increasing with time from $200km/h$ to $250km/h$. The train takes $64s$ to make the $4km$.

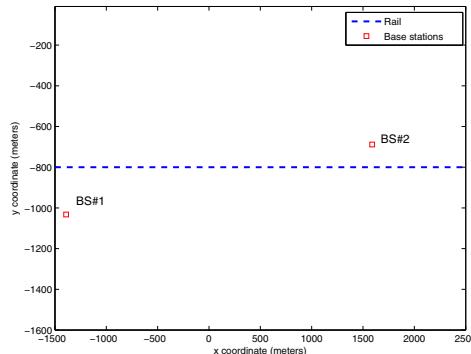


Fig. 2. Rail and base stations deployment

The shadowing has a standard deviation equal to $\sigma_{sh} = 4dB$ and is correlated according to the exponential function (11) and we set $d_{corr} = 200m$. The cross-correlation coefficient from two different base stations is constant and equal to 0.5, for simplicity. The initial atlas mean is equal to zero and the atlas covariance matrix is constructed according these parameters values.

The measurement error \mathbf{n}_k is a zeros mean white Gaussian process with a diagonal covariance matrix and diagonal entries equal to $4dB^2$. In fact, to have a white process, the error remaining after averaging out the fast fading has to be white, and this can be obtained by a displacement of the antennas of about $\lambda/2$ perpendicular to the motion direction, as shown in Fig. 3.

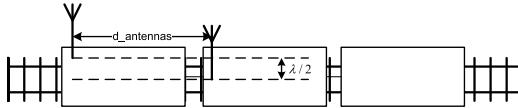


Fig. 3. Train antennas with a displacement of $\lambda/2$

For the path loss, we consider the Macro-cell system simulation model defined in [11] with omnidirectional base station antennas.

For the particle filter implementation, we take $M = 100$ particles. This number is sufficient since the dimension of the state vector is low (equal to two). We also replace the Gaussian process $p(\Sigma_k | \Sigma_{0:k-1}, \mathbf{s}_{0:k}^i)$ in (17) by a first-order process $p(\Sigma_k | \Sigma_{k-1}, \mathbf{s}_{k-1:k}^i)$. At low train speeds, it might be useful to consider higher orders of this process as the antennas at two non-consecutive time steps can be overlapping as shown in Fig. 4(b).

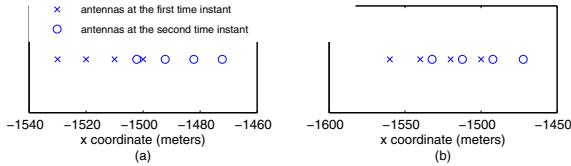


Fig. 4. Antennas positions at two consecutive time instants with separations of (a) 10m and (b) 20m. The velocity is 200 km/h.

In Fig. 5, we plot the Root Mean Square Error (RMSE) of the position. The RMSE of the DR solution based on the accelerometer observations is increasing with time. The initial position and speed are perfectly known.

The antennas placed on the train are equidistant. We can see that with 4 antennas, the performance is better when the distance between two consecutive antennas is 20m ($d_{antennas}$ in Fig. 3). A possible justification of this result is that the antennas are overlapping at two consecutive time steps when the separation is equal to 20m leading to a better estimation of the shadowing, while

this overlap does not occur for a separation of 10m as shown in Fig. 4. Moreover, the shadowing values affecting the different antennas measurements are less correlated for larger separations, and thus, leading to a higher diversity. We remark that the RMSE begins to decrease after about 40s as the train approaches $BS\#2$ since the path loss decreases logarithmically with the distance which is better estimated near a base station.

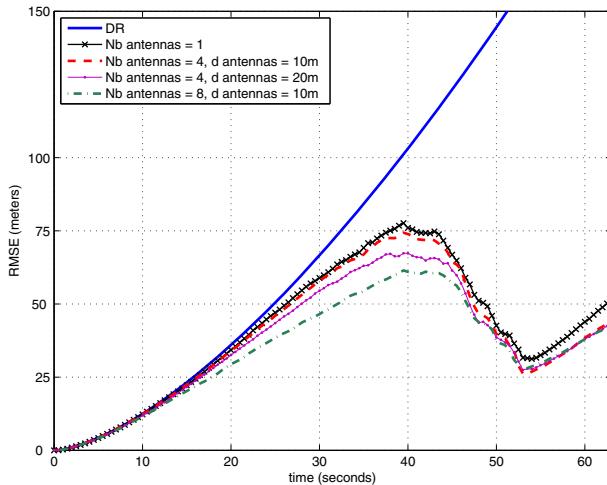


Fig. 5. RMSE of the train position

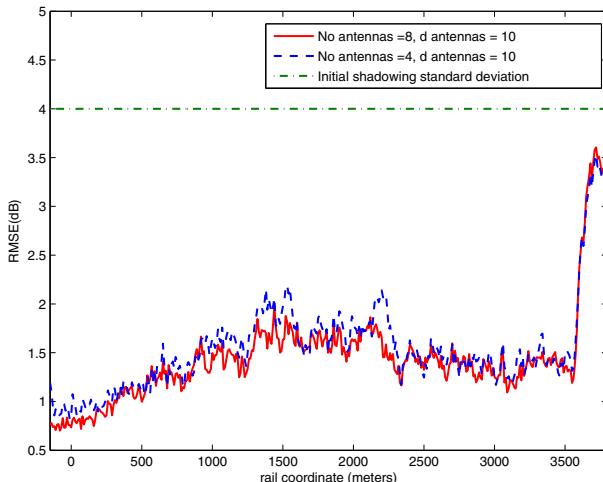


Fig. 6. RMSE of the Shadowing map estimation of $BS\#1$

In Fig. 6 the RMSE of the shadowing map of *BS#1* is plotted after one passing of the train. Later on, the obtained atlas distribution can be used as an initial distribution for other trains passing through the same railway.

6 Conclusions

We presented an on-board solution for train tracking based on the RSS measurements in a radio communication system. It can be integrated with INS sensors and can serve as a long time substitute of the GPS as it does not suffer from error accumulation of DR solutions. We also perform a joint estimation of the shadowing and construct the shadowing atlas.

The atlas estimation can be processed either on-line or off-line where the measurements made during the train travel are saved and processed by a server at the railway station. The initial distribution of the atlas can be downloaded at the railway station. An operation similar to the fingerprinting can be also performed by making RSS measurements at known positions and updating the atlas accordingly.

References

1. Mirabadi, A., Mort, N., Schmid, F.: Application of sensor fusion to railway systems. In: IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems, pp. 185–192 (December 1996)
2. Ernest, P., Mazl, R., Preucil, L.: Train locator using inertial sensors and odometer. In: Intelligent Vehicles Symposium, 2004 IEEE, pp. 860–865 (June 2004)
3. Gudmundson, M.: Correlation model for shadow fading in mobile radio systems. Electronics Letters 27(23), 2145–2146 (1991)
4. Rappaport, T.: Wireless Communications: Principles and Practice. Prentice Hall PTR, Upper Saddle River (2001)
5. Doucet, A.: On sequential simulation-based methods for bayesian filtering. Technical report (1998)
6. Kay, S.M.: Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory (v. 1). Prentice Hall, Englewood Cliffs (1993)
7. Sorensen, T.B.: Slow fading cross-correlation against azimuth separation of base stations. Electronics Letters 35(2), 127–129 (1999)
8. Jiang, T., Sidiropoulos, N.D., Giannakis, G.B.: Kalman filtering for power estimation in mobile communications. IEEE Transactions on Wireless Communications 2(1), 151–161 (2003)
9. Goldsmith, A.J., Greenstein, L.J., Foschini, G.J.: Error statistics of real-time power measurements in cellular channels with multipath and shadowing. IEEE Transactions on Vehicular Technology 43(3), 439–446 (1994)
10. Goller, M.: Application of GSM in high speed trains: measurements and simulations. IEE Colloquium on Radiocommunications in Transportation, 5/1–5/7 (May 16, 1995)
11. 3GPP. Tr 25.814, physical layer aspects for evolved universal terrestrial radio access (utra) (release 7) (September 2006)

Delivering Broadband Internet Access for High Speed Trains Passengers Using an Innovative Network Mobility Solution

Bernadette Villeforceix

France Telecom/ Orange Labs Networks and Carriers R&D
38-40 Rue du Général Leclerc
92794 Issy Les Moulineaux, France

Abstract. This paper presents a description of the connectivity solution designed in Orange Labs R&D to enable the delivery of broadband internet access services on board high speed trains. Since a couple of years, several technical studies have been launched on the network mobility management across different access networks for the implementation of broadband on board-ground communications solutions. These studies are supported by an internal project, Mobile Router in Multiple Access, launched by Orange Labs in 2005 in order to develop technical solutions for the delivery of broadband access services and which led to the build of a connectivity solution to answer requirements for the equipment of high speed trains. Orange deployed an infrastructure supporting these internet connectivity services. This paper explains in detail the technical features that are implemented in the mobility solution to provide a broadband internet access and guarantee an efficient and seamless mobility across heterogeneous access networks. A further evolution of the mobile router towards ITS architecture is approached in the paper.

Keywords: On board-ground communications, broadband Internet access service, network mobility management, mobile router.

1 Introduction

During the last few years, Orange Labs was involved in internal studies in the domain of mobility management architectures, for the design and development of an end to end connectivity solution capable of providing broadband access through various radio access networks.

These studies concerned the network access technologies, MAC and PHY layers of radio access networks such as two way satellite, WiFi and 3G/HSPA and also the mobility management protocols that were assessed to design and implement a complete solution for the delivery of internet connectivity in the transportation domain, for trains, cars and planes.

In the meantime, many railway operators have launched RFP for the delivery of commercial services based on board-ground broadband connectivity solutions aiming at providing onboard internet services to train passengers, high speed trains included.

Based on these technical choices, Orange has implemented an end to end connectivity solution that will be deployed in commercial services delivering Internet services to passengers onboard high speed trains.

One of the most critical challenges to solve at technical level is to hold a radio link in high speed conditions where the propagation channel is subject to fast fading.

Another critical challenge is to achieve a seamless handover between various radio access links, keeping the change of access network transparent for the end users.

A third critical challenge is to deliver connectivity speeds of a few Mbps to a single train. These throughputs are comfortable for the delivery of Internet access to some tens of passengers, but needed data-rates are going to grow very fast along with users demand.

These connectivity solutions will have to face a fast increase of the number of passengers willing to connect in the next years, in particular due to the very large success of smartphones.

To meet this future growing demand for this kind of services onboard trains, cars and planes, mainly in terms of bandwidth, future communication systems will need to exploit different access technologies at the same time, especially in dense traffic areas, to provide the required quality of service.

Orange Labs solution, using enhanced mobility management protocols, will be able to take up the challenge of this growing demand in high bit rate in high speed conditions.

Following chapters of this paper will introduce the architecture of the connectivity solution and the different functionalities supported by the mobile router implementation to cope with high bit rate requirements in high speed conditions.

2 The Mobility Architecture

As already mentioned, to design an end to end connectivity solution for the transport domain, the studies were oriented in two main axes:

- the study of radio access technology, two way satellite links, IEEE WiFi 802.11a/b/g/n and 3G/HSPA for the provision of the on board-ground link. PHY and MAC layers of these technologies were assessed to select the best configurations for high speed environments
- the study of the different mobility management architectures and protocols to design a mobile router capable of managing network mobility and supporting communications on simultaneous access networks while guaranteeing a seamless handover between those links at very high speeds.

The resulting network mobility architecture is presenting in figure1, it is composed of two network components for the mobility management: the mobile router (MR) that is on board the train and the Home Agent (HA) at the ground side, these network equipments are responsible for handling mobile IPv6 tunnels between them.

On board train, customers are connected in WiFi and their transmitted data streams are intercepted by the on board mobile router, encapsulated then routed in mobile IP tunnels (MIP tunnels) across access links towards the home agent that is the edge router of MIP tunnels. These data streams are then de-encapsulated and transmitted towards servers and internet according to applications required by passengers.

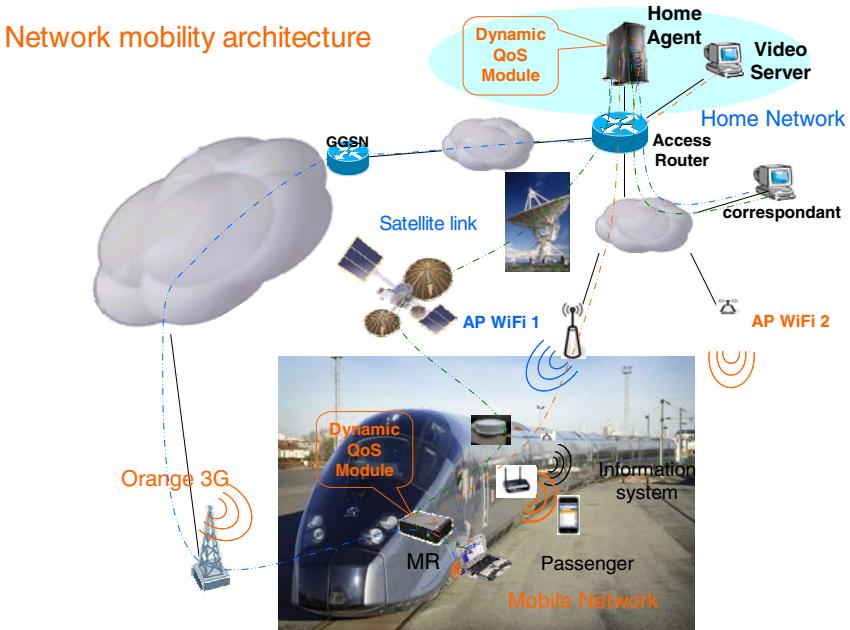


Fig. 1. Global mobility architecture

MIP tunnels are two way signaling links and support communications from board to ground and from ground to the board.

During the movement of the train, each time a new access network is available, its link interface in the mobile router gets the status "connected", the mobile router sends a signaling BU (Binding Update) to the HA.

The HA checks the BU and processes a Binding Acknowledgement BA transmitted to the MR, if the BA validates a previous sent BU, the status link is active and data are transmitted on that link.

At first step, mipv6 specified in IETF RFC 3775 only supported host mobility then it was ported on user space and renamed UMIP.

The evolution of MIPv6 protocol, adding mobile network prefix in IPv6 was then integrated in the solution so that the mobility architecture is now based on a NEMO platform according to RFC 3963 (implementation for mobility management) with MR integrated inside the vehicle and HA integrated in the operator facilities or at control center side. The mobile router developed in Orange Labs entails hardware technologies and software functionalities giving the benefit of supporting several access networks and managing multiple interfaces at the same time.

The Multi Care of address (MCoA), multihoming, multiple Care of Address registration specified in draft-ietf-moam6-multiplecoa, an extension of MIPv6 allows handling simultaneous mobile IPv6 links (NEMO) and multihoming over them and also switching between them according to their status.

To be capable of routing and addressing an IPv4/v6 address scheme on access interfaces and be compatible of IPv4/v6 networks, DSMIP (dual stack MIPv6

extension of MIPv6 capabilities to MN) specified in ITEF RFC 5555 was added to the previous implementation.

The scalability of the mobility architecture is achieved through the use of IPv6 protocol- capable of addressing billions of vehicles and sensors- therefore the adaptation to current network links by using DSMIP enables the transparent management of IPv4 or IPv6 routing rules on link interfaces.

2.1 Complementarity of Networks

The mobility architecture implemented in Orange Labs solution is capable of supporting several radio access links such as satellite link/WiFi/3G/HSPA to achieve the connectivity and handling several of them at the same time.

The multi interface router is capable of running multiple links and sending the data in an ordered scheme on the available links while being managed by the MCoA protocol; this function contributes to provide an increased network availability and capacity combined with an improvement of the radio coverage, taking benefit of network complementarity.

In case a radio access link fails, a handover is carried out to the resulting links according to re-routing rules defined in a module in charge of intelligent addressing and routing packets through the MIP tunnels.

The figure 2 illustrates the network complementarity.

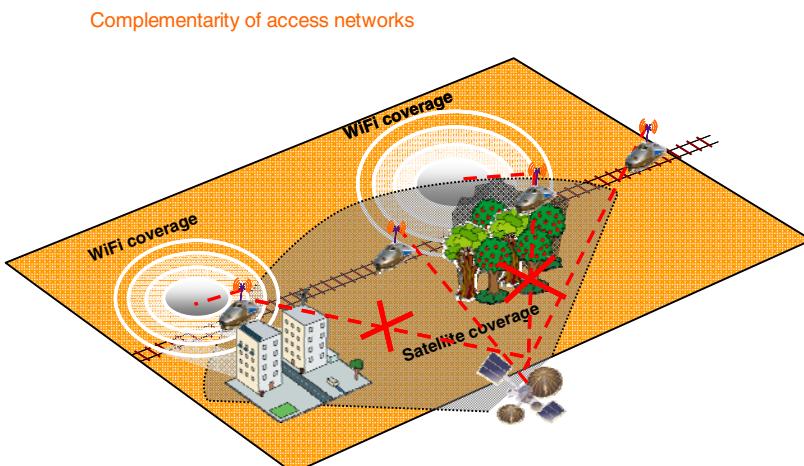


Fig. 2. Complementarity of access networks

This figure 2 illustrates the use of heterogeneous complementary access networks; the solution relies on a satellite link to communicate with the ground, the satellite coverage is considered as a global coverage but when the train enters an urban area or crosses a tree shadowing zone, the link is blocked so a handover occurs to switch the communications inside the WiFi coverage areas. A real case of this network complementarity is performed when the train enters a station where satellite coverage is not longer active.

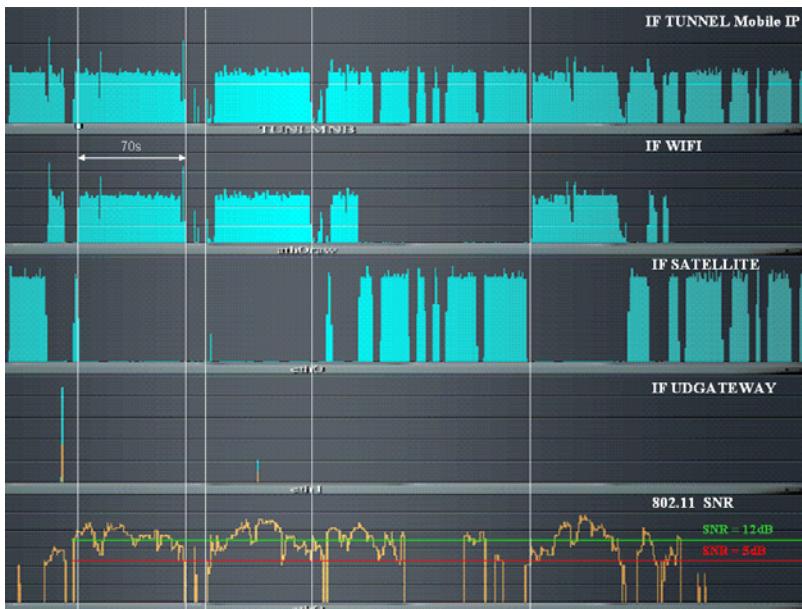


Fig. 3. Complementarity of satellite and WiFi access networks in high speed conditions (320Kmph)

A measured curve of the complementarity of satellite and WiFi access networks is illustrated in figure 3. The satellite and WiFi interfaces are monitored and the mobile IP tunnel is registered on the curves.

2.2 Soft Handover Management

To fit the mobility management requirements to fast variations of wireless links, the radio link interfaces are continuously monitored by the mobile router to which links data packets are to be transmitted to.

Handover is carried out by detection of a change in signal level (at L1/L2 layer) and/or a change of connection status (at L3 layer).

In case an interface status becomes inactive, the MIP tunnel is cancelled and all data packets are re routed to other active interfaces.

The only parameter that the end user will experience is the variation of the connection speed due to the change of radio access network and as the available bandwidth is shared by several end users, it can be smoothed with the help of upper layer protocols.

The robustness of the designed Handover mechanism applied to 3G/HSPA/WiFi/satellite has been successfully proved in high speed trains until a 320Kmph speed. The figure 4 shows the soft handover management.

The following figure 5 shows the registered handover in high speed conditions, from satellite to WiFi and from WiFi to satellite. This test has been performed with a downlink traffic composed of a 300Kbps UDP stream plus a 530Kbps TCP (video streaming).

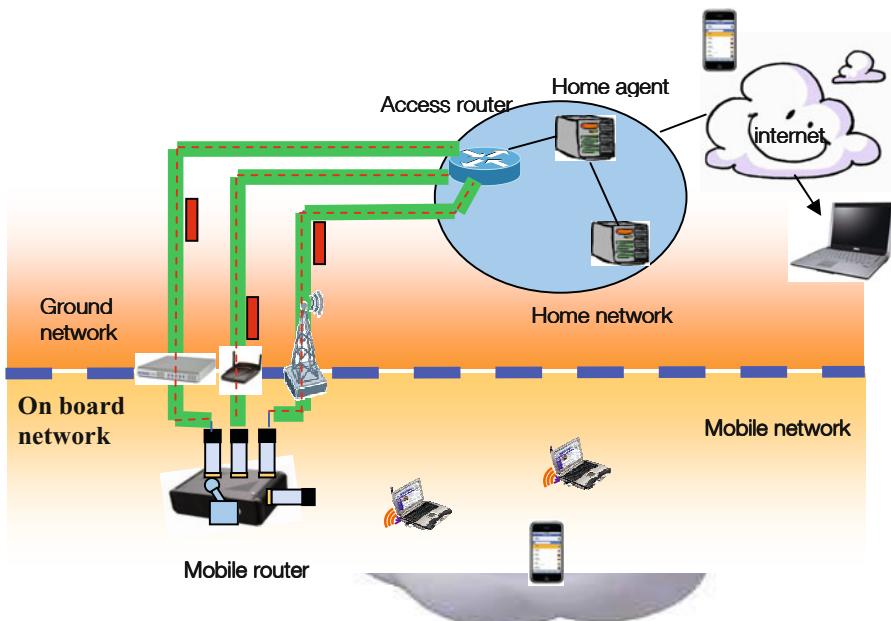


Fig. 4. Software handover management

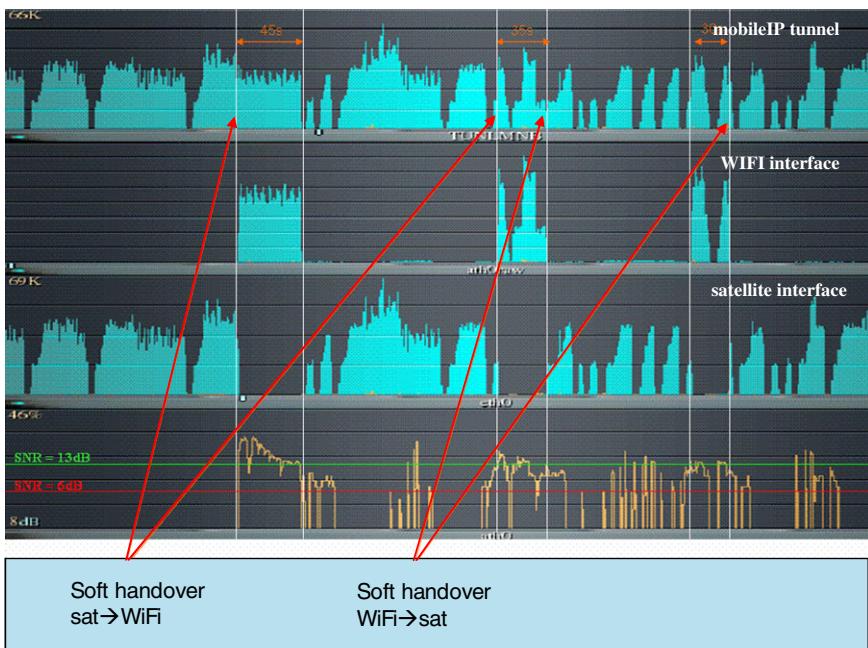


Fig. 5. Measured handover between different access links in high speed conditions

The following table 1 gives some performance values of the measured soft handover.

Table 1. Measured delays of horizontal and vertical handovers in high speed conditions (320Kmph)

	WiFi→WiFi	WiFi→3G	3G→WiFi
Scanning delay	Delay < n beacons (n*100ms)		
Handover delay	A few hundred of ms (~700ms)	A few tens to hundred of ms depending on 3G cell load	A few hundred of ms

2.3 Intelligent Routing Engine

A specific module is in charge of the routing policy by applying pre-determined rules to route the data streams to the active link interfaces; it entails a policy enforcement algorithm to react to sudden changes on interface status.

The MR and the HA continuously exchange routing policies to process BU/BA (Binding Update/Binding Acknowledgement) transmission and creation/cancelling of MIP tunnels.

2.4 QoS Policy

To enhance the mobile router with Quality of Service functionality, studies have been launched to design a solution capable of optimizing the use of the available bandwidth, in a constraint and frequently changing environment.

The access networks being the most sensitive links to radio changes due to dynamic environment (fast fading, interferences), the QoS function was applied in priority to them.

The mobile router, before routing data information to a particular radio link, 3G/WiFi/satellite, applies QoS policies by mapping specific interface parameters (real time, required throughput, jitter, network capacity) to specific application requirements.

It was shown in paragraph 2.2 that the mobility management is achieved by establishing and maintaining MIP v4/v6 tunnels over each physical interface between MR and HA. Each tunnel has its specific routing and QoS rules, all of them being automatically re-adjusted regarding the network interface load and the application requirements, by taking into account the available bandwidth variation.

Queuing files are automatically managed by handling traffic classes with different priorities on the MIPv4/v6 tunnels. The QoS management is performed by attaching queuing disciplines on each link interface in a hierarchical structure.

The enhanced capability of the QoS module is to assess dynamically the available bandwidth of the links to perform allocation and fair sharing of the packets streams into the links.

The figure 6 illustrated the queuing scheme of MR QoS policy.

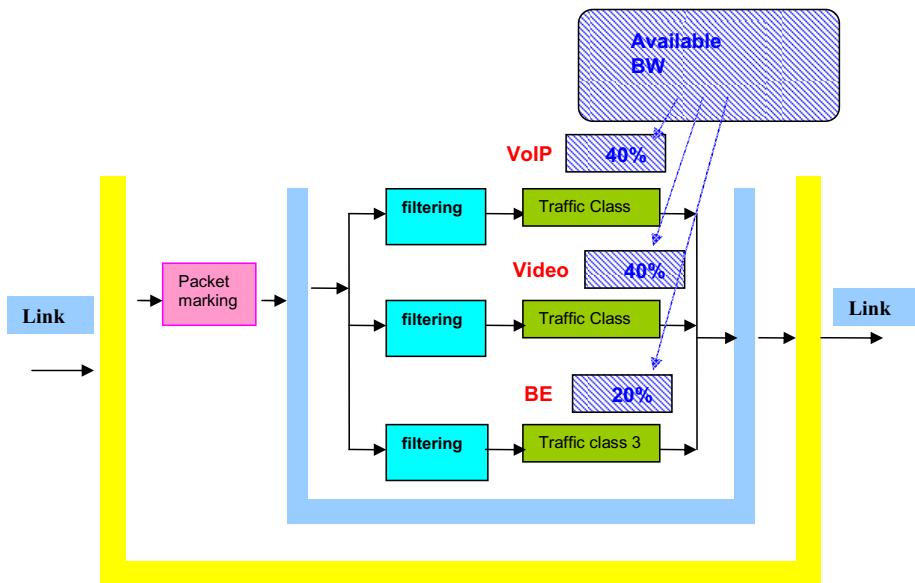


Fig. 6. QoS policy

The following figures 7 and 8 show the allocation of bandwidth regarding the application flows applied to a WiFi 802.11a (20Mbps) and 802.11b/g access link (5Mbps).

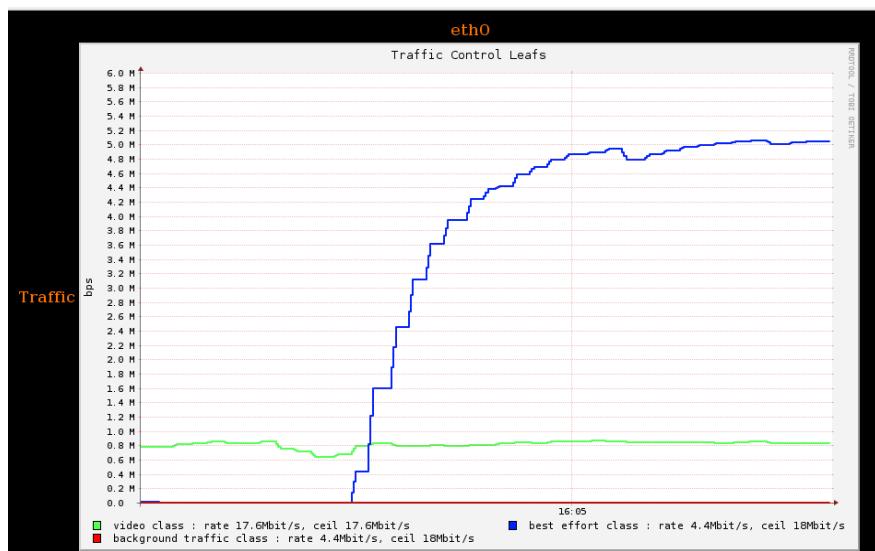


Fig. 7. Prioritized video traffic class according to the QoS policy in a 20Mbps available bandwidth

The prioritized flow, a video traffic class, has a guaranteed bandwidth 800Kbps while the 5Mbps BE traffic is filling the rest of the available bandwidth, on a 20Mbps radio link as shown in figure 7.

On the figure 8, the prioritized flow, the same video traffic class holds the same guaranteed bandwidth and the BE traffic is kept in a reduced bandwidth of 600Kbps, according to the available bandwidth dynamically measured on a 802.11b access link.

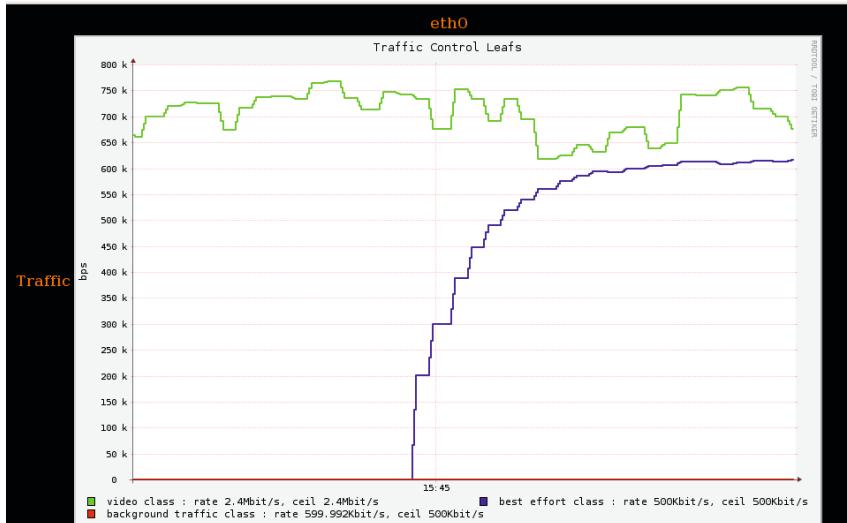


Fig. 8. Prioritized video traffic class according to the QoS policy in a 3Mbps available bandwidth

The main features of the mobile router are summarized by:

- MIP (Mobile IP) v6 capability (RFC 3775)
- NEMO implementation for mobility management (RFC 3963)
- MCoA (Multi Care of Address) IPv4/v6: multihoming, *Multiple Care-of Addresses Registration (MCoA) draft-ietf-monami6-multiplecoa*
- DSMIPv6 : Dual stack MIPv6 extension of MIPv6 capabilities to MN (RFC 5555)
- Access link Monitoring in MR (WiFi, satellite, 3G) at L2 and L3
- Handover decision and execution (horizontal and vertical HO)
- Routing, addressing and tunneling
- Flow management on link interfaces (policy enforcement)
- QoS policy : optimization of bandwidth utilization, dynamic management of interface bandwidth by queuing disciplines and traffic classification

The software implementation of the mobile router is illustrated in the figure 9:

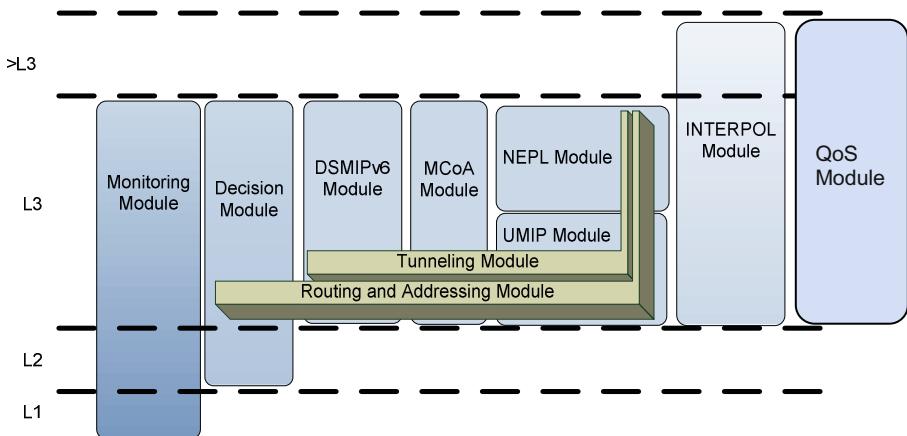


Fig. 9. Software implementation of the mobile router

3 Conclusion and Perspectives

This paper gave a complete description of the mobile router developed in Orange Laboratories for its application in trains, cars and planes.

All the implemented functionalities bring the following assets to this connectivity solution.

- Agility is given by the switching capability between several access technologies; the developed software solution can be fitted to different hardware according to the requirements and the possible selection of functionalities associated to considered services
- Efficiency is brought by the configuration of optimized radio parameters to achieve complementarity of access technologies, bit rate aggregation by handling several access links at the same time to achieve higher capacity, inter-technology handover and multihoming leading to seamless mobility
- Flexibility of this solution to fit the different transportation segments by a modular design of the mobile router, open integration of new access technologies and new IP functionalities (QoS and security), possible adaptation to various business models and role models

All these advanced functionalities enable the connectivity solution to support V2I and I2V communications. They can be applied to different transport segments, answering to various technical requirements. They can fulfill different declensions with fitted specifications.

The mobile router is now evolving towards a new domain of telecommunication ITS (intelligent Transport Systems).

Further studies are in progress to enhance the mobility solution towards ITS requirements to be able to support V2V communications and comply with ITS reference architecture.

References

1. IETF RFC 3775. Mobility Support in IPv6
2. IETF RFC 3963. Network Mobility (NEMO) Basic Support Protocol
3. IETF RFC 5555. Mobile IPv6 Support for Dual Stack Hosts and Routers
4. Multiple Care-of-Addresses Registration : draft-ietf-monami6-multiplecoa

Measurement and Analysis of the Direct Train to Train Propagation Channel in the 70 cm UHF-Band

Andreas Lehner, Cristina Rico García, Thomas Strang, and Oliver Heirich

German Aerospace Center DLR, Institute for Communications and Navigation,
Muenchner Str. 20, D-82234 Wessling, Germany

{andreas.lehner,cristina.ricogarcia,thomas.strang,
oliver.heirich}@dlr.de

Abstract. In this paper we present first analyses and results of a comprehensive measurement campaign investigating the propagation channel in case of direct (base station free) communication between railway vehicles. The measurements cover urban, suburban and rural environments along a multifaceted regional railway network in the south of Bavaria. Beside different operational conditions like front, rear, and flank approaches of trains, we investigated several topological scenarios on both, single and double track sections along the line. We will also discuss the observed characteristic changes in narrow band signal attenuation and Doppler spectra for passages through forests, hilly areas, stations and a tunnel.

Keywords: Propagation, channel, measurement, UHF, Train-to-Train, railway, path loss, model, V2V, RCAS.

1 Introduction

The design of Vehicle-to-Vehicle communication systems requires new knowledge about the specific propagation characteristics, which are certainly different to the widely measured and accurately modeled conditions in cellular terrestrial networks, where mobile nodes communicate with base stations at advantageous levels above ground. But not only has the low height of antennas raised uncertainties. Especially in the railway environment we face very specific conditions that influence the wave propagation along the lines. Shadowing by platforms with roofs in stations, or reflections from other railway vehicles might cause high losses. On the other hand, cuttings along the lines of railway networks with relatively large curve radii, catenaries, and tunnels are likely to have guidance effects [1] that could result in unexpected low path-loss exponents at larger distances.

New safety related applications in transportation are often based on direct communication among vehicles and have stringent requirements on the communication range. As an example the German Aerospace Center (DLR) is currently developing a Railway Collision Avoidance System (RCAS) [2] [3], that will allow the train driver to have an up-to-date accurate knowledge of the traffic situation in the vicinity, and act in consequence. The basic idea of RCAS is to calculate the own position and movement vector and broadcast this information as well as additional data like

vehicle dimensions to all other trains in the area. Thus, the train driver's cabin could be equipped with a display showing the position of the other vehicles in the region. Computer analysis of the received information, the own position and movement vector and an electronic track map allows to detect possible collisions, displaying an alert signal and advising the driver of the most convenient strategy to avoid the danger. In order to work reliable, the system requires a communication range that is in all cases larger than twice the maximum braking distances [4], which are usually in the order of 1-2 kilometers in regional networks.

Another challenge in the design of vehicular ad-hoc networks is to effectively control the common media access. While in cellular networks the media access is usually controlled by base stations, and the cell arrangement and frequency reuse is optimized to cause minimum interference from nodes in neighboring cells, the character of vehicular ad-hoc networks requires distributed solutions that strongly depend on the node density and on the communication range, respectively the interference range of nodes in the neighborhood [5]. Thus, it is important to have accurate knowledge of path loss and fading characteristics under realistic conditions, e.g. when regional lines meet shunting areas or large stations with high user densities, where both, communication with a fast approaching train and the information exchange to the multiple nodes in the close vicinity have to be reliably maintained at the same time.

2 Measurement Campaign

In the design phase of RCAS several questions arose concerning possible hindrances in fulfilling the requirements on the direct Train-to-Train communication link: Can the communication range be guaranteed to be significantly larger than the braking distances under all topological and operational conditions in different environments? How do passages through urban areas, forests, hilly areas or tunnels affect the statistics of successfully transmitted packets?

In order to be able to get an answer to these questions we performed a measurement campaign in an area covering all the interesting environmental, topological and operational aspects. We identified the regional network of the "Bayerische Oberlandbahn (BOB)" to be ideally suited for our measurement campaign and appreciated the extraordinary level of support we received from BOB upon our request for cooperation.

2.1 Railway Network and Environment

The route network of the BOB can be seen in **Fig. 1**. The 3 lines of the BOB run as a combined train-set from Munich central station to Holzkirchen. This part of the network has two tracks, is electrified and is used together with the Munich S-Bahn service. In Holzkirchen the combined train-set separates with one train-set heading off to the east running to Bayrischzell. The two remaining train-sets continue on to Schäftlach where they separate again, with one train-set going to Lenggries, while the last train-set heads off to the southeast to Tegernsee. On these 3 branches the lines are single track and are not electrified.

The BOB operates 20 diesel-hydraulic train-sets, which carry about 15000 people per day, with speeds up to 140 km/h [6]. The total network has 27 stations and covers 120 km, connecting the metropolis of Munich with the pre-alpine region in the south of Bavaria. Starting from the pure urban environment near the Munich central station the line passes through a tunnel, gradually leaving the city through suburban areas into the forest rich part around Holzkirchen. Finally in the south the three lines enter a hilly area with lakes in the alpine upland. **Fig. 2** shows two typical environmental conditions on the BOB network.

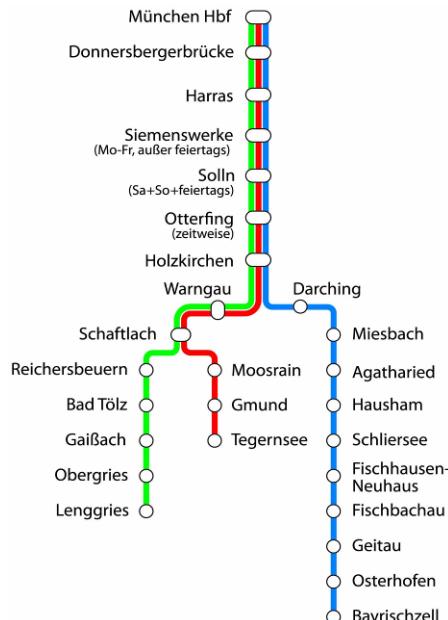


Fig. 1. Railway network of the “Bayerische Oberlandbahn (BOB)” in southern Bavaria [6]



Fig. 2. A BOB train-set on the electrified two track line between Munich and Holzkirchen (left) and in the hilly alpine upland near Bayrischzell (right) [7]

2.2 Measurement Setup

In regular operation the BOB train-sets meet at 3 distinct locations: At the stations in Holzkirchen and Schaftlach, where the train-sets are separated and connected, respectively, and at a point north of Holzkirchen, where 2 trains pass by each other on their way to and from Munich. In order to measure a variety of scenarios at many different locations along the network, we decided to use a car instead of a second train as receiving station for the measurements as depicted in **Fig. 3**. The car also allowed us to simulate hypothetical railway collisions, which would not be realizable with two trains during regular operations, because of safety regulations and the general railway network topology.

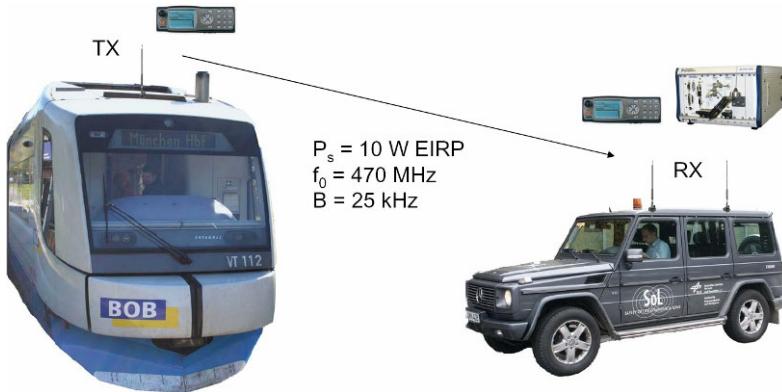


Fig. 3. Setup for the “Train-to-Train” channel measurement campaign: A car replaces the second train to allow for a variety of measurement scenarios at different locations

The transmitter was installed into one of the trains with the antenna mounted above the driver’s cabin. The measurement signal was transmitted by a Tetra radio at a carrier frequency close to 470 MHz with a power of 10 W EIRP, and had a bandwidth of 25 kHz. With a repetition rate of 1 Hz we transmitted a 150 bit SDS (Short Data Service) message containing traffic relevant information for the RCAS system as described in [8]. Prior to each RCAS message a synchronization sequence was transmitted, which is necessary for reception of the message in the Direct Mode of Operation (DMO), where the involved Tetra terminals are unsynchronized [9].

On the receiver side another Tetra radio recorded the successfully decoded messages. In addition we used a RF vector signal analyzer to record the received signal in time domain as shown in an example in **Fig. 4**. According to the call description for unacknowledged short data in direct mode of the Tetra standard [10], the synchronization covers frames 17 and 18, followed by the SDS data in slot 1 of frame 1. The pre-emption signaling indicates the priority request for further transmissions. The example in **Fig. 4** was recorded during a pass by at moderate speed, and it clearly shows the fading due to multipath. In the later analyses of the propagation channel characteristics this synchronization sequence is evaluated.

Both vehicles also carried a GPS receiver and stored PVT (Position, Velocity and Time) information during the measurements.

Although it does not scale for areas with dense railroad traffic, the first RCAS system demonstrator was built using Tetra radios. Since the provided data rate by the Tetra standard is very similar to the RCAS system requirements [8], this allowed us to investigate the message transmission performance under realistic propagation conditions. In this way both, the performance evaluation of the Tetra DMO-SDS messaging for Train-to-Train conditions and the corresponding channel characterization could be achieved.

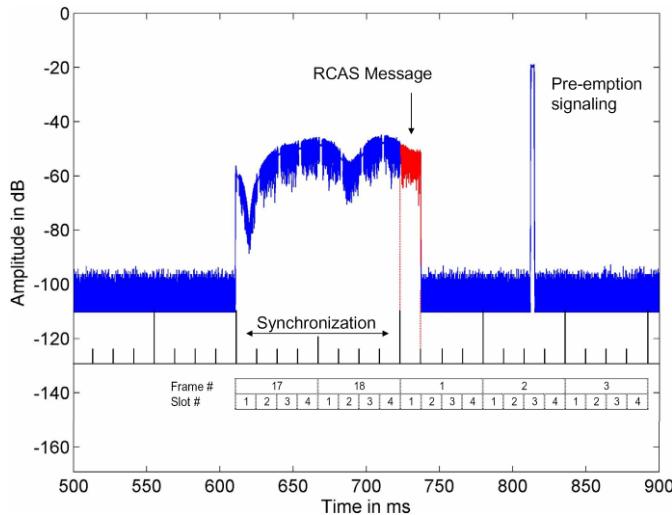


Fig. 4. Example of the recorded measurement signal during a pass by with moderate speed, clearly showing signal fading due to nearby multipath

2.3 Measurement Scenarios

During the whole measurement campaign 26 scenarios were measured, collecting 16 hours of data on more than 700 travelled kilometers. Two methods were applied:

- Static RX measurements: The car with the RX equipment was placed just aside the railway track or on a bridge above the track, and recorded the measurement signal while the train was approaching and receding the measurement location.
- Dynamic RX measurements: The car was driven along roads just aside the railway track to simulate a moving train. Thus we could investigate hypothetical collision scenarios like front, rear-end, and flank collisions with the approaching train.

Propagation conditions in railway environment, especially shadowing effects and multipath, may vary significantly depending on the local topology of the railway

network. For example in stations the tracks are usually separated by platforms with roofs that might block the LOS (Line of Sight), while in shunting areas LOS is reasonable for larger distances, but it is usually accompanied by strong multipath from other vehicles and the particular environment. The speed and therefore the signal fading is significantly slower near stations than on the sections of the regional network between them. Urban, suburban and rural passages are expected to show different path loss characteristics, depending if there are houses, trees or slopes alongside the track. Thus measurements were conducted in the following environments:

- Urban main station (Munich central station)
- Suburban stations (Solln, Holzkirchen, Bad Tölz)
- Shunting area (near Munich central station)
- Tunnel (near Donnersberger Brücke)
- Forest and farmland passages (multiple locations)
- Hilly alpine upland (Bayrischzell, Lenggries, Tegernsee)

3 Data Analysis

Prior to the measurement campaign we made a calibration measurement, where we recorded an undisturbed version of the specific Tetra synchronization sequence $s(t)$ that is sent by the TX radio. The power spectral density of $s(t)$ can be seen in **Fig. 5**.

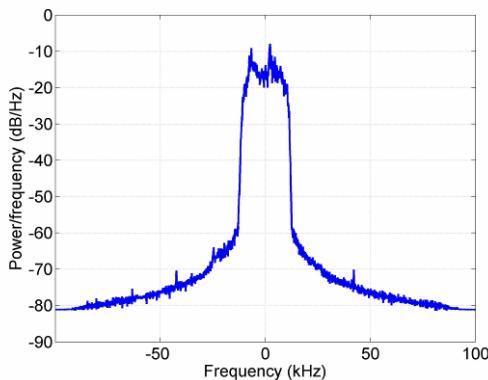


Fig. 5. Power spectral density of the Tetra synchronization burst that is used as measurement signal for the characterization of the Train-to-Train propagation channel

Distorted by the propagation channel and additional noise, the received signal $y(t)$ is given by

$$y(t) = h(t) * s(t) + n(t) , \quad (1)$$

where $h(t)$ denotes the momentary channel impulse response and $n(t)$ is an additive white Gaussian noise (AWGN) signal. For AWGN the Maximum-Likelihood

estimation of the channel impulse response is given by cross-correlation of the received signal $y(t)$ and the transmitted signal $s(t)$, or expressed in the frequency domain, the momentary channel frequency response $H(f)$ is given by

$$H(f) = Y(f)S^*(f)/|S(f)|^2 = Y(f)/S(f), \quad (2)$$

with $Y(f)$ and $S(f)$ being the Fourier transform of $y(t)$ and $s(t)$, respectively [11]. Assuming perfect synchronization between transmitter and receiver, $H(f)$ can be calculated from the DFT (Discrete Fourier Transform) of the sampled receive signal divided by the DFT of the transmit signal.

In our measurements the TX and RX stations were clocked by standard internal crystal oscillators, which causes an error when calculating $H(f)$. This error is mainly a slowly varying offset in Doppler due to drift changes of the oscillators. In the measurement data this offset can be easily determined, provided that PVT data of both stations is available. For example in sections when both, RX and TX were at a stop, the Doppler of the received signal shall be zero.

Fig. 6 illustrates this correction. In the left plot we see the channel frequency response in dB over a complete measurement run that lasted more than 30 minutes. Each second the received synchronization burst was evaluated. During this measurement the receiver was static at one location. The white lines indicate the maximum possible Doppler according to the velocity of the transmitter. Thus, each stop of the train at a railway station is represented by a section where the white curves equal zero. The corrected data can be seen in the right plot of **Fig. 6**, after estimating the clock drift and removing the corresponding Doppler offset. It shows a very good match and clearly allows to distinguishing between LOS and multipath components, as we will see in the following analysis. Note that the SNR is approximately 30 dB and that the frequency resolution in these plots is about 6 Hz due to the finite length of the synchronization sequence $s(t)$ and the sample rate of 200 kHz. Thus, according Fig. 6, 8, 10 and 13 there seem to be spectral components outside the theoretical Doppler bandwidth. However, these are effects of the limitations of the sampling.

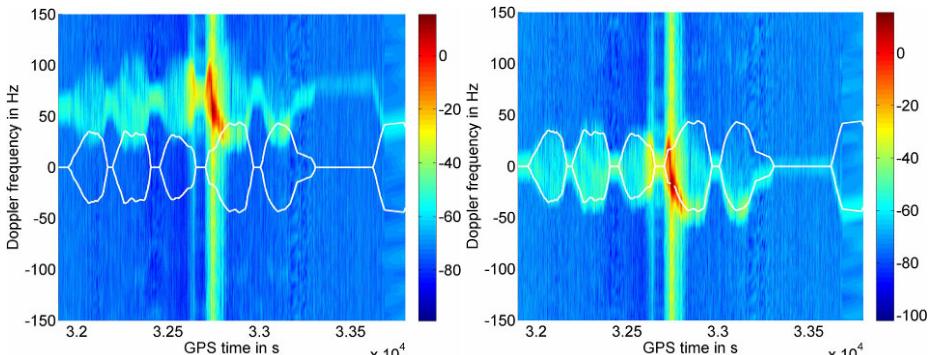


Fig. 6. Changes in the channel frequency response $H(f)$ in dB during a complete measurement run. The white lines indicate the maximum Doppler due to RX and TX movement. Left: offset due to RX and TX clock drift; Right: corrected clock drift offset.

3.1 Suburban Railway Station and Forest

The following data was recorded at the railway station in the small Bavarian town of Bad Tölz in a static RX measurement. The train approached from the south west, stopped at the station just 200 meters in front of the RX position, which is marked by the yellow triangle in **Fig. 7**, and finally continued passing by the RX position, leaving the town into some minor forest passages. The TX track is indicated with red diamonds, the size of which show the HDOP (Horizontal Dilution of Precision) recorded by the GPS receiver. It can be seen, that the HDOP significantly increases in forest sections and even when passing beyond a street (see right top corner of **Fig. 7**).

When the train approached the station between time stamps 100 to 150 seconds in **Fig. 8**, we can see a clear increase in power. The direct wave components are at approximately +20 Hz, but there are as well several components from scatterers that reach



Fig. 7. Suburban measurement scenario in the vicinity of the railway station in Bad Tölz. Red diamonds indicate the route of the train; the yellow triangle shows the RX position.

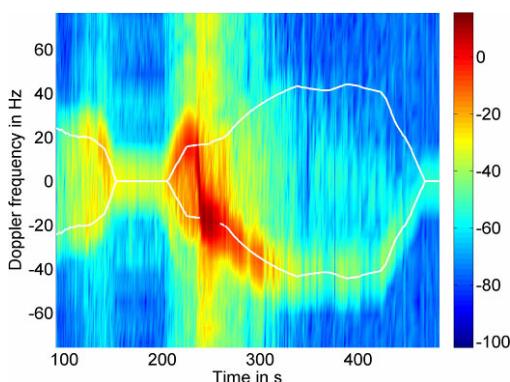


Fig. 8. Measured channel frequency response in dB at the railway station in Bad Tölz

almost the same power level. Please note, that the power in **Fig. 8** is normalized to the hypothetical FSL (Free Space Loss), i.e. we only show the additional loss due to diffraction and reflection. In other words, the red parts with 0dB refer to LOS condition.

During the halt from 150 to 200 seconds, we can see a characteristic lower power caused by the roofs on the platforms, which block the LOS signal. Immediately after leaving the station the LOS path appears, quickly changing its Doppler from +20 to -20 Hz when the train passes by the receiver at time stamp 240. In the whole area of the station strong multipath was present from different directions causing a maximum Doppler spread, until the LOS signal vanished as the train entered a forest passage.

In the last part of the measurement between time stamps 350 and 450, the relative Doppler spread is lower, but we can still see scattering from objects and vegetation on both sides of the track with a Doppler frequency around zero.

3.2 Tunnel in Urban Environment

The following measurement example was recorded near the Munich central station. Again the receiver was static at a location very close to the railway track (see yellow triangle in **Fig. 9**). The train started at Munich central station and had its first stop just beyond the Donnersberger Brücke in the upper part of the image. Right after this stop the train entered a tunnel that crosses under other tracks, allowing to directly going south passing by the receiver location. This tunnel is about 280 meters long and is indicated by the white line in **Fig. 9**. The red diamonds again show the TX track, showing increased HDOP values, when passing under the bridge, at the tunnel entrance and exit, and near higher buildings, where less GPS signals can be tracked.

The measured channel frequency response of this scenario is shown in **Fig. 10**. The train starts to leave Donnersberger Brücke at time stamp 130. Between time stamp 140 and 150 some fading is notable on the direct signal at the moments when the



Fig. 9. Urban measurement scenario near Donnersberger Brücke in the city center of Munich. The white line marks the tunnel that is used to cross multiple other tracks.

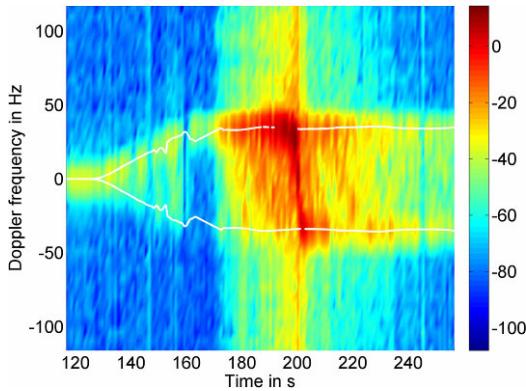


Fig. 10. Measured channel frequency response in dB in the urban environment near the tunnel at Donnersberger Brücke in Munich

antenna passed behind the abutments of the bridge. At the same time right under the bridge the GPS velocity measurements become erroneous, which can be clearly seen in the white line that marks the maximum possible Doppler. Soon after that, at time stamp 160, the train-set with the antenna entered the tunnel via a ramp. A short significant signal loss is visible at that moment and all the multipath components disappear. It is very interesting to note, that in the following 10 seconds, when the train passes through the tunnel, the received signal level is almost as high as before. This can be best observed in **Fig. 11**, which shows the evolution of the normalized received power for the same section of the measurement. Note that this corresponds to the evolution of total path loss including the FSL.

After the train exits the tunnel, the signal strength increases rapidly, as we begin to have LOS condition. At time stamp 200 the train passes by the RX station heading further south. Very strong multipath exists in this urban environment (see **Fig. 10**), which tends to show a typical Jakes spectrum characteristic [12], as the transmitter departs from the receiver location after time stamp 220.

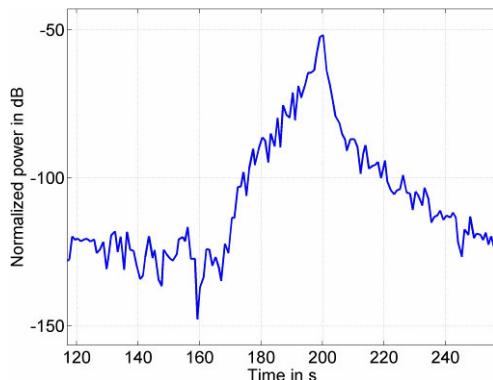


Fig. 11. Measured relative power in urban environment near the tunnel in Munich, showing the evolution of path loss in this scenario

3.3 Hilly Alpine Upland

In this scenario both, RX and TX moved towards each other. The yellow line in **Fig. 12** shows the track of the car following the railway line south along lake Schliersee, while the train approached the valley from the east. The hills east of the lake are obscuring a direct propagation between RX and TX in this scenario.

Despite these hills and the large distance of more than 10 km, we received most of the transmitted SDS messages. The channel frequency response in **Fig. 13** shows that there is reception of relatively strong multipath components in this situation. The strength of the multipath increased continuously until the train reached the village south of the lake, still under non-LOS condition. Around time stamp 250 we can also see a longer signal outage, which occurred during the halt of the train in Fischbachau (at the right lower corner in **Fig. 12**).

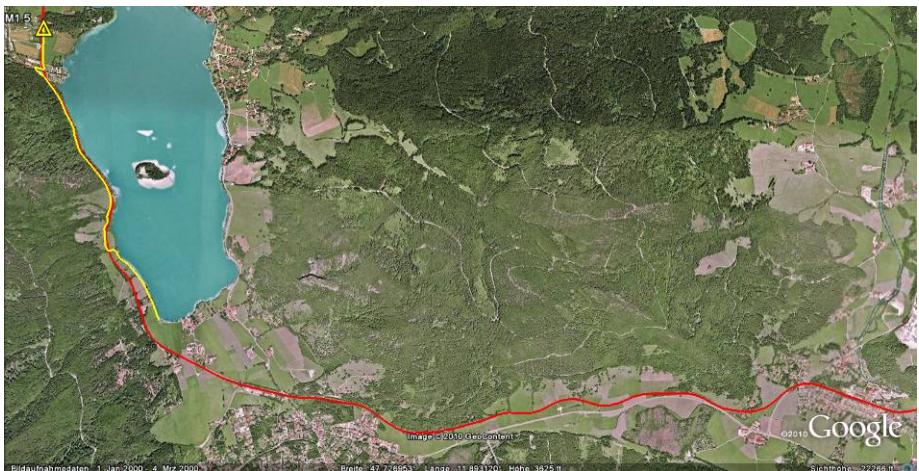


Fig. 12. Dynamic RX measurement scenario in the hilly alpine upland south of Schliersee

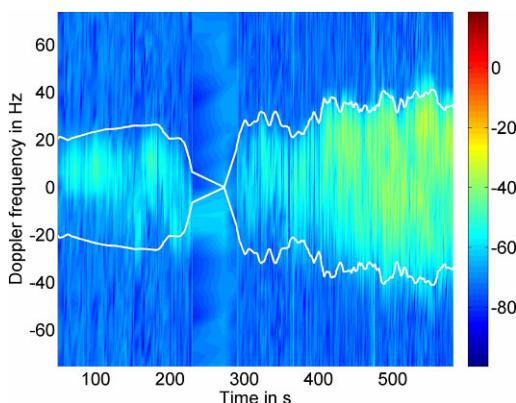


Fig. 13. Measured channel frequency response in dB in the hilly alpine upland near Schliersee

4 Path Loss Model

In [13] existing channel models for terrestrial communication and their applicability to Train-to-Train scenarios is discussed. As a result Hata-Okumura models were proposed for modeling the path loss in different railway environments. Hereby the suburban characteristic best fits the mix of scenarios that we measured during our campaign. In **Fig. 14** we compare the averaged path loss of the entire measurement data (blue dots) with the theoretical model. First of all, we note that in case of direct Train-to-Train communication the LOS condition is maintained for larger distances than in other terrestrial applications, typically up to more than 100 meters. While the exponential decay is in the expected order, the path loss for non-LOS condition was observed to be 10 dB lower than predicted, which in fact allows for a 100% increase of the communication range.

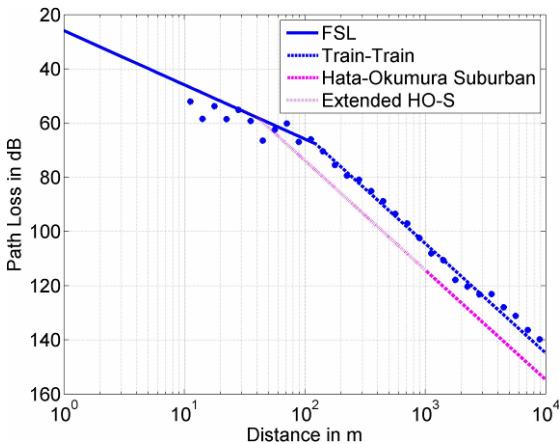


Fig. 14. Comparison of theoretical path loss model and Train-to-Train measurement result

5 Conclusion

The analysis of data from the presented channel measurement campaign shows, that applications requiring reliable direct Train-to-Train communication can benefit from the specific environmental conditions. Due to the layout of railway lines and the typical character of passages through urban, suburban and rural areas, the wave propagation for UHF frequencies is significantly better than in other Mobile-to-Mobile applications. Moreover, for higher velocities the train routes need to be straighter and the resulting longer breaking distances are likely to be still covered by the enhanced communication range. In fact, we see an unexpected high margin for timely warning of train drivers in case of a collision threat from the successfully received RCAS data during the measurement campaign. Even in critical environments with obstacles like hills, or in case of a tunnel, the performance of Tetra SDS-DMO messaging could be proved to be highly reliable.

Acknowledgments. The authors express their thanks to the staff and management of the “Bayerische Oberlandbahn” for their help and cooperation in performing the measurements, which lead to a unique collection of data.

References

1. Lienard, M., Degauque, P.: Propagation in Wide Tunnels at 2 GHz: A Statistical Analysis. *IEEE Transactions on Vehicular Technology* 47(4) (1998)
2. Strang, T., Meyerzu Hörste, M., Gu, X.: A Railway Collision Avoidance System Exploiting Ad-hoc Inter-vehicle Communications and Galileo. In: Proceedings, 13th World Congress on Intelligent Transportation Systems, London, UK, October 8-12 (2006)
3. <http://www.collision-avoidance.org>
4. Lehner, A., Strang, T., Rico García, C.: A reliable surveillance strategy for an autonomous Rail Collision Avoidance System. In: Proceedings of the 15th ITS World Congress, New York, USA, November 16-20 (2008)
5. Rico García, C., Lehner, A., Robertson, P., Strang, T.: Performance of MAC protocols in beaconing Mobile Ad-hoc Multicast Networks. In: Vinel, A., Bellalta, B., Sacchi, C., Lyakhov, A., Telek, M., Oliver, M. (eds.) MACOM 2010. LNCS, vol. 6235, pp. 263–274. Springer, Heidelberg (2010)
6. <http://www.bayerische-oberlandbahn.de>
7. http://en.wikipedia.org/wiki/Bayerische_Oberlandbahn
8. Lehner, A., Rico García, C., Wige, E., Strang, T.: A Multi-Broadcast Communication System for High Dynamic Vehicular Ad-hoc Networks. In: Proceedings of the First IEEE Vehicular Networking Conference IEEE VNC, Tokio, Japan, October 28-30 (2009)
9. ETSI EN 300 396-2 V1.3.1: Terrestrial Trunked Radio (TETRA); Technical requirements for Direct Mode Operation (DMO); Part 2: Radio aspects (2006-2009)
10. ETSI EN 300 396-3 V1.3.1: Terrestrial Trunked Radio (TETRA); Technical requirements for Direct Mode Operation (DMO); Part 3: Mobile Station to Mobile Station (MS-MS) Air Interface (AI) protocol (2006-2008)
11. Parsons, J.D.: The mobile radio propagation channel, 2nd edn. John Wiley & Sons, Inc., London (2000)
12. Jakes, W.C.: Microwave Mobile Communications. John Wiley & Sons, Inc., New York (1974)
13. Rico García, C., Lehner, A., Strang, T.: Channel Model for Train to Train Communication using the 400 MHz Band. In: Srinivasan, Vikram (Hrsg.) IEEE 67th Vehicular Technology Conference VTC Spring, May 11-14, pp. 3082–3086. IEEE Conference eXpress Publishing, Singapore (2008); ISBN 978-1-4244-1645-5, ISSN 1550-2252

WiMax'ble Pervasive Cloud – Empowering Next Generation Intelligent Railway Infrastructure

Subrahmanyam Venkata Radha Krishna Rao¹ and Vivek Diwanji²

¹ Global Technology Office, Cognizant Technology Solutions, Chennai

² Engineering & Manufacturing Solutions, Cognizant Technology Solutions, Pune

{subrahmanyavrk.rao,vivek.diwanji}@cognizant.com

Abstract. One of the major expectation from any next generation Railway Intelligent Infrastructure will be to reduce track side infrastructure to an extent possible by empowering on-board infrastructure as this enhances operational efficiency while reducing cost. If a broadband technology is made available in a dynamic train environment that could address critical signaling and control system requirements as well as empowers the train passengers - with on-board broadband it will boost business avenues. Pervasive Cloud combines the Power of the Cloud through Pervasive Mobile access . WiMax empowered pervasive access could address critical signalling/control systems requirements while Pervasive cloud itself addressing demands of passengers including but not limited to voice and data services at high speeds, thereby addressing the goals of next generation intelligent infrastructure for railways. This paper while giving a futuristic view of related work in our Labs, intends to argue that WiMax'ble Pervasive Cloud (WiMax empowered Pervasive Cloud) could probably be a right candidate to address the expectations and aspirations of a next generation intelligent infrastructure for Railways.

Keywords: WiMax; Pervasive Cloud; Intelligent Infrastructure; High Speed Trains; Signalling.

1 Introduction and Scope

Large distributed-asset-base industries such as rail, utilities are moving towards building integrated proactive asset management frameworks making use of enhanced capabilities of networking and centralized data processing. Better understanding of asset condition and deterioration, using both historical and operational data, will lead to predictive maintenance and renewal of assets, rather than reactive/fixed cycle of maintenance [1]. Such proactive approach will result into improvements in operational performance through enhanced reliability, availability and fault handling and also result in improved safety of maintenance staff due to reduced manual inspections. For this reason, asset condition monitoring and maintenance practices are increasingly adapting technological advances in smart devices, wireless communications and distributed computing for building intelligent asset monitoring systems called intelligent infrastructure. Such systems are ensuring that maintenance-related information is

made available anywhere, anytime and to authorized users so that actions can become more and more proactive and thus help reduce cost.

In addition, the advancements in data communication are driving train control to move more and more on-board from track-side operations. In rail industry, usually trackside signals require real-time low bandwidth data communications (data rate such as 56kbps) and are expensive to maintain. The blending of communication and information technologies is enabling on-board solutions with more frequent and accurate information such as that about train location resulting into reduced train headways or better train separation to enhance capacity without compromising safety of operations. For example, the safe distance between trains is determined using train's position, its relative speed with other trains and direction of movement. The position of a moving train in a rail network is continuously relayed by wireless signals to other trains. By using this continuous digital signal it is possible to reduce inter-train intervals, and increase traffic capacity without huge infrastructure investments [2]. Such communications are increasingly data based, the role of analogue radio being reduced to local use [3].

Apart from this, there is a growing demand for on-board data connectivity which indeed makes a great business sense for the rail operator as well as empowering a passenger on-board with a spectrum of on-demand data as well as voice services.

In a nut-shell, the major challenges encountered by railway industry include (a) Providing Intelligent and Integrated Operations Management (b) Check integrity of a running Train at any given time (c) Ensuring the harmonic operation between various Railway Transportation and business subsystems (d) Real Time and machine critical Supervisory and Control System ensuring train safety (e) Empowerment of Customer i.e., Passenger Information Service - Travel Guidance, train information and station information, Realtime freight information for merchants and passengers (f) Disaster Prevention and control - GPS/GIS (g) Interoperate with other modes of transport such as air, bus, ship etc. towards facilitating better service to the passenger etc.

If a broadband technology is made available in a dynamic train environment that could address critical signaling and control system requirements as well as empowers the train passengers - with on-board broadband it will boost business avenues, thereby, addressing the goals of the next generation Railway Intelligent infrastructure. Current pursuit within rail industry is towards making use of wireless technologies with an eye on interoperability without compromising safety and security. WiMax is a standards based broadband technology ensuring last mile connectivity [4]. WiMax either as a back haul [5] or while interoperating with Satelite / Fiber / WiFi / other wireless networks has the potential to provide broadband on-board of a train at high speeds [3,8-20].

Pervasive Cloud [6,7] is to access the Power of the Cloud through Pervasive Mobile access. In the railway industry context, WiMax empowered pervasive access could address critical signalling/control systems requirements while Pervasive cloud could address passengers demands including but not limited to on-board voice and data services at high speeds. This paper while giving a futuristic view of related work in our Labs, intends to argue that WiMax'ble Pervasive Cloud (WiMax empowered Pervasive Cloud) could probably a right candidate to address the expectations and aspirations of a next generation intelligent infrastructure for Railways.

2 WiMax'ble Pervasive Cloud

2.1 WiMax

Worldwide Interoperability for Microwave Access (WiMax) is an IEEE 802.16 standards based wireless technology and is promising towards providing high speed last mile broadband connectivity to homes, businesses as well as for mobile wireless networks. WiMax offers larger bandwidth, stronger encryption and improved performance over longer distances in a Non Line of Sight mode as well. WiMax uses Orthogonal Frequency Division Modulation (OFDM) technology, which has a lower power consumption rate. WiMax can be used for various applications such as last-mile broadband connections, hotspots and cellular backhaul and high-speed enterprise connectivity for business etc. It supports broadband services like VoIP or Video. WiMax is essentially a next-generation wireless technology which can deliver ultra-fast Internet access over many miles and fills a critical gap in the end-to-end communications network. Lower equipment costs for a standards based solution and assured interoperability [Figure 1] with other technologies such as 3G/WiFi etc. greatly reduce the risks for operators wishing to roll out WiMax services.

WiMax is a cost-effective technology which could be deployed quickly and efficiently in regions that otherwise would not have broadband access [4].

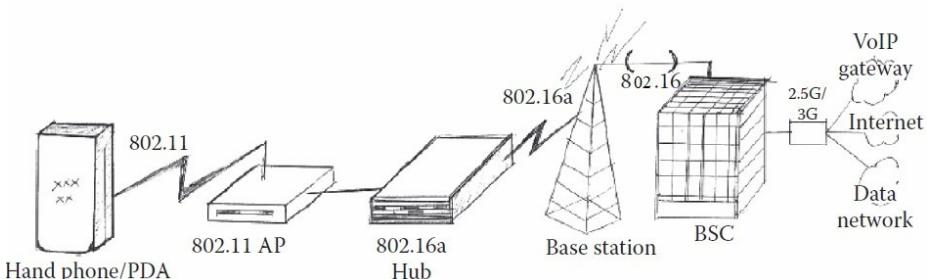


Fig. 1. A Back-haul configuration of WiMax with WiFi and 3G networks [4, 5]

2.2 Pervasive Cloud

Pervasive Cloud [6] is to access the Power of the Cloud through Pervasive Mobile access i.e., in other words Pervasive Cloud is to “Get connected seamlessly – Get your data and computing taken care ubiquitously”. Pervasive Cloud is based on convergence [Figure 2] of emerging technologies such as Cloud and Mobile computing and could be used to provide inexpensive solutions with minimal infrastructure. This leads to a low cost solutions, and has advantages such as Mobility, Faster Response, Zero maintenance, Fault tolerance, Scalability etc [6].



Fig. 2. Architectural Model of Pervasive Cloud

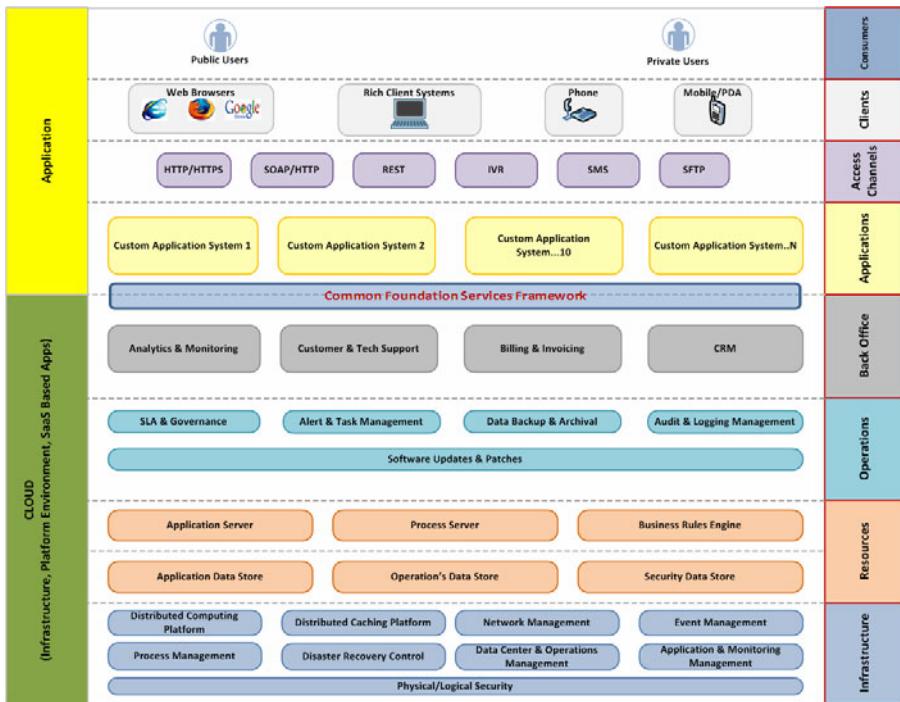


Fig. 3. Pervasive Cloud Platform – Layered View [6]

Figure 3 depicts the proposed solution as a set of collaborative components working at various architectural layers - Consumers, Clients, Access Channels, Applications, BackOffice, Operations, Resources and Infrastructure layer. These

layers are designed in accordance with the principle of enabling increased cohesion within the layers and reduced coupling across the layers. This logical separation of the architecture enables a clear separation of concerns thereby partitions the responsibility of the system more appropriately. This technique also aids in the infrastructure design by mapping these logical layers to the physical tiers. It is also important to note that the layers represented here are more of logical in nature and hence a layer can be present across different physical tiers or more than a layer can be present in a single physical tier. The components provided in this illustration [Figure 3] are just logical in nature that provides the functions provided by these components without providing any details on how these components would be realized. Both the Application specific components and the Cloud specific components are depicted here across the layers. The Cloud specific components are depicted under green section [Figure 3] whereas the Application-specific components are depicted under yellow section [6]. The details of the individual components are described below.

'Consumers' layer shows various actors who will access the Pervasive Cloud for specific purposes that are catered through the various application components.

- ✓ Public Users
- ✓ Private Users

'Access Channels' layer represents the various modes through which consumers can access the Pervasive Cloud. The access to the application will be done using:

- ✓ Internet Explorer, Safari and Firefox web browsers running on top of Windows or Linux- desktop
- ✓ Web Services using SOAP/HTTP and accessing web services through an authentication client device (Mobile-phone/POS-device/Computer with a biometric scanner)
- ✓ With an operator assistance through IVR (Interactive Voice Response)
- ✓ By submitting a request through SMS
- ✓ REST style URLs exposing all publishable public information for consumption of 3rd Party developers

'Applications' layer represent various Custom Applications Systems which has been built using the Cloud Platform. This layer can be consists of sub-layers such as Presentation, Business (Service Layer, Orchestration), Data Access, and Integration Layers for a given Custom Application which can interact with other layers.

'Back Office' layer represents the Back Office Automation of Pervasive Cloud. The back office layer consists of administrative modules such as Analytics and Monitoring, Billing & Invoicing, CRM and Customer & Tech Support.

'Operations' layer is responsible for providing operational related services like SLA monitoring and verification, Alert & Task Management, Audit & Logging Management, Data Backup & Archival Management, and Software and Service pack updates and remote management of the Application infrastructures.

'Resources' layer represents the various sources of resources such as Application Server, Business Rules Engine, Process Server, and various Data Stores such as Application Data Store, Operations Data Store, and Security Data Store etc.

'Infrastructure' layer is responsible for providing Infrastructure related services like Distribute Computing Platform, Distributed Caching Service, Event Management, Network Management, Various Process Management Services, Disaster Recovery Control, Data Center & Operations Management, Application Monitoring Management, Physical/Logical Security Management etc. of the Infrastructure.

Pervasive Cloud was earlier implemented in Rural Healthcare Scenario [6] as well as an e-Business Scenario [7]. For both these implementations we did use WiFi as well as GPRS connectivity. We believe that WiMax is promising towards providing pervasive wireless access. Having said that, we define WiMax'ble Pervasive Cloud as an implementation of Pervasive cloud using a WiMax network where WiMax could be in form of a Back-Haul or interoperating with existing Satelite / Optical Fiber / WiFi or any other wireless network. That said, a WiMax'ble Pervasive Cloud leverages and inherits the power of WiMax (pervasive access) as well as the power of a Cloud. Having seen the success of [6,7] a Pervasive cloud (through WiFi and GPRS), we believe that WiMax'ble Pervasive Cloud could address the needs of a dynamic Railway environment such as High Speeds as well as the maintenance of machine critical signaling and control systems.

Broadband communication makes possible the advanced control signals [8-18]. A communications management unit (which in-turn contains multiple antennas for multiple standard of technologies) mounted on the roof of a train can interoperate with a train LAN (connecting various wagons of the Rail). Also, the roof mounted train antennas can communicate well with track-side communication systems including WiFi, Satelite, Optical fiber systems, ensuring high quality communication possible even at high speeds of the train [8-20].

3 Intelligent Railway Infrastructure

Intelligent Infrastructure is conceptualized to achieve important technological and economic objectives such as increased capacity and asset utilization, improved reliability and safety, higher customer service levels, better energy efficiency and less emissions, and increased economic viability and profits. Intelligent Infrastructure is expected to bring together an array of technologies in the form of sensors, communications, computing, and intelligent control to address various aspects of rail operations and control such as predictive maintenance, planning and scheduling, and fuel management.

Remote condition monitoring has been around [1] that uses principles of real-time monitoring of systems for control and diagnostics with the help of programmable logic controllers, distributed as well supervisory data acquisition and control systems. The advancements in technologies for sensors, systems and networks is enabling a

more holistic view to model and manage interconnectivity of systems and thus derive more meaningful information out of interdependent events and actions.

We believe that there are two core technology challenges (and hence opportunities) for rail technology i.e. (a) making the applications Intelligent and (b) ensuring the connectivity at dynamic high speeds. Any next generation Intelligent Railway Infrastructure should address these two core challenges while encashing the related opportunities. In Figure 4, we propose functional architecture of intelligent infrastructure. While this functional architecture inherits the architectural features of the Pervasive Cloud (as illustrated in Figure 3), the WiMax empowers the architecture through ‘connectivity’ (as shown in the Figure 4). The Intelligent Infrastructure system comprises of logical blocks in the form of data acquisition, aggregation and processing, data storage, data analytics and decision support, and reporting to various users.

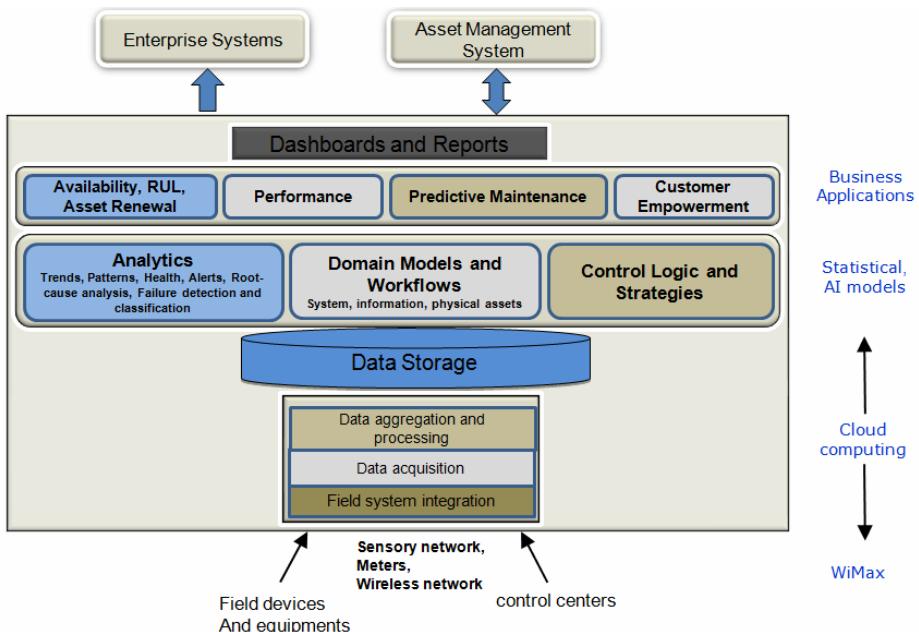


Fig. 4. A high level view of proposed Intelligent Infrastructure

3.1 Example Deployment Scenario – Railway Condition Monitoring

As an illustration we state scenario of deployment of this architecture in railway condition monitoring application. Such system will gather:

- ✓ Monitoring data from fixed assets such as points, signaling systems, track circuits etc. through sensors or wireless devices
- ✓ Reported events from control center
- ✓ Data from on-board systems through sensors or wireless devices

The data collected is converted to a common engineering format as different measurement systems co-exist and processed. The data for events, asset conditions, performance and their maintenance activities, etc. are stored for analytics layer to analyze trends, patterns while monitoring asset conditions, detect and classify failures, perform root cause analysis of different failures and host models that predict the availability of assets for intended operation. Such knowledge-base is very useful to designers as well as domain experts. Continuous improvement to this knowledge-base can be done effectively by learning from experiences captured as part of asset management system. Intelligent Infrastructure linked with asset management systems will create alert messages for technicians and maintenance engineers through web client and mobile devices.

In view of the technology advantages offered by Cloud computing such as on-demand high availability, zero maintenance, fault tolerance, scalability, we have chosen a Cloud environment for our ‘intelligent’ data processing. We are leveraging our Pervasive Cloud initiative [6] towards processing various data pieces to churn out ‘intelligence’. Towards ‘connectivity at dynamic high speeds’, we are banking on WiMax pervasive connectivity as WiMax has proved it’s potential towards connectivity at high speeds [8-18].

4 Discussion

For Railway Industry, Wireless technologies could be of significant value to [16]:

- ✓ Deliver Vital and Non-vital data services
 - a) train control and location information
 - b) Improve passenger safety and security
 - c) On-board media /advertising services
 - d) Remote software upgrade for applications
- ✓ Use of standardized protocols & technologies for Air-Interface
- ✓ Lower capital costs, rapid deployment
- ✓ Connects with multiple bearer services (satellite, fixed wireless, cellular backbone)

Popular Internet applications may not be available at high speeds due to (a) lack of bandwidth (b) poor quality of service and (c) frequent handoffs. These problems could be partially addressed by (a) increasing network bandwidth using smart antenna systems and MIMO technologies (b) improved handoff protocols that prevent connection loss when moving from one base station to another. Fast fading, Doppler shift, train penetration loss and tunnels have an effect on the dynamic wireless train environment. The train penetration loss significantly reduces reception quality for users inside trains due to metalized windows [8, 10-20]. Further, moving mobile stations cause Doppler shift and Doppler spread. Typically, customized WiFi and WiMax can lead to a feasible solution [8-18].

Motivational factor of WiMax include

- ✓ IP transport layer
- ✓ Broadband video on Demand
- ✓ Last mile Broadband
- ✓ Broadcast TV
- ✓ Backhaul for Cellular and WiFi
- ✓ VoIP
- ✓ Lower CAPEX and deployment time

WiMax is already implemented on Trains at Dubai Metro with a coverage for Smartphone, PDA and Laptops [16]. Usage of customized WiMax and Wi-Fi access solutions offered by a dedicated network can improve their ranges, and become a competitive technology [8-18]. There are identified features in the IEEE 802.16 specification such as Mobile to Mobile support, License Exempt Operation, Public Emergency Services Support and 700 MHz Profile, which currently contribute to consider the IEEE 802.16 access technology as the best competitive access technology for the railway domain [8-18]. The WiMax technology has emerged extending the reach of WiFi, and is a very suitable technology to establish radio links, given its potential and high-capacity at a very competitive cost when compared with other alternatives [8-20].

Smart Instrumentation, better Interconnections and added intelligence are enabling opportunity to think and act in newer ways within Rail context, for example:

- ✓ Can on-board decision support systems be developed for driver assistance for energy savings or travel optimization?
- ✓ Can passengers plan their end-end journey using mobile devices, while traveling can they continue to receive real-time updates about best routes and fares?
- ✓ Can historical travels be used to understand customer preferences for travel and provide suggestions accordingly?
- ✓ Can one determine inventory of spare part across different locations based on maintenance history and asset degradation model?

Also, specifically from WiMax perspective

- ✓ On-demand services – music, video streaming, chatting, internet browsing, internet gaming, weather updates are feasible
- ✓ Real-time train and route updates to passengers on board – essentially helping to plan further connecting travel

Table 1. Sample applications for Wimax'ble Pervasive Cloud [17]

Category	Application	Communication requirement	Computing/Cloud Requirement	Business Model
<i>User navigation</i>	Providing information to passenger and goods personnel for decision making before departure	High speed networks and require high bandwidth May not be time-critical	Storage, Analysis and Computation are involved and Cloud makes sense both towards storage and computation	Transactional On-Demand etc.
	Providing information to passenger when aboard			
	Navigation of passengers on station			
<i>E-business</i>	Passenger traffic	High Bandwidth and not time-critical	Computing involved and Cloud makes sense	Transactional and On-Demand etc.
	Freight traffic			
<i>Train control and supervision</i>	Train and traffic control	Low bandwidth but Time critical	Cloud usage is evolving	Not applicable
	Supervision at rail-road crossings			
<i>Resource management</i>	Signaling infrastructure management – of other assets	Low bandwidth but Time critical	Cloud usage is evolving	Not applicable
	Maintenance work order management			

References

1. Ollier, B.D.: Intelligent Infrastructure - The Business Challenge. In: International Conference on Railway Condition Monitoring, Birmingham, UK (November 2006)
2. Shafiullah, G., Gyasi-Agyei, A., Wolfs, P.J.: Survey of Wireless Communications Applications in the Railway industry. In: 2nd International Conference on Wireless Broadband and Ultra-Wideband Communications (AusWireless), Sydney, Australia, August 27-30 (2007)

3. Railway Communications Strategic Directions Project - published by Booz – Allen – Hamilton (2003)
4. Subrahmanyam, G., Rao, V.R.K., Radhamani, G.: WiMax: A Wireless Technology Revolution, pp. 1–300. Auerbach Publications, New York (2006)
5. Smith, C., Meyer, J.: 3G Wireless with WiMax and WiFi, pp. 1–234. McGraw Hill Publishing, New York (2005)
6. Subrahmanyam, G., Rao, V.R.K., Sundararaman, K., Parthasarathi, J.: Dhatri: a Pervasive Cloud Initiative for Primary Healthcare Services. In: IEEE-ICIN 2010, Berlin, October 11–14 (2010)
7. Subrahmanyam, G., Rao, V.R.K., Sundararaman, K., Parthasarathi, J.: Context-Aware Ubiquitous Musafir. In: IEEE-ICOS 2010, Malaysia, December 3–7 (2010)
8. Fokum, D.T., Frost, V.S.: A Survey on Methods for Broadband Internet Access on Trains, pp. 1–40 (August 2008)
9. Rail Safety and Standards Board, Guidance on Digital Wireless Technology for Train Operators. GE/GN8579 (1), 1–94 (June 2008)
10. De Greve, F., De Turck, F., Moerman, I., Demeester, P.: Design of Wireless Mesh Networks for Aggregating Traffic of Fast Moving Users. In: Proceedings of MobiWAC 2006, pp. 35–44 (October 2006)
11. Aguado, M., Jacob, E., Higuero, M.V., Saiz, P.: Broadband communication in the high mobility scenario the WiMax opportunity, pp. 429–442
12. Li-min, J.I.A., Ping, L.I.: The System Architecture of Chinese RITS. Proceedings of the Eastern Asia Society for Transportation Studies 5, 1424–1432 (2005)
13. Salaberria, I., Carballedo, R., Gutierrez, U., Perallos, A.: Wireless Communications Architecture for “Train-to-Earth” Communication in the Field of Railways
14. Verstrepen1, L., Joseph, W., Tanghe, E., Van Ooteghem, J., Lannoo, B., Pickavet, M., Martens, L., Demeester, P.: Making a well-founded choice of the wireless technology for train-to-wayside data services. In: Proceedings of 9th Conference on Telecommunications Internet and Media Techno Economics, CTTE (June 2010)
15. Subrahmanyam, G., Rao, V.R.K.: WiMax- A journey Towards Ubiquitous Computing. In: International Conference on Convergence Information Technology, Korea (2007)
16. Michael, F.: Wireless Technologies in Rail Transport. Invited talk at IET Toronto chapter (May 2010)
17. Wang, F.-Y., Zeng, D., Ning, B., Tang, T., Gao, Z., Yan, F.: Intelligent Railway Systems in China. IEEE Intelligent Systems, 80–83 (2006)
18. WiMax Forum, Deployment of Mobile WiMax Networks by Operators with Existing 2G & 3G Networks (2008)
19. Wajima, T., Bekki, K., Yokosuka, Y.: Leading-edge Solutions for Next-generation Railway Systems, pp. 153–160
20. Gumaste, A., Lodha, A., Mehta, S., Wang, J., Zheng, S.Q.: Light-trains: An Integrated Optical-Wireless Solution for High Bandwidth Applications in High-Speed Metro-Trains

The MIH (Media Independent Handover) Contribution to Mobility Management in a Heterogeneous Railway Communication Context: A IEEE802.11/802.16 Case Study

Marina Aguado, Jasone Astorga, Jon Matias, and Maider Huarte

University of the Basque Country

{marina.aguado, jasone.astorga, jon.matias, maider.huarte}@ehu.es

Abstract. In this paper we propose the use of the IEEE802.21 protocol in the railway context to enhance the highly frequent, repeated and foreseeable handovers between different radio access technologies. This standard specifies IEEE802 media access-independent mechanisms that optimize handovers between heterogeneous IEEE802 systems and between IEEE 802 systems and cellular systems. The global aim is to contribute to develop a seamless layer that provides independence to the application layers from the radio access technology underneath. A case study with two out of the most popular radio access technologies, IEEE802.11 and IEEE 802.16 is undertaken.

Keywords: handover, MIH, railway, IEEE 802.21.

1 Introduction

Traditionally, railway transport is an industry sector with a great demand on telecom services due to the intrinsic mobile nature of the resources involved. Moreover, the railway domain introduces quite specific and challenging requirements to a general wireless communication architecture, system or technology, such as: high mobility, high handover rate, compatibility with legacy or non-conventional applications, stringent quality of service (QoS) indicators and reliability. These legacy applications are related to signalling and train control and command systems. Such signalling systems highly demand communication availability - any communication loss disrupts the signalling system and the trains stop. In fact, the information contained within these systems governs train movement and thus, it must conform to very strict safety rules.

The high mobility and safety constraints are, on the other way round, balanced with some favorable features from the point of view of quality of service (QoS) and telecom resources provisioning. For example, trains are constrained to move within rails and to perform a previously known route. Not only the mobile node's mobility pattern and the corresponding handover sequence is predictable, but also its data traffic profile. Secondly, except in busy yards, the railway network is not a heavily loaded telecom network as the traditional telecom ones. Besides this, normally just one railway operator is responsible for the full infrastructure and equipment deployed all along the line.

Next generation railway communication scenarios are envisaged to consist of a wide variety of radio access technologies deployed all along the railway line; each technology with quite different performance characteristics and most of the times with overlapping coverage support. The major key challenge under this context is how to meet the railway specific high stringent reliability and quality of service (QoS) constraints during the handover processes in this multi technology scenario.

Last years, in the telecom arena, a huge amount of research raised on the integration of heterogeneous wireless technologies. The IEEE 802.21 [1] is the IEEE standardized proposal to enable seamless service continuity among heterogeneous radio access technologies including the IEEE 802 family of standards and 3GPP networks. However the IEEE 802.21 standard does not address implementation details. In fact, in the standard, it is possible to find the guidelines or reference model for each link layer technology but no implementation details. Consequently it is necessary a further analysis that assists in deploying the IEEE 802.21 standard in the specific railway scenario.

To the best of our knowledge this is the first time that this standard is introduced in the railway context. The main contribution of this paper is to identify this railway context as a suitable scenario for a complete MIH (Media Independent Handover) services implementation, to specify the design implementation of this standard for this specific context and to evaluate MIH contribution in a specific case study, a IEEE802.11/IEEE802.16 case study. Our future research challenges are related to handover policies integration in this MIH framework.

This article is structured as follows. Second section introduces background information and related work on handover enhancement techniques. Section 3 describes the IEEE802.21 standard. Section 4 explains how the railway context forces specific design considerations when implementing the MIH standard for this use case. Section 5 provides detail on simulation framework implementation. Finally, Section 6 evaluates handover performance in a specific IEEE 802.16/IEEE802.11 case study and Section 7 concludes the paper.

2 Background and Related Work

A handover process in mobile wireless networks consists basically in the process that a Mobile Node or Station (MS) carries out when moving from one Point of Attachment (PoA) or Base Station (BS) to another point of attachment (PoA) while the mobile node moves across the BSs' cell boundary. An Intra-RAT handover (also known as horizontal handover) takes places when the two points of attachment share the same radio access technology. In the Inter-RAT (Radio Access Technology) handover or vertical handover, the mobile node is equipped with multiple interfaces that support different technologies. Nevertheless, although multiple interfaces are required, just one interface it is used at a time. The handover process takes place between two points of attachment that use different access technologies. The handover may take place between UMTS and GSM or UMTS and WLAN or WLAN and WiMAX.

When mobile broadband wireless technologies migrate from a nomadic scenario to a vehicular usage scenario, the dwelling time within a cell decreases. Consequently,

the need to seamlessly hand over sessions from one cell to another as the user moves across their boundaries increases. Therefore, a heavy burden is placed upon the performance of mobility management solutions. In order to enhance the handover performance and reduce handover latency, a huge amount of different initiatives can be found in literature. One of the main difficult issues is the high number of standardization groups that are involved. The MIH initiative [1] is leading this effort. IEEE802.21 working group is a regulatory group that started its work in March 2004 and was standardized in 30th Jan 2009. This standard main goal is to provide independent mechanisms that enable the optimization of intra-RAT handovers between media types specified by Third Generation (3G) Partnership Project (3GPP), 3G Partnership Project 2 (3GPP2), future LTE initiative and both wired and wireless media in the IEEE802 family of standards, included IEEE 802.16 specification.

Within the inter-RAT handover context, the mobility management problem has to be solved at the link layer and at the network layer. It could be understood that the MIH standard occupies a L2.5 layer in the OSI protocol stack. In this regard, there are a set of IETF network layer initiatives for solving the mobility management problem between different domains (Mobile IP, NEMO, Mobike), and network layer initiatives for solving the mobility management problem inside the same domain (Cellular IP, HAWAI, Mobile-MLPS, HMIP, FastHMIP, MIPSHOP). However, most initiatives, such as MobileIPv6, lack of lower layer information [2,3]. There are some initiatives [2,3] to overcome this problem by proposing cross layer based handover schemes. The problem with these approaches is the lack of standardization. On the other hand, the last IETF group, MIPSHOP [4], has already defined a layer 3 protocol for carrying MIH payload and supporting IEEE 802.21 services at layer 3 between the different access networks.

It is necessary to take into account that Intra-RAT handover are not within the scope of the MIH standard. They are handled by the specific standard: IEEE802.11r, IEEE802.16e, 3GPP, 3GPP2. It is also worthy to remark that handover control, handover policies and other algorithms involved in handover decision making are generally handled by communication system elements that do not fall within the scope of the IEEE802.21 standard. The IEEE802.21 contribution to the handover decision process is that MIH services provide the information about different networks and their services thus enabling more effective handover decision to be made across heterogeneous networks. The IEEE802.21 based Media Independent Handover (MIH) mechanism presents different types of triggers on Layer 2. However, IEEE802.21 does not specify how or when to generate these handover triggers. In this article we name the entity responsible for deciding when to generate these handover triggers as *handover policy engine*. The handover execution itself is also out of the MIH standard scope. IEEE 802.21 can be understood as a handover decision support framework.

To conclude, it is interesting to point out that there cannot be found currently in literature enough MIH performance evaluations, neither in real testbeds nor in simulation framework. However, it is currently been developed an open source MIH function implementation Odtone [5]. Also, in most articles in the literature it is possible to find studies and implementations focused on IEEE 802.21 in the Mobile nodes but no consideration for external servers.

Regarding handover specific strategies for the railway context, most of the identified ones are implementation dependant, mostly patented protected and in any case they don't cover the multi technology scenario but specific horizontal handover strategies [6,7,8].

3 IEEE802.21 Overview

The IEEE 802.21 protocol [1] was designed to aid mobile terminals during the handover processes between heterogeneous networks, including 3GPP, 3GPP2 and the whole IEEE 802 standard family, with the final aim of providing mobile users with an enhanced mobility experience, especially in terms of seamless service continuity. In order to achieve this purpose, the IEEE 802.21 defines the concept of a *Media Independent Handover Function* (MIHF), which provides an abstraction layer to upper level protocols so that they can interact with lower layers without needing to know the specific characteristics of the underlying technologies. Fig. 1 shows how the MIH layer is logically located between the higher layer mobility management protocols and the link layer protocols of each specific network.

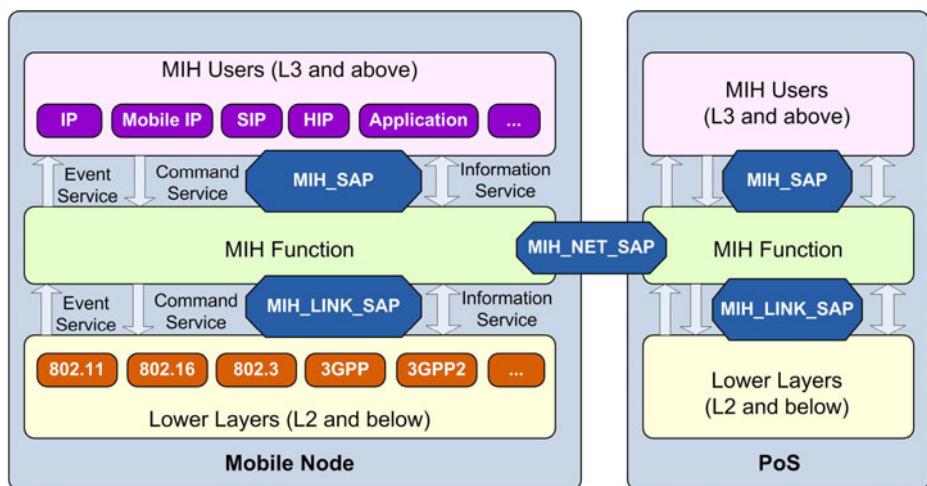


Fig. 1. IEEE 802.21 Standard General Architecture

Therefore, the IEEE 802.21 standard relies for its operation in three key elements: (1) a framework, based on a protocol stack, which handles seamless service continuity during handover between heterogeneous access technologies. This framework must be implemented in all the elements involved in the handover process as it provides the necessary interactions among the involved elements so that the handover decisions can be optimized. (2) A new link layer SAP (Service Access Point) consisting of a technology-independent interface for link layer functions. Then this SAP is mapped to the corresponding technology-specific primitives for each of the technologies

considered within the IEEE 802.21 protocol. (3) A set of functions that enable the communication between the upper layer protocols and the MIHF layer. These functions allow triggering, via the IEEE 802.21 framework, the corresponding local or remote link layer technology-specific primitives considered before.

As it can be observed in Fig. 1, from the MIHF perspective, the upper layer mobility management protocols behave as users of the MIH function, which exploit the services provided by the MIHF making use of a set of defined service primitives grouped in a Service Access Point known as the MIH_SAP. In the same way, the MIH_LINK_SAP has been defined to allow the communication between the MIHF layer and the lower link layer protocols. Additionally, a third type of SAP has been specified, MIH_NET_SAP, for the communication between different MIHF entities.

On the other hand, the IEEE 802.21 standard defines also, as observed in Fig. 2, a set of network entities, including:

- MIH Mobile Node (MIH MN): it refers to a mobile device which implements MIH functionalities and has multiple wireless interfaces.
- MIH Point of Service (MIH PoS): it denotes a network entity with MIH capabilities that exchanges MIH messages with the MN. It must be noted that a given Mobile Node may have more than one PoS at a time, as it may be exchanging messages with multiple network entities.
- MIH non-PoS: it designates a MIH capable network entity that does not exchange MIH messages with a given Mobile Node.
- MIH Point of Attachment (MIH PoA): it refers to the endpoint of a L2 link that has the Mobile Node as the other endpoint.

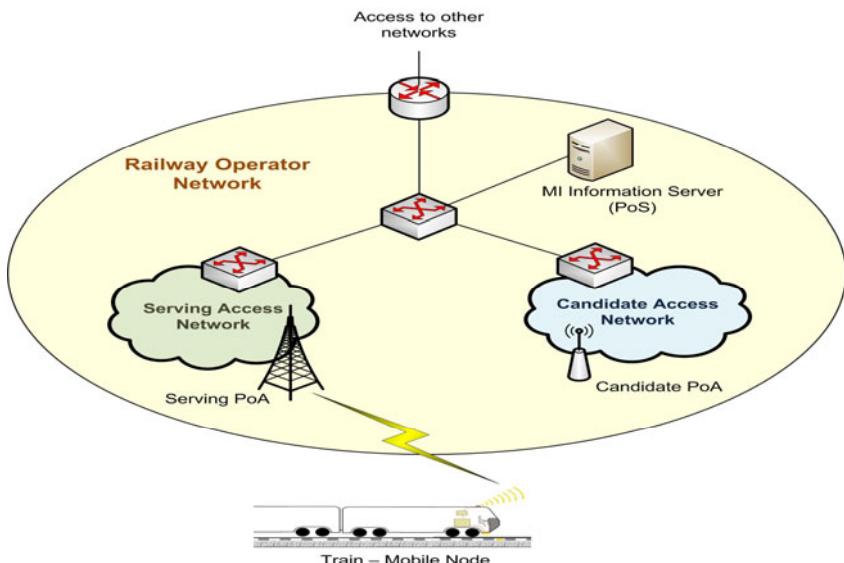


Fig. 2. IEEE 802.21 Network Entities

Finally, the IEEE 802.21 protocol defines three different types of communications with their associated semantics, which are known as the MIH services. The aim of these services is to help MIH users in determining the need for handovers, initiating handovers and deciding the target network by allowing the MIH users to access handover related information, as well as to deliver commands to the link layers or to the network. The specific services defined in the standard are the Media Independent Event Services (MIES), the Media Independent Command Services (MICS) and the Media Independent Information Services (MIIS).

Media Independent Event Services (MIES) are used to report dynamic changes in the link status, link characteristics and link quality, and they are divided in two sub-categories: link events, generated within the link layer and received by the MIHF; and MIH Events, generated by the MIHF and addressed to the MIHF users. The types of reported events can be state changes in the MAC (Medium Access Control) or Physical layers (“link up”, “link down”), changes in the link parameters (“Link Parameters Change”), synchronous events regarding link layer information, such as the native link layer handover methods, or link transmission events, such as information of link layer losses in the ongoing handover.

Media Independent Command Services (MICS) represent commands sent from the higher layers to the lower layers in order to determine the status of links or control and configure the terminal. These commands are especially useful for network initiated handovers and network assisted handovers and they can be further classified in two categories: MIH Commands, sent by the higher layers to the MIHF (“MIH Handover Initiate”, “MIH Handover Prepare”); and Link Commands, originated in the MIHF on behalf of the MIHF user, to configure and control the lower layers.

Media Independent Information Services (MIIS) in turn, allow a MIHF to obtain and distribute network information within a geographical area. This way a given Mobile Node is able to gain knowledge of all the different access networks available in the area of interest from its currently active network interface, facilitating handovers when roaming across these networks. MIIS are based on a new type of information structures called Information Elements (IEs). IEs provide information related to lower layers (neighbor maps, coverage zones and other link parameters) and higher layers (lack of internet connectivity or availability of certain services).

Therefore, following the IEEE 802.21 standard, mobility management protocols should consider two types of information in order to take the handover decision: (1) *Static* information regarding network status, network operators or higher layer service information; and (2), *dynamic* information regarding link status and parameters.

4 Design Considerations

The railway environment presents some specific characteristics that frame the implementation of MIH support in the different nodes involved in railway communication architectures.

The first issue to take into account is that, as already mentioned in the introduction, the access networks deployed in railway environments present special characteristics regarding ownership. They are not usually public access networks, but private

networks frequently owned by the railway operator itself. Therefore, the railway operator has full control over the configuration and administration of the different access nodes deployed along the track, such as for example, IEEE 802.16 Base Stations or IEEE 802.11 Access Points. This fact solves one of the main challenges in IEEE 802.21 implementations that it is the availability of an external server fully aware of the complete network topology in a multi provider scenario. Consequently the previously identified as *static* information provided by the MICS and MIIS is easily achievable.

Another issue is that solving mobility management at layer 3 (L3) is still unsolved in the general mobile network scenario and also in the specific high rate handover railway scenario. Many protocols are under study, which try to deal with it in the most efficient way: Mobile IP, HIP, SCTP, NEMO, etc. However, none of them has achieved widespread deployment today and there is not a clearly prevailing technology, in part due to the complexity of the existing specifications. Therefore, mobility between different IP networks is not a straightforward issue to solve and it frequently involves disruption of ongoing communications.

However, there can be found several successful initiatives to solve mobility management in this scenario at layer 2 (L2) [9,10,11]. Consequently, our approach in this study is based on solving mobility management at L2. Thanks to the implementation of MIH support in all the elements that take part in the handover processes, it is possible to keep communications at L3 and above totally oblivious to the changes that take place at L2 and below.

As a result, the MIH capable mobile nodes involved in our proposed framework can be understood as multi-technology switches or bridges. Each of these devices implements interfaces and the subsequent MAC protocols of each of the different access networks it is connected to (i.e. IEEE 802.3, IEEE 802.11, IEEE 802.16, etc). Additionally, these devices integrate some kind of “intelligence” implemented by the MIH Function in order to deal with the handover between all the different access networks they are connected to more efficiently. More specifically, the MIHF implemented in all MIH-capable nodes allows them to exchange information regarding next handovers as for example information about the coverage areas of all the different access technologies deployed, configuration details regarding the access nodes corresponding to the different technologies, etc. This way, mobile nodes are able to anticipate handovers, knowing the configuration details of the target access technology beforehand and selecting the best instant to perform the change from the currently used access technology to the new one. This decision takes places in each node in what we have named as *handover policy engine*, see Fig. 3.

In this context, dealing with mobility at L2, it does not exist the functional overhead introduced by MIH_SAP, as this SAP’s main function is to provide an interface between the MIHF and upper layer mobility management protocols such as Mobile IP, SIP, etc. In the same way, regarding the MIH Services defined by the standard, we will discard from our implementation MIH Commands and MIH Events, being the former events generated within the MIHF and received by the MIH users, and the latter commands sent by the higher layers to the MIHF. Our railway MIH implementation includes the three different types of MIH services (event, command and information) in the interface between the MIHF and the link layer protocols.

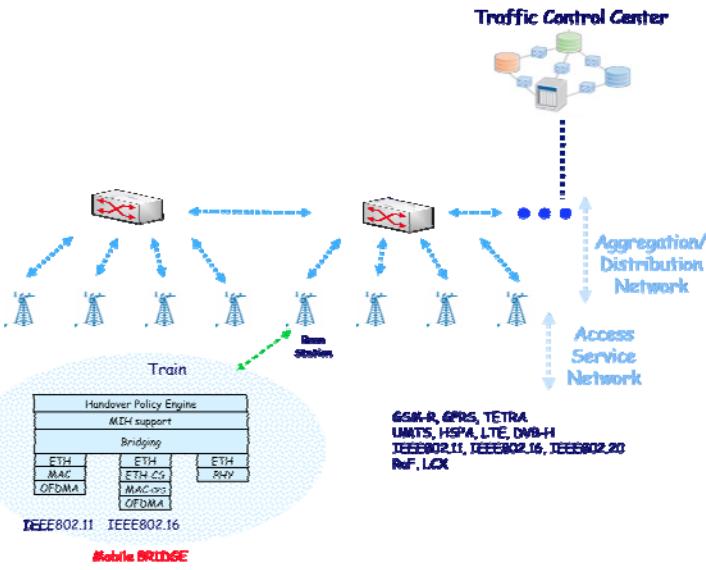


Fig. 3. Railway communication architecture based on solving mobility at L2 and the multi technology proposed mobile bridge

5 Implementation of MIH in the OPNET Simulation Tool

We have implemented the described MIH functionality in OPNET Modeler simulation tool as a new process model included in all the nodes that take part in the MIH-enhanced handover: Mobile Node and PoSs. This process model implements the MIH_LINK_SAP and MIH_Net_SAP in order to allow the communication between the MIHF module and the lower layer MAC protocols, as well as between MIHFs located in different network nodes. Fig. 4 shows the designed communication model among internal modules within the Mobile Node, which is in fact the most critical element of the handover process.

As it is shown in Fig. 4, the developed Mobile Node has interfaces to three IEEE 802 access network technologies: IEEE 802.3, IEEE 802.11 and IEEE 802.16. The implemented MIHF module communicates with the MAC processes corresponding to these technologies by using the MIH_LINK_SAP.

Regarding the developed MIHF process model, it consists of three MIH related processes (MIH_Main, MIH_MN and MIH_PoS) and a handover policy engine. The MIH_Main process is the basic process in charge of invoking the corresponding child processes depending on the received messages and events. Therefore, this process starts the appropriate MIH_MN or MIH_Pos child process depending on the role undertaken by the node it is currently running on. That is, in the Mobile Node, the MIH_Main process will start the MIH_MN child process, while in the MI Information Server the MIH_Main process will start the MIH_Pos child process. Additionally, whenever a handover is imminent, after the reception of a Link_Detection.indication generated by the candidate access network MAC protocol, the MIH_Main process

starts the handover decision process. This process is in fact responsible for determining the optimum instant to perform handover, starting at that moment the establishment of a L2 connection with the candidate access network. In order to take the handover decision, this process takes as inputs the SNR (Signal to Noise) values measured by the MAC layers of the currently serving and candidate access technologies.

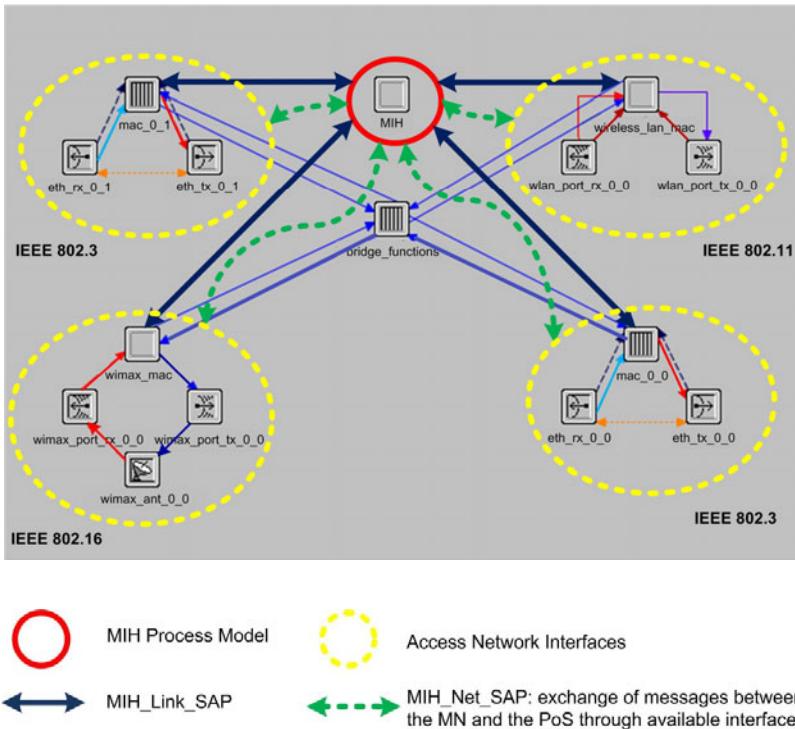


Fig. 4. IEEE 802.21 OPNET Model Implementation

6 Case Study: IEEE802.11/802.16 Handover

As a specific example, we will study the case of a L2 handover between IEEE 802.16 and IEEE 802.11 networks. In the railway environment, these situation could take place, for example, if a train connected to a IEEE 802.16 network enters a tunnel were IEEE 802.16 network coverage is very poor. Therefore, in order to maintain service connectivity the train would stop using the IEEE 802.16 network and connect to the IEEE 802.11 access point deployed inside the tunnel. In the same way, when leaving the tunnel, the train moves out of the coverage of the IEEE 802.11 AP and thus, it re-connects to the IEEE 802.16 network. This situation is represented in Fig. 5. Another possible situations covered under a similar case study are, for example, a possible

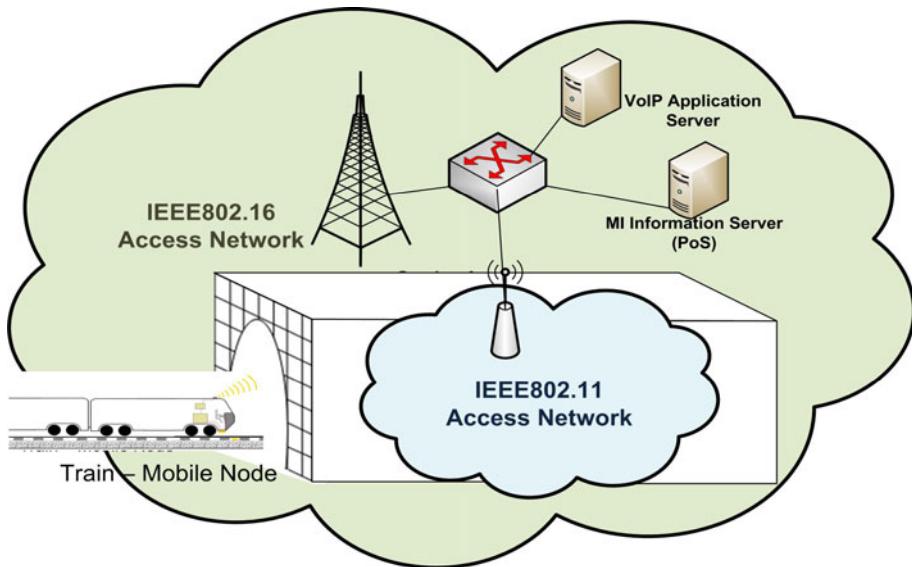


Fig. 5. Train communication alternatives within a tunnel

handover from the IEEE 802.16 wide area access network to the IEEE 802.11 network when a train arrives at a train station or a malicious electromagnetic attack to one of the IEEE 802.16 or IEEE 802.11 access networks forcing the mobile node to migrate to other redundant available access network.

In a normal operation, without MIH support, the handover process starts when the bridging functionalities implemented in the mobile node detected that the link through the IEEE 802.16 interface is down. In such case, the bridge would start broadcasting frames through all its currently active interfaces, which in this case would include the IEEE 802.11 link with the AP located inside the tunnel. As a result, in this case, the handover would not start until the IEEE 802.16 is so degraded that the link is down. Therefore, ongoing communications suffer packet losses and high transmission delays derived from the degradation of the IEEE 802.16 link until it reaches the link down threshold and the communication through the IEEE 802.11 link is initiated. The behavior when the train leaves the tunnel is the same one: handover to the IEEE 802.16 access network is not initiated until the degradation of the IEEE 802.11 link reaches the link down threshold. As before, this fact causes the degradation of the service provided to higher layer applications as a result of the poor IEEE 802.11 link quality.

Thanks to the implementation of MIH functionalities both in the Mobile Node as well as in the access network elements the handover process is enhanced as the bridging functionalities implemented in the Mobile Node can anticipate impending handovers and do not have to wait until the link down condition is reached to change ongoing communications to a better quality link. In fact, thanks to the information regarding coverage areas provided by the MI Services, the Mobile Node can take

advantage of overlapped coverage areas and determine the optimum instant to change from the currently serving access network to the candidate access network.

6.1 MIH-Enhanced Handover Process

Next we describe the designed handover operation that takes advantage of MIH Services. The aim of this MIH-assisted handover is to guarantee service continuity by minimizing packet losses when the Mobile Node is traveling across different access networks.

Fig. 6 (a) shows the message exchange involved in a mobile initiated handover from a IEEE 802.16 network to a IEEE 802.11 network that would represent the situation of the train entering the tunnel, as explained before. The handover process is started by the MIHF located in the Mobile Node, which queries the MI Information Server placed in the railway operator network about surrounding access networks (MIH_GET_Information Request). The answer to this request (MIH_GET_Information Response) includes information regarding all the different access networks that provide connectivity in the geographical area of the train and in its near future locations along the track. Therefore, when the train approaches the tunnel, this response will include information about the IEEE 802.11 network, and thus the train switches on its IEEE 802.11 interface and starts listening for beacons. Additionally, the MIH_GET_Information Response message

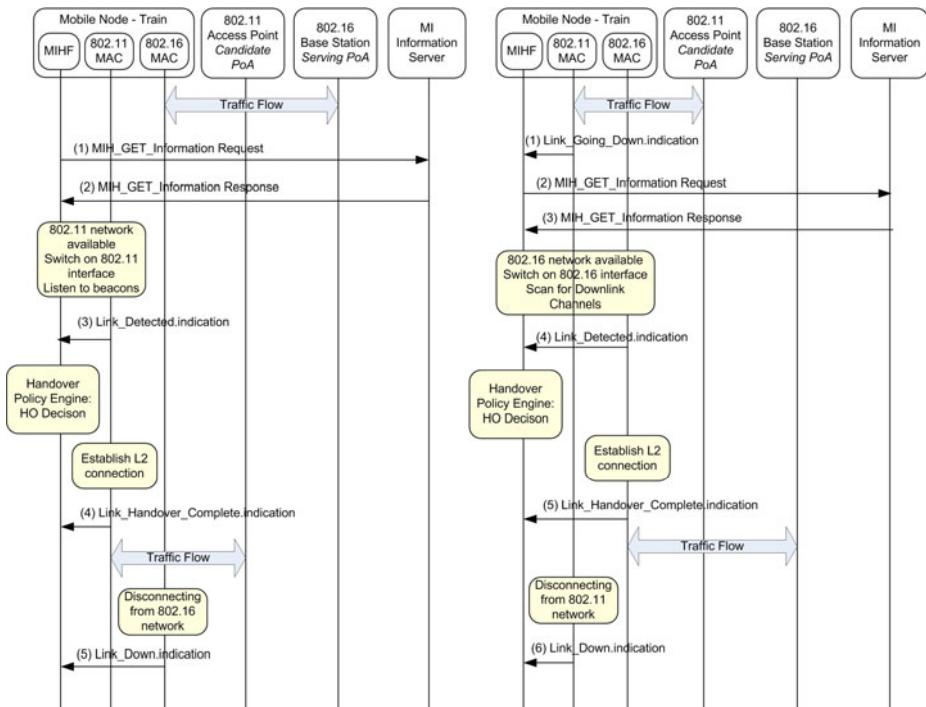


Fig. 6. Message exchange involved in a mobile initiated handover from a IEEE 802.16 network to a IEEE 802.11 network (a) and vice-versa (b)

includes also information regarding the configuration of the target IEEE 802.11 network such as the location of the AP PoA and the transmission channel range. This way, when switching on the 802.11 interface, the Mobile Node can discover the target 802.11 AP more efficiently than if it had to scan all channels.

When a beacon is detected, the 802.11 MAC layer informs the MIHF of the newly detected link through the MIH_LINK_SAP, by generating a Link_Detected.indication event. At the reception of this event the MIHF starts a handover decision process which is not defined in the IEEE 802.21 standard, but which is a crucial element of our implementation. This process measures the quality of the signals received by both IEEE 802.11 and IEEE 802.16 links and decides when handover must be performed. In order to avoid ping-pong effect this process does not make use of instantaneous SNR values, but it integrates the received measures over predefined periods.

Once the decision process implemented inside the MIHF has decided that the Mobile Node should handover to the IEEE 802.11 network, it triggers the IEEE 802.11 connection. When the connection is established the 802.11 MAC layer sends to the MIHF a Link_Handover_Complete.indication and the MIHF proceeds to disconnect from the currently serving IEEE 802.16 network. Once disconnection is performed, the IEEE 802.16 MAC layer informs of this fact by generating a Link_Down event.

Fig. 6(b), in turn, shows the operation of the MIH assisted handover from the IEEE 802.11 network to the IEEE 802.16 network, that is, the process that would take place when the train exited the tunnel. A difference from Fig. 6(a) is that in this case handover is initiated by a Link_Going_Down event generated by the IEEE 802.11 MAC layer when the link condition starts to degrade, and thus, the connection loss is imminent. This event causes the transmission of a MIH_GET_Information Request by the MIHF in order to obtain information about surrounding candidate access networks to which handover. From this point on, the handover procedure follows as in the case represented in Fig. 6(a). Note therefore that MIH_GET_Information Request messages can be spontaneously generated by the MIHF or generated as a result of certain events received by this layer.

6.2 Handover Performance Evaluation Results

In this section we present some simulation results that show how the handover performance is enhanced thanks to the use of the MIH functionality. These results are obtained by simulating the case in which the train exits the tunnel, as explained in previous section. Specifically, we have evaluated the data traffic loss, end to end transmission delay and delay variation occurred during the handover process, both in the case of performing MIH assisted handover, as well as in the case of handover without MIH support.

The results shown in the following graphics correspond to the case of a train which communicates through a IEEE 802.11 Access Point during a minute. In minute 1, the train exits the tunnel and IEEE 802.11 coverage is lost. Consequently, in order to continue its communication service, the train connects to the IEEE 802.16 access network available outside the tunnel. The data traffic used to evaluate the handover performance consists of a VoIP communication between the train and a fixed node

connected to the train operator wired network. VoIP traffic is characterized by owning a constant transmission rate, and thus, when represented in a graphic, it presents a flat data traffic rate where it is easy to identify packet losses.

Fig. 7. shows how thanks to the implementation of MIH functionalities, it is avoided the data loss in the handover between a IEEE 802.11 and a IEEE 802.16 network.

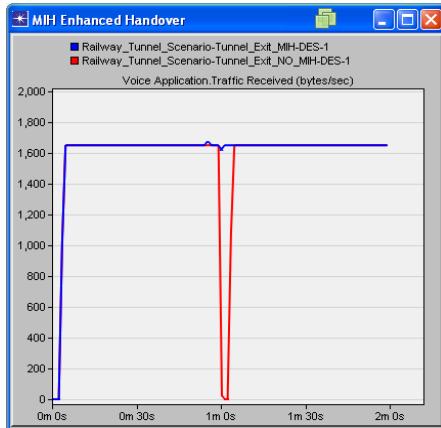


Fig. 7. Traffic Received (bytes/sec)

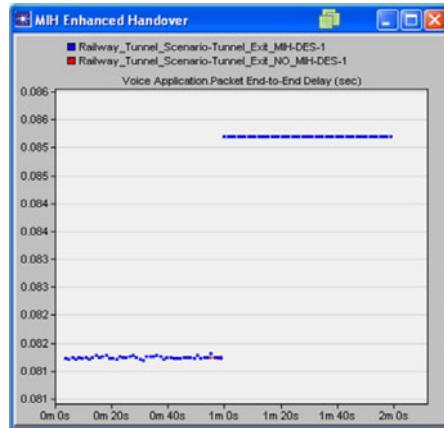


Fig. 8. End-to-end delay (sec)

Fig. 8 in turn, shows how the end to end transmission delay increases significantly when the mobile node starts using the IEEE 802.16 network. This Fig. 8 also shows that the utilization of the MIH functionality does not have a major effect in the end to end data transmission delay, as the results obtained corresponding to MIH assisted handover and to handover without MIH support practically overlap. The increase in the transmission delay is in fact caused by the specific characteristics of each of the access networks used. IEEE 802.16 is a connection oriented technology based on TDD multiplexing and IEEE 802.11 is a contention-based technology, consequently it presents better performance in case of low traffic loads, which is the specific case under study.

7 Conclusions and Further Steps

Our specific implementation of the IEEE802.21 standard for the railway context demonstrates its utility for improving handover performance in a multi technology access network environment. A specific IEEE 802.16 and IEEE 802.11 case study modeled in simulation framework is used to validate our approach. The handover performance evaluation carried out demonstrates that, specifically the data traffic loss, is clearly enhanced when making use of the MIH functionality. Further steps considered within the context of this study are related to analyze the proposed MIH implementation with additional data traffic coming from operational complementary

services such as video transmission and maintenance data transmission for surveillance systems. And another line of research is on heterogeneous handover policies for the railway context been supported by our MIH implementation.

Acknowledgements. This work is partially funded by the ECOTRANS research framework, a *CENIT* Program project from the Spanish Ministerio de Industria. Its main goal is to develop the necessary technologies to be able to offer, in a close future, urban transport public solutions more attractive to passengers and with less environmental impact.

References

1. IEEE 802.21 Standard Local and Metropolitan Area Networks: Media Independent Handover Services (2009)
2. Bernaschi, M., Cacace, F., Iannello, G.: Vertical handoff performance in heterogeneous networks. In: MWN 2004 (2004)
3. Chang, B.-J., Lin, S.-Y.: Mobile IPv6-based efficient vertical handoff approach for heterogeneous wireless networks. Wireless Communications & Mobile Computing 6, 691–709 (2006)
4. MIPSHOP Mobility for IP: Performance, Signaling and Handoff Optimization, <http://tools.ietf.org/wg/mipshop/>
5. Odtone. Open Source Implementation of Open Source Implementation of the IEEE 802.21 Media Independent Handover Framework. Current release: beta version (0.2), <http://helios.av.it.pt/projects/odtone>
6. Patent JP2007235541 by NEC (2007)
7. Patent JP2007194754 by NEC (2007)
8. Patent JP 2007324635 by MITSUBISHI ELECTRIC CORP (2007)
9. Van Quickenborne, F., De Greve, F., De Turck, F., Moerman, I., Demeester, P.: Management of aggregation networks for broadband Internet access in fast moving trains, pp. 273–283 (2005)
10. Aguado, M., et al.: ASME Joint Rail Conference, WiMAX role on CBTC systems (2007)
11. Ritesh Kumar, K., Angolkar, P., Das, D., Ramalingam, R.: SWiFT: A Novel Architecture for Seamless Wireless internet for Fast Trains. In: IEEE 67th Vehicular Technology Conference VTC-2008, Singapore (2008)

Multiple Description Coding and Scalable Video Coding Combined with Multiple Input Multiple Output Techniques: Two Strategies to Enhance Train to Wayside Video Transmissions in Tunnels

Imade Fahd Eddine Fatani¹, Yann Cocheril², Crépin Nsiala²,
Marion Berbineau², François-Xavier Coudoux¹,
Marie Zwingelstein-Colin¹, and Patrick Corlay¹

¹ Université Lille Nord de France, F-59000, UMR 8520, IEMN, Department OAE,
UVHC, F-59313, Valenciennes

{ImadeFahdEddine.fatani,francois-xavier.coudoux}@univ-valenciennes.fr

² Université Lille Nord de France, F-59000, INRETS, LEOST,
F-59666, Villeneuve d'Ascq

{yann.cocheril,crepin.nsiala,marion.berbineau}@inrets.fr

Abstract. Video monitoring applications for underground rely on wireless train-to-wayside communication systems which require high data rate as well as high Quality of Service (QoS) level. In order to satisfy both constraints we propose a combined source and channel coding approach in the context of MIMO (Multiple Input Multiple Output) video transmission. In the present case, MIMO transmission is based on the PHY layer of IEEE 802.11n Wi-Fi standard currently deployed in a railway tunnels. Two different strategies are studied: first, the association between Multiple Description Coding (MDC) and a STBC (Space Time Block Code) MIMO scheme is considered when no channel information is available at transmitter side. In the case when perfect channel information is available at transmitter side (CSIT), a Singular Value Decomposition of the MIMO channel is possible. This transmission scheme is then associated with scalable video coding, which consists here in the separation of the scene into different Regions Of Interest (ROI). The creation of the regions of interest is based on the Flexible Macroblock Ordering (FMO) technique introduced in the new H.264/AVC compression standard. The stream associated to the area with the maximal perceptual relevance is transmitted on the eigen-channel with the higher gain. Consequently, this strategy which provides unequal protection against channel errors, allows guaranteeing better robustness and acceptable reconstructed video quality at the control-centre. The two different strategies of transmission have been evaluated thanks to realistic simulations. Two antenna configurations representative of real cases encountered in railway tunnels are considered. The channel model is generated by using the correlation based Kronecker model obtained by computing the channel matrix with a 3D ray tracing tool. Simulation results show that the two proposed solutions allow enhancing the reconstructed video quality

compared to conventional transmission schemes with no increase of the transmitted power and of the number of radio access points along the infrastructure, even in tunnels in presence of spatial correlation.

Keywords: train-to-way-side communications, video communications, Multiple Input Multiple Output (MIMO) system, Multiple Description Coding (MDC), Scalable Video Coding (SVC), Kronecker model.

1 Introduction

The monitoring of the inside of trains for security purposes is today widely developed for public transports. The video streams are transmitted to a control and security center thanks to a train-to-way-side wireless communications deployed along the lines (CCTV-Closed-Circuit TeleVision). With the development of WLAN (Wireless Local Area Network) IEEE 802.11a/b/g/n/p, railway industry has developed CCTV solutions based on free propagation in tunnels using IEEE 802.11 (a/b/g) modems considered as COTS (Commercial of The Shelf). We can mention Urbalis system for ALSTOM, Airlink system used by SIEMENS designed to migrate easily towards standards evolution (IEEE 802.11n/p...). In this paper we consider a video transmission in an underground tunnel at 2.4 GHz and we propose to enhance video quality at reception thanks to the association of MIMO (Multiple Input Multiple Output) techniques with video coding techniques such as MDC (Multiple Description Coding) then ROI (Region Of Interest) techniques. Two efficient strategies are chosen, whether information about the transmission channel state (CSIT) is available at the transmitting side (MIMO-ROI) or not (MIMO-MDC). The paper is organized as follows. Section II will present briefly some basics in MIMO systems, and the channel model considered in tunnel is explained. MDC and ROI techniques are described in Section III. Section IV will detail the principles of the two association strategies: MIMO-MDC and MIMO-ROI. Monte Carlo simulation results obtained in ideal Rayleigh channel and in realistic tunnel environment are given in Section V. Finally, Section VI concludes this paper and gives perspectives.

2 Basics on MIMO Techniques

2.1 MIMO System Principles

MIMO is known as a promising technology able to offer important channel capacity increase [24] in rich multipath environments without consuming extra bandwidth or transmitting power. This technology is today available in several standards (IEEE 802.11n, WiMAX and LTE) but has not yet been sufficiently tested under realistic propagation conditions and, hence, their integration into real applications can be considered to be still at the beginning [2].

In this work we consider a non-selective MIMO channel perfectly described by the $(N_r \times N_t)$ channel matrix \mathbf{H} which contains the time-variant individual

complex impulse response $h_{ij}(t)$, $1 \leq i \leq N_r$, $1 \leq j \leq N_t$ of each individual SISO (Single Input Single Output) channel. N_t and N_r are the number of antennas at the transmitter and receiver sides, respectively.

The capacity of MIMO channels depends on the number of eigen-channels available, which corresponds to the rank r of the channel matrix \mathbf{H} [3]. The rank is maximal in a channel rich of multipaths like a Rayleigh channel. With spatial correlation or keyhole effect in the channel, the \mathbf{H} matrix will be degenerated [5][22]. This is applicable in an empty tunnel environment considered in this work. The number of scatters is generally low as well as the spread of the angle of arrival of the rays due to the guided effect [16]. Channel correlation brings more challenges to the application of high data rate services in MIMO systems. The authors in [6] give the correlation variations in the MIMO channel in tunnel versus the antennas positions in the tunnel and on the train.

Spatial Multiplexing (SM) [9][10] and Space-time Block Codes (STBC) and particularly the Alamouti scheme [11] are the most well known algorithms for MIMO scheme. The aim of SM is to increase data rate without extra bandwidth consumption or additional transmitted power. STBC aims at increasing the robustness of the system by increasing the global diversity order bounded by $N_r \times N_t$ without need for channel knowledge at the transmitter. MIMO systems offer an interesting tradeoff between data rate and robustness, commonly known as diversity/multiplexing tradeoff. For more details, the interested reader is referred to [25]. In [8], the authors have shown that in the same tunnel configurations than the ones presented here, STBC always outperforms the SM scheme. Thus, we choose to use STBC technique when no CSI is available at the transmitter (see 4.2). When this information is available, it is possible to compute the Singular Value Decomposition (SVD) of the \mathbf{H} matrix to highlight the eigen-channels and their associated gains. In Section IV, we show that the best eigen-channels can be exploited in the association with ROI scheme to improve performances of video transmission.

The simulation chain considered is the same than the one we have considered later in this paper and as close as possible to the IEEE 802.11n PHY layer. The data stream is coded with a 1/2 rate convolutional code with constraint length $K = 7$, and defined by the generator polynomials $g_0 = 171_8$ and $g_1 = 133_8$ followed by a pseudo random interleaver and a 4-QAM modulation.

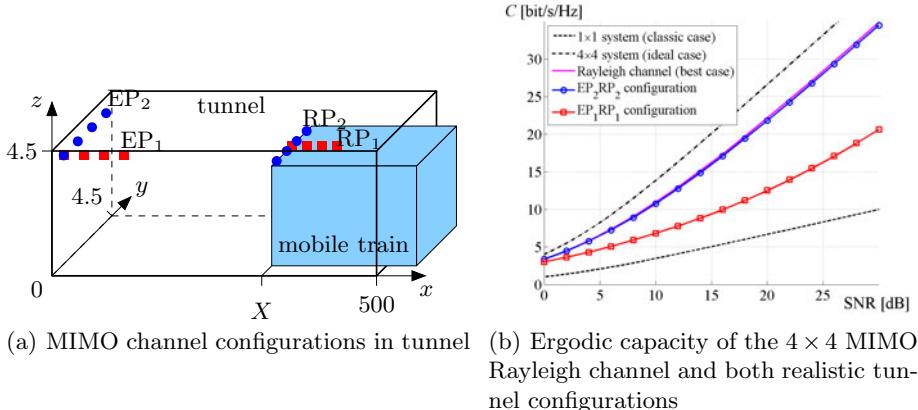
2.2 Modeling MIMO Channel in Tunnels

Reference [2] gives a survey of channel and radio propagation modeling for MIMO wireless systems. In this work, we consider the correlation based analytical model of Kronecker [15]. This model is currently used in MIMO literature to model in a transmission chain the correlation inherent to the channel. A 3D ray tracing based software [4] permits to generate the individual complex impulse responses $h_{ij}(t)$ of the channel matrix \mathbf{H} , from which we are able to set up the model. The use of 3D ray tracing techniques to model the electromagnetic propagation in tunnel has been validated in tunnels by [17][18].

Previous works [7,6] conducted at 2.4 GHz and 5.8 GHz for several tunnel configurations have shown that the correlation degree ρ in the channel is related to the antenna positions. In this paper, two configurations in an empty 1-track tunnel illustrated in Fig. 1(a) are chosen at 2.4 GHz. These two configurations, named EP₁RP₁ (antenna elements parallel to the tunnel longitudinal axis) and EP₂RP₂ (antenna elements perpendicular to the longitudinal tunnel axis), provide high ($\rho = 0.95$ for EP₁RP₁) and low ($\rho = 0.5$ for EP₂RP₂) correlation values respectively, as illustrated in Fig. 1(b), where the capacity versus signal to noise ratio is given for the two configurations and compared with the Rayleigh case. These results highlight the difference between the two realistic tunnel configurations chosen for the performance evaluation.

The 1-track tunnel dimensions are $4.5 \times 4.5 \times 500$ m (width \times height \times length). The train ($3 \times 4 \times 120$ m) under the receiver has not been taken into account in this scenario due to its weak contribution in the computation of the electromagnetic field [12]. The main electric characteristics of the tunnel walls are a relative permittivity equal to $\epsilon_r = 10$, and a conductivity equal to $\sigma = 10^{-2}$ S/m. The number of interactions allowed for ray tracing simulations is fixed to 10 reflections (good trade-off between computing time and accuracy). Transmitters and receivers are composed with two and four elements antennas, respectively. The transmitter is fixed on the tunnel roof and each element is 1 m spaced. The receiver is fixed on the roof of the train and the elements are 65 cm spaced. The main channel characteristics in the tunnel are the followings: RMS delay spread: 2.8 ns, Delay spread: 26 ns and Coherence bandwidth: 79 MHz.

The Kronecker model is based on the computation of the spatial correlation matrix respectively at transmission \mathbf{R}_t and reception \mathbf{R}_r sides, to generate the channel matrix $\tilde{\mathbf{H}}$ with the hypothesis of independence between the correlation at both sides. The total channel correlation matrix is given by equation (11). For each configuration, the correlation matrices at transmitter and receiver sides \mathbf{R}_t



(a) MIMO channel configurations in tunnel (b) Ergodic capacity of the 4×4 MIMO Rayleigh channel and both realistic tunnel configurations

Fig. 1. Realistic tunnel configurations and their propriety in terms of capacity

and \mathbf{R}_r are computed using the channel matrix \mathbf{H} obtained from ray tracing simulations.

$$\tilde{\mathbf{H}} = \frac{1}{\sqrt{\text{Tr}(\mathbf{R}_r)}} \mathbf{R}_r^{\frac{1}{2}} \mathbf{H}_w \left(\mathbf{R}_t^{\frac{1}{2}} \right)^T \quad (1)$$

$$\mathbf{R}_t = E\{\mathbf{H}^T \mathbf{H}^*\} \quad (2)$$

$$\mathbf{R}_r = E\{\mathbf{H} \mathbf{H}^H\} \quad (3)$$

Here, \mathbf{H}_w is a $(N_r \times N_t)$ random matrix with independently identically distributed complex Gaussian entries, $(\cdot)^T$ denotes transpose, $(\cdot)^H$ hermitian transpose, $(\cdot)^*$ conjugation, $\text{Tr}(\cdot)$ matrix trace, $(\cdot)^{\frac{1}{2}}$ matrix square root and $E\{\cdot\}$ the expectation operator.

3 Basics on Multiple Description Coding and Scalable Video Coding

3.1 MDC Technique

Multiple Description Coding (MDC) [11][14] is a video coding scheme where the input video data stream is split into a set of sub-streams with equal importance, also called “descriptions”. Each description can be decoded independently, giving a moderate but acceptable video quality at the receiver. Furthermore, the higher the number of correctly received descriptions, the better the video quality. The cost of MDC is the insertion of a certain amount of redundancy between descriptions. Different MDC techniques have been recently proposed in the literature based on splitting, quantization or transformation [11][26]. In the present work, we adopted the so-called Spatial Splitting (SS) method. It generates two descriptions by separating the odd- and even-line parts of each frame of the video sequence as illustrated in Fig 2. The SS method is fully compatible with

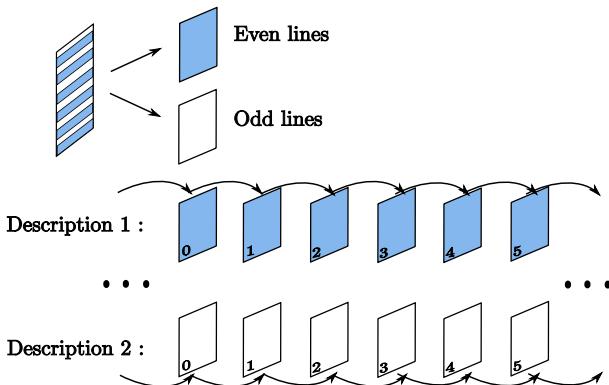


Fig. 2. Description of the spatial splitting method

any coding standard. It is very easy to implement and it is more robust to temporal discontinuities: at a given time, one could expect that at least one half of the entire image should be available to display visual information.

3.2 SVC Technique

Scalable Video Coding (SVC) introduces a hierarchical layered structure between the sub-streams compared to MDC. Indeed, the video signal is split into a Base Layer (BL) and one or several Enhancement Layers (EL). The BL provides coarse video representation in terms of resolution or quality, while the EL can be used progressively to refine the achieved final video representation. In an SVC scheme, the reception of the base layer is crucial. Therefore, BL must be protected against transmission errors by channel coding, or ARQ mechanisms must be used to guarantee good BL reception. SVC has recently been standardized as an extension of the H.264/AVC coding standard. It has been developed in order to provide scalability and flexibility for video transmission over error-prone heterogeneous networks. In particular, in wireless networks, it allows us to cope with reduced bandwidth constraints through unequal error protection (UEP). Hence, the BL can be transmitted with higher protection to guarantee minimal but acceptable video quality.

Traditionally, three different kinds of scalability are provided: SNR, temporal and spatial scalability [21]. A new scalability method, called contextual scalability, consists in splitting the video signal into several streams corresponding to the different regions of an image [23]. Hence, it is possible to distinguish different so-called Regions Of Interest (ROI), and then protect them according to their perceptual relevance. ROI scalability is possible, thanks to the introduction of a new H.264/AVC coding tool called Flexible Macroblock Ordering (FMO). FMO allow to split each image of a video sequence into several separated groups of macroblocks, each of these corresponding to a given spatially localized part of the image as illustrated on Fig.3. In the case of video surveillance, the main ROI (which typically corresponds to an unusual event in the scene, a fight for example) can thus be treated separately from the rest of the scene (the background).



Fig. 3. Principle of ROI scheme based on FMO tool

4 Multiple Description Coding and Scalable Video Coding Associated with MIMO Techniques

4.1 Introduction

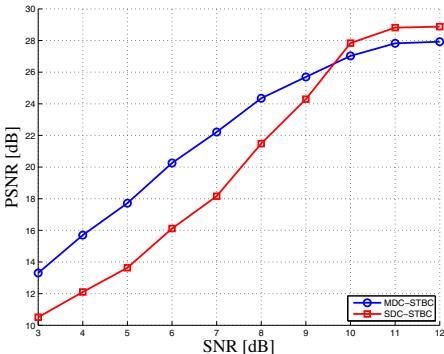
In this section we consider the transmission of a H.264/AVC video stream in the air and we present the two proposed strategies. In both cases, we use the well-known peak signal-to-noise ratio (PSNR) as an objective metric for measuring the quality of received video signal. The PSNR criterion is very popular, due to its simplicity and its ability to provide reasonable assessment of video quality [27]. Hence, it is widely used by the scientific community as a benchmark tool for evaluating the performance of any video transmission system.

4.2 MDC Associated with MIMO

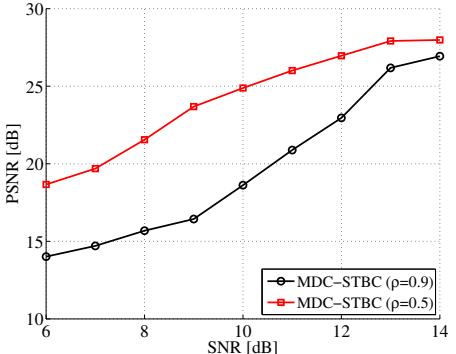
Considering a MIMO system, the probability that all channels between antenna pairs are off simultaneously is very small. Hence, MDC provides an efficient way to enhance error resiliency of the transmitted video stream in order to cope with the MIMO channel impairments. This operation is achieved without using additional error correction techniques; hence the presence of a feedback channel is not mandatory. In this section we consider the association of MDC and STBC in the Alamouti case for a 2×2 MIMO system. We used the well known sequence Garden (352×240 pixels, 100 images) coded with the H.264/AVC standard using only Intra coding with one slice per line. Then, the stream corresponding to the two descriptions is transmitted on the MIMO transmission chain (see Sect. 2.1). In reception, the video sequence is reconstructed. Error concealment based on spatial interpolation is performed in case of packet loss. For several values of SNR, and different channel characteristics, the visual quality of the reconstructed sequence can vary a lot from a simulation to another. Each point of the following figures corresponds therefore to an average computed over 50 simulations.

Fig 4(a) presents the evolution of PSNR as a function of channel SNR in the Rayleigh case for Single Description Coding and MDC. We observe a gain of more than 4 dB for the MDC case compared to the SDC case. The comparison is performed for an equal video throughput for SDC and MDC. Consequently, for high SNR, SDC outperforms MDC.

Fig 4(b) gives the results obtained with correlation as a function of channel SNR. We consider 2 cases: $\rho = 0.5$ and $\rho = 0.9$. The simulation conditions are similar to the one in the Rayleigh case. When the correlation increases, the performance decreases. When the correlation is high, the number of eigen-channel decreases and it is no more possible to receive at least two descriptions on independent channels. Nevertheless this strategy allows us to enhance the video quality on the basis of a well known MIMO algorithm implemented in the IEEE 802.11n norm.



(a) Comparison of MDC and SDC in a Rayleigh channel

(b) MIMO-MDC in a low ($\rho = 0.5$) and a high ($\rho = 0.9$) correlated channels**Fig. 4.** Performance of MIMO-MDC in an ideal Rayleigh channel and both uncorrelated ($\rho = 0.5$) and correlated ($\rho = 0.9$) realistic configurations

4.3 ROI Coding Associated with MIMO

In this part, we present the results obtained with the ROI scheme using the classical Silent QCIF video sequence. Here, the strategy consists in adapting the video content to the transmission channel characteristics. Recently, the association of MIMO and SVC has been considered in the literature [28][19], but the effect of spatial correlation has not been investigated.

As illustrated in Fig. 3, with ROI coding we can separate different areas of different interest in the scene. The data corresponding to the first priority area will be sent on the eigen-channel offering the highest gain and the low priority area on the other eigen-channel. This process is equivalent to a hierarchical protection of the video streams. It allows us to control the degradation due to the channel and offers a better transmission quality with a maximal diversity order for the Region Of Interest (ROI). The streams corresponding to the two ROI are transmitted on the MIMO chain (see Sect. 2.1). We consider QCIF 176x144 sequences encoded using Intra images only. The exact size of each region has been experimentally chosen such that the corresponding output bit rates after Intra-only H.264/AVC encoding are almost equivalent.

At reception, the two streams are decoded with a modified and enhanced version of JM decoder [13]. The reconstruction of the whole video sequence is done applying error concealment in the pixel domain: the parts of the image lost during the transmission are replaced by the same parts of the previous images correctly received and decoded.

Fig. 5(a) presents the evolution of the average PSNR versus E_b/N_0 ratio obtained in a Rayleigh channel in three cases: without FMO tool, for the region of interest (slice group 0 in Fig. 3) and for the stream corresponding to the background (slice group 1 in Fig. 3). The proposed scheme allows us to maintain a high quality level for the ROI. The difference can reach 10 dB compared to the

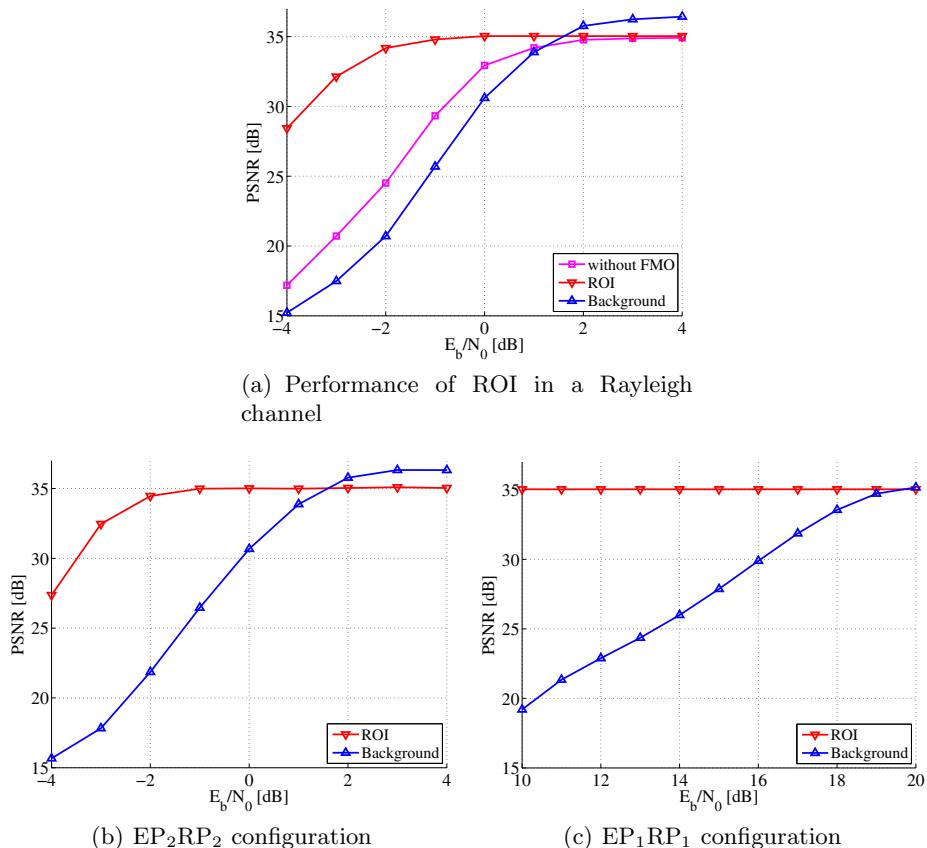


Fig. 5. Performance of MIMO-ROI in a Rayleigh channel and in the both EP₂RP₂ ($\rho = 0.5$) and EP₁RP₁ ($\rho = 0.95$) configurations

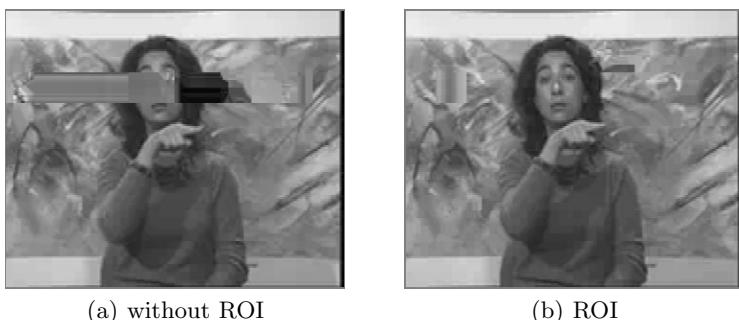


Fig. 6. Visual comparison of decoded frame (image 71) for $E_b/N_0 = 0$ dB in a 4×4 Rayleigh channel with and without ROI approach

classical scheme (without FMO). Obviously, as a consequence, the low priority region (background) experiences distortions. We observe a 4 dB degradation in term of PSNR compared to the classical case. The results obtained in the two tunnel configurations are given Fig. 5. We plot the evolution of the average PSNR of the two regions versus E_b/N_0 ratio in the channels. For the EP₂RP₂ configuration (low correlation scenario), we can observe that the quality of the region of interest remains high for a large scale of E_b/N_0 ratio compared to the quality of the background that is severely distorted after $E_b/N_0 = 2$ dB. For the EP₁RP₁ configuration (high correlated scenario), the ROI quality is maximal at the expense of the deterioration of the background. This is due to the high degree of correlation in the channel that leads to very low gains for the three last eigen-channels. However, the ROI is kept untouched, which constitutes the aim of the proposed strategy. As an example, Fig. 6 present visual comparison of decoded received frames for $E_b/N_0 = 0$ dB in a 4×4 Rayleigh channel with and without ROI approach. We can see that in the classical case, visual distortions can appear in the high priority region (ROI). This area remains free of error when the ROI approach is considered. In the embedded video surveillance application, such degradation in the monitor area will introduce perturbations in the scene analyzed by the security teams.

Finally, the results presented show that the proposed MIMO-ROI scheme permits a selective protection of the video streams. We are able to obtain a perfect protection of the region of interest with a gain of 10 dB compared to the classical scheme in a Rayleigh channel. Furthermore, the ROI based structure is resistant to the channel impairments due to correlation.

5 Conclusion and Perspectives

In this work, we have proposed to enhance the quality of train-to-wayside wireless video transmissions in tunnels based on two strategies of association of MDC or ROI with MIMO transmissions considered for CCTV applications in metros.

First of all, we described briefly the MIMO principles and the two 1-track tunnel configurations considered. The generation of the corresponding Kronecker models is explained. Then, we detailed the MDC and SVC techniques.

Finally, the two proposed transmission schemes are introduced. It is important to notice here that the use of each scheme rely on different transmission philosophy in terms of complexity and reachable performances. In the MDC case, the aim is to receive at least one correct description of the video sequence. While, in the ROI approach, the target is to guarantee the optimal quality of the high priority region of the video sequence.

MDC and ROI techniques divide the original video sequence into several sub-streams in order to exploit the MIMO sub-channels. In the MDC case, each sub-stream corresponds to a description with equal weight. Thus, there is no need to implement hierarchical protection. The MDC is more adapted than the ROI scalability when no feedback link is available. When a return link is available on the transmission system, the ROI scalability approach will take advantage of

the different gains offered by the MIMO eigen sub-channels. Thus, this approach offers high protection of the priority region with no additional coding. The MDC associated with MIMO-SVD can be also envisaged. In this case, one of the descriptions is sent on the best sub-channel. Then, the quasi-perfect reception of one description is guaranteed, even in correlated channel.

In both cases, we have shown in this paper that the association of MDC or ROI respectively with Alamouti and SVD MIMO scheme depending on the availability of CSIT, offers two strategies to enhance clearly the quality of the received video streams in the case of CCTV applications for metros in tunnels even in high correlated MIMO channels. In order to avoid any space-time error propagation in the video sequence, the compression is performed using only intra prediction, compliant with the H.264/AVC standard.

Future works will focus on real hardware implementation of the proposed schemes for real demonstration in tunnels. The system will be enhanced by considering ROI with dynamic segmentation based on the automatic detection of dangerous situations in the video scenes. Precoding techniques are also good candidate and particularly QoS precoder [20] in the case of ROI-MIMO scheme.

References

1. Alamouti, S.: A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Commun.* 16(8), 1451–1458 (1998)
2. Almers, P., Bonek, E., Burr, A., et al.: Survey of channel and radio propagation models for wireless mimo systems. *EURASIP Journal on Wireless Communications and Networking*, 19 (2007)
3. Berezansky, Y., Sheftel, Z., Us, G.: Functional analysis, vol. 1. Birkhauser Verlag, Basel (1996)
4. Chartois, Y., Pousset, Y., Vauzelle, R.: A siso and mimo radio channel characterization with 3d ray tracing propagation model in urban environment. In: Proceedings of ECPS 2005. IEEE, Brest (2005)
5. Chizhik, D., Foschini, G., Valenzuela, R.: Capacity of multi element transmit and received antennas: Correlation and keyholes. *Electron. Lett.*, 1099–1100 (2000)
6. Cocheril, Y., Berbineau, M., Combeau, P., Pousset, Y.: On the importance of the mimo channel correlation in underground railway tunnels. *Journal of Communications* 4(4), 224–231 (2009)
7. Cocheril, Y., Combeau, P., Berbineau, M., Pousset, Y.: MIMO channel propagation characteristics in tunnels. In: Proceedings of ITST 2007, pp. 405–410. IEEE, Sophia-Antipolis (2007)
8. Cocheril, Y., Langlais, C., Berbineau, M., Moniak, G.: Advantages of Simple MIMO Schemes for Robust or High Data Rate Transmission Systems in Underground Tunnels. In: 2008 IEEE 68th Vehicular Technology Conference, pp. 1–5. IEEE, Los Alamitos (2008), <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4656914>
9. Foschini, G.: Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal* 1(2), 41–59 (1996)
10. Golden, G.D., Foschini, C.J., Valenzuela, R.A., Wolnianski, P.W.: Detection algorithm and initial laboratory results using v-blast space-time communication architecture. *Electronics Letters* 35(1), 14–15 (1999)

11. Goyal, V.K.: Multiple description coding: compression meets the network. *IEEE Signal Processing Magazine* 18(5), 74–93 (2001)
12. Hairoud, S., Combeau, P., Cailbault, J.F., Pousset, Y., Cocheril, Y., Berbineau, M.: Acceleration method for a radio propagation simulator based on 3d ray tracing to predict the performances of mimo channels in dynamical railway environment (November 2010)
13. Joint Video Team Reference Software,
<http://iphome.hhi.de/suehring/tm1/download>
14. Kim, J., Mersereau, R., Altunbasak, Y.: Distributed video streaming using multiple description coding and unequal error protection. *IEEE Transactions on Image Processing* 14(7), 849–861 (2005)
15. Kermoal, J., Schumacher, L., Pedersen, K., Mogensen, P., Frederiksen, F.: A stochastic MIMO radio channel model with experimental validation. *IEEE J. Sel. Areas Commun.* 20(6), 1211–1226 (2002)
16. Lienard, M., Degauque, P., Molina-Garcia-Pardo, J.M., Maria, J.: Wave propagation in tunnels in a MIMO context – a theoretical and experimental study. *Comptes Rendus Physique* 7(7), 726–734 (2006)
17. Mariage, P., Lienard, M., Degauque, P.: Theoretical and experimental approach of the propagation of high frequency waves in road tunnels. *IEEE Transactions on Antennas and Propagation* 42(1), 75–81 (1994)
18. Masson, E., Combeau, P., Cocheril, Y., Berbineau, M., Aveneau, L., Vauzelle, R.: Radio wave propagation in arch-shaped tunnels: Measurements and simulations by asymptotic methods. *Comptes Rendus Physique* 11(1), 44–53 (2010)
19. Park, J., Oh, T., Lee, S., Bovik, A.C.: Optimal power allocation for minimizing visual distortion over MIMO communication systems. In: *IEEE ICIP*, pp. 1833–1836 (November 2009)
20. Sampath, H., Stoica, P., Paulraj, A.: Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion. *IEEE Trans. Commun.* 49(12), 2198–2206 (2001)
21. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Transactions on Circuits and Systems for Video Technology* 17(9), 1103–1120 (2007)
22. Shiu, D., Foschini, G., Gans, M., Kahn, J.: Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.* 48(3), 502–513 (2000)
23. Bae, T.M., Thang, T.C., Kim, D.Y., Ro, Y.M., Kang, J.W., Kim, J.G.: Multiple Region-of-Interest Support in Scalable Video Coding. *ETRI Journal* 28(2), 239–242 (2006)
24. Telatar, I.: Capacity of multi-antenna gaussian channels. *European Transactions on telecommunications* 10(6), 585–595 (1999)
25. Tse, D., Viswanath, P., Zheng, L.: Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Transactions on Information Theory* 49(5), 1073–1096 (2003)
26. Wang, Y., Reibman, A., Lin, S.: Multiple Description Coding for Video Delivery. *Proceedings of the IEEE* 93(1), 57–70 (2005)
27. Wu, H.R., Rao, K.R.: *Digital Video Image Quality and Perceptual Coding*. CRC Press, Boca Raton (2006)
28. Yang, D., Nasruminallah, Yang, L.L., Hanzo, L.: SVD-aided unequal-protection spatial multiplexing for wireless video telephony. In: *IEEE 69th Vehicular Technology Conference Spring* (April 2009)

VANET Architectures and Protocol Stacks: A Survey

Sajjad Akbar Mohammad, Asim Rasheed, and Amir Qayyum

Center of Research in Networks and Telecom (CoReNeT),

Mohammad Ali Jinnah University (MAJU),

Islamabad, Pakistan

sajjad_ahr@yahoo.com, asim@corenet.org.pk, qayyum@gmail.com

Abstract. Intelligent transportation systems (ITS) provide a set of standards for vehicular communications. The main focus of research activities, within ITS, has been on development of safety, traffic efficiency and infotainment related applications. Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications are the main research goals of ITS. This paper reviews some popular architectures of VANETs (Vehicular Ad hoc NETworks) i.e., WAVE by IEEE, CALM by ISO, C2CNet by C2C consortium / GeoNet. It also includes some recent research regarding these standards, specially focusing on Network and MAC layer issues. This paper also discusses safety related application protocols, i.e. WSMP by WAVE, CALM FAST by ISO and C2CNet by C2C consortium. Various recommendations regarding the above protocol stacks are presented. The recommendations are based on different parameters like flexibility, implementation etc.

Keywords: VANET, ITS, WAVE, CALM, C2C-CC.

1 Introduction

Intelligent transportation systems (ITS) address the critical problem of traffic safety. Bodies (like IEEE, CALM and C2C-CC) under ITS made tremendous efforts for achieving this goal by making the road traffic management efficient through help of different applications and protocols. The ratio of road accidents can be reduced by using proper traffic management applications. ITS is working on such traffic management solutions to make our vehicular systems safe and better. ITS also discusses the security issues related to these safety applications. Organizations like, International Standard Organization (ISO), Institute of Electrical and Electronic Engineers (IEEE) and Car-to-Car Communication Consortium / GeoNet are working on ITS architecture proposals.

IEEE introduced a complete protocol stack of 1609 protocol family and named it ‘WAVE’ (Wireless Access in Vehicular Environment). Standard is divided in different sub standards to ensure a modular handling of the diverse issues at different layers. It supports dedicated short range communications (DSRC). WAVE enlists two modes of communication:

- i. Safety applications (Non-IP), and
- ii. Non-safety applications based on IPv6.

The approved frequency band is 5.9 GHz (in Europe 5 GHz).

ISO's proposal for VANETs is CALM, a Continuous Air Interface for Long to Medium range. ISO CALM was designed to address the issues of continuous communication among vehicles and between vehicles & roadside infrastructure. The concept of CALM is based on heterogeneous cooperative communication framework to provide continuous communication to user transport. Different implementation projects like COOPERS and SAFESPOT included the main concept of CALM in their work. CALM proposes the use of all available interfaces as opposed to the use of single 802.11p proposed by WAVE.

The Car-to-Car Consortium (C2C-CC) [1] is backed by European car industry. They are working on Car-to-Car and Car to Infrastructure communication, specifically for safety applications. They proposed C2CNet architecture for safety applications. C2CNet is further incorporated by GeoNet in their comprehensive architecture under project COMeSafety. C2CNet is based on geographical routing, in addition to uni-cast and broadcast design proposed by other two architectures. C2C-CC deals with safety as well as non-safety applications. Non-safety applications use IPv6 on wireless multi-hop links to communicate with OBUs (on Board Units), RSUs (Road Side Units) and can access internet for other applications.

IETF is another standard body working on mobile networks and introduced the concept of NEMO (Network Mobility) in MANET. Different projects were started by European countries to realize ITS's proposals. Some of these are NOW, COMeSafety, CVIS, SAFESPOT, COOPERS, GST, GeoNet, FleetNet, GrooveSim, CARLINK, CarTalk2000 etc.

The rest of the paper is structured as follows. Section 2 reviews the VANETs protocol stacks in detail. In section 3, a comparative analysis of VANETs architectures is presented. Section 4 concludes the paper and enlists the goals which may be pursued in future.

2 VANETs Protocol Stacks

This section discusses the VANETs protocol stacks in detail. The discussion focuses on Network, MAC and PHY layers for WAVE, CALM and C2CNet. The approved frequency band is 5.9 GHz for all standards [2]. It was initially approved by U.S Federal Communications Commission (FCC) under Dynamic Short Range Communication (DSRC) concept. The spectrum is divided into six service channels (SCH) and one control channel (CCH) with equal bandwidth of 10 MHz each. For emergency messages (originated by safety related applications) and control messages, CCH is used. SCH is used for other applications' packets. The entire spectrum is divided into time slots of 50 ms. If the CCH channel is active, all nodes are bound to stop their communication during CCH time frame to receive and transmit emergency messages on CCH channel.

2.1 WAVE by IEEE

IEEE introduced a complete protocol stack of 1609 protocol family and named it as WAVE (Wireless Access in Vehicular Environment). There are six sub-standards under 1609 family named as IEEE 1609.1,2,3,4,5,6. Each one handles different issues at different layers. Fig. 1 provides an insight into the six sub-standards and their relationship with respect to the tasks at the various OSI layers [3].

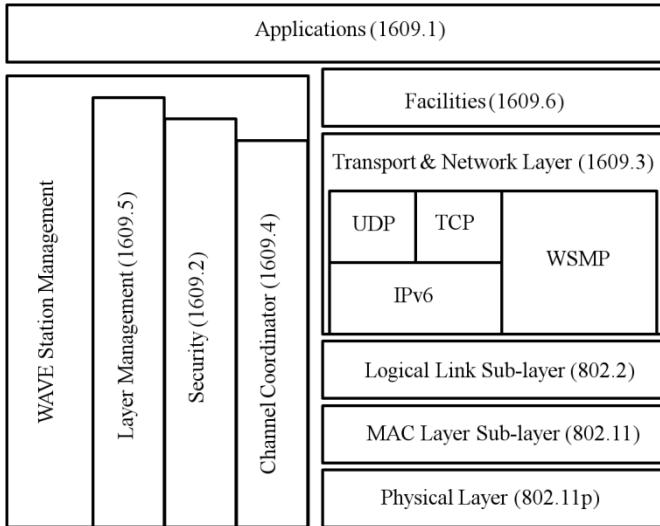


Fig. 1. WAVE Architecture

The types of applications are divided into two main categories as defined by IEEE 802.11p [4]

- i. Safety
- ii. Infotainment.

IEEE 1609.1 details the management activities required for the proper operation of the applications. 1609.2 describes the considerations to be taken into account for communication security. For Transport and Network Layer handling of traffic safety related applications, 1609.3 provides a dedicated single protocol, named as WSMP (Wave Short Messages Protocol). 1609.4 defines the coordination between the multiple channels of the spectrum. 1609.5 deals with Layer Management while 1609.6 offers an additional middle layer between transport and application layer, for handling of additional facilities at the Applications Layer.

IEEE 802.11p details the MAC Layer operation of the WAVE architecture [5]. Vehicular nodes in VANETs can move very fast, leading to fast topological changes. WAVE uses two available services sets [6] for network topology handling. WAVE Basic Service Set (WBSS) is defined for communication between RSUs and OBUS and closely resembles the 802.11a specifications for communication of nodes with the

Access Points (AP). After listening to a beacon message, any new user can join WBSS without authentication process. The second service set is called WAVE independent basic service set (WIBSS). This service set supports the communication between two nodes in a mesh network i.e. V2V communication without the involvement of an RSU.

IEEE WAVE allows only one option at the MAC layer, i.e. 802.11p. Though this restricts the degree of freedom of research activities, open source simulators, like NCTUNs, often provide extended protocol stack support, which assist researchers to use other options at the MAC layer as well. However, here it must be noted that 802.11p is based on a time tested standard (802.11a) which has proven its suitability for short range communications.

2.1.1 Literature Review

Many researchers have investigated the various aspects of WAVE. Some have also done performance evaluation at lower layers and presented their modification proposals. A few of them will be discussed here.

As VANETs mainly focus on traffic safety applications, W. Yi, et al [6] proposed to reduce back-off window for such applications. A priority procedure is discussed in [6] which are based on the modification of back-off windows. For analysis and performance evaluation, different back-off values were assigned to different types of applications and node densities. It was concluded that back-off values and node densities significantly affect throughput. For applications that require quick dissemination and response (like traffic safety related applications), smaller back-off values lead to increased throughput.

M. Todd et al. [7] considered average throughput, average delay and packet loss ratio as performance evaluation parameters for their research. They proved that the speed is insignificant factor, whereas vehicles density is significant factor in vehicular communications considering above parameters. By changing vehicle speed, no significant change occurs with respect to average throughput while average delay converges and packet loss is increased.

E. Stephan et al. [8] concentrated on the collision probabilities in V2V communication. A simple simulation was conducted to prove that in dense and high load scenarios, throughput per node decreases, whereas delay and collision probability increases. Due to the increase in collision probability, performance of safety applications using CCH can seriously suffer. Messages for CCH queue up further during the SCH. Authors suggested a re-evaluation mechanism for messages to provide support to high priority messages.

L. C. Kang et al. [9] introduced Telematics technology with WAVE DSRC by using mechanisms like CAS and EAS. Authors concluded that the Telematics technology is helpful for safety and efficiency scenarios.

2.2 ISO CALM

ISO CALM was designed to provide continuous V2V, V2I and V2O (where O stands for Other Interfaces) communication. The concept of CALM [2] is based on heterogeneous cooperative communication framework. Different projects like COOPERS and SAFESPOT included the main concept of CALM in their work. Even

C2C-CC incorporated the CALM idea for use of multiple-interfaces for vehicular communication networks. The ISO working group TC204 WG16 is working on seamless connectivity issues. CALM standard is also principally based on 5.9 GHz DSRC concept. This standard has been adopted in many countries including Australia, some countries in Asia and South America. For short and medium distances, CALM recommends Infrared. Similarly, for long distances, it prefers the use of GSM, UMTS or any other technology available at PHY layer. Use of satellite links, especially as the backbone, is costly; however, CALM has kept the option open.

A management entity (CCME) is defined in CALM [2] to provide flexibility and adaptability features. Basically CCME consist of three components which are as follows:

i. CALM Interface Manager

CALM interface manager monitors and stores the status of each communication interface (CI), along with its channel quality which aids in making decisions.

ii. CALM Network Manager

The process of handover to alternate media is managed by CALM Network Manager.

iii. CALM/ Application Manager

CALM Application Manager ensures application transmission requirements. It interacts with interfaces to get information about most suitable medium and instructs CALM Network Manager to establish the connection. CALM networking allows IPV6 nodes to move from one IP subnet to another. IETF NEMO protocol is concerned with managing the mobility of a network.

Fig. 2 shows the inter and intra layer entities and their relationships in CALM [10].

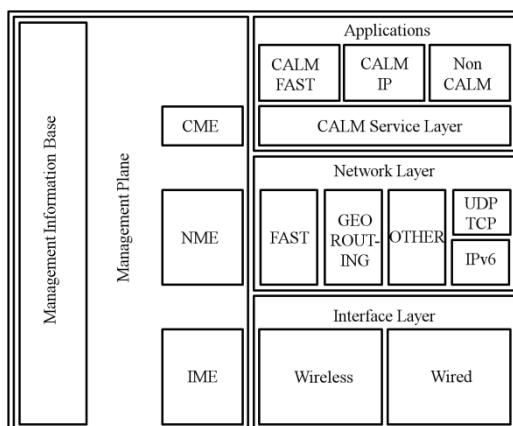


Fig. 2. CALM Architecture

ISO CALM is apparently very flexible in design, with less than 1ms link setup time. It is popular in its design as it provides combination of different technologies for communication as well as it has space in its design to incorporate future technologies. However, researchers are still facing a lot of confusions for its implementation. As CALM is a combination of different technologies so its implementation and provision of different interfaces is a tough job. There is still no support in any simulator for complete CALM protocol stack.

2.2.1 Literature Review

B. Olivia et al. [2] presented CALMnet; a comprehensive network centric simulation environment for CALM based vehicular systems. The authors used OPNET modeler simulation tool, which addressed and identified different problems in cooperative vehicular networks. CALMnet is a simulation tool that was initially implemented to evaluate the performance of lower layers. The simulations were run on Microsoft Windows Server 2003 machine with VMWARE virtual machine running Windows 2000 at 2GHz. It was shown in [2] that the number of simulated nodes has a major affect on the simulation run times and amount of memory required. In future, the authors plan to use IPV6 for their simulations.

2.3 Car to Car Consortium (C2C-CC)

Car 2 Car Communication Consortium (C2C-CC) aims to establish an open European industry standard, focused on development of active safety applications. The C2C-CC is backed by European automobile industry. C2C-CC is designing C2CNet protocol that differs from IP. The protocol is destined to be used for both safety and non safety applications. Also, a single protocol is being considered for use with both the infrastructure based and infrastructure less communication modes. For routing, position based algorithms are used. A dedicated bandwidth of 30 MHz will be available for safety applications in the 5.9 GHz band. In C2C-CC, Network Layer (C2CNet) provides multi-hop communication based on geographical addressing and routing. Network layer also provides beaconing for vehicle movement. Beaconing, location service and forwarding of data are main components of geographical routing. The features that will be provided by C2C-CC are as follows:

- i. Fast data transmission for V2V and V2I communications.
- ii. Support for the transmission of different types of messages including safety messages and infotainment.
- iii. Different short range wireless LAN technologies including IEEE 802.11p, traditional wireless LAN technologies i.e. IEEE 802.11a/b/g/n and other radio technologies e.g. GPRS or UMTS

Fig. 3 shows the C2C-CC architecture [11].

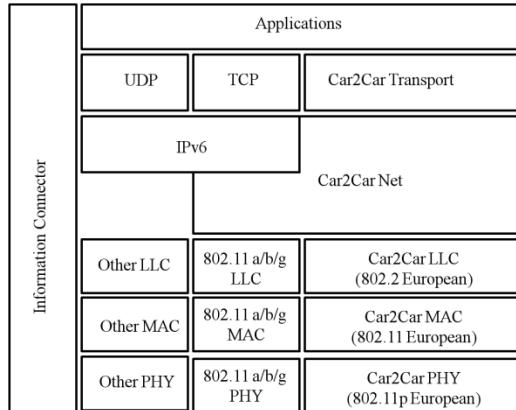


Fig. 3. C2C-CC Architecture

Unlike WAVE and CALM, Safety applications are not restricted to use C2C-CC transport and network layer. Similarly, non-safety applications can use traditional TCP/IP to access wireless multi-hop communications for communication with OBUs, RSUs and can also use C2CNet below IP stack.

2.3.1 Literature Review

T. Manabu et al. [12] referred to project of GeoNet with the aim of combining IPV6 with C2CNet. They focused on V2V and V2I communications. While testing for indoor environment, their test results show that IPV6 over C2CNet does not have too much delay and feasible for vehicle communication. They proposed their future work to evaluate IPV6 over C2CNet in out-door environment.

B. Roberto et al. [13] proposed a solution for application of the NEMO basic support protocol with geographic routing in VANET. They used mobile NEMO router with infrastructure through mutli-hop communication. They showed that mobility does not cause packet loss. Security, privacy, NEMO route optimization and disconnection / isolation, are the challenges of network mobility for VANET.

3 Analysis

In this section we will conduct a comparative analysis of above protocol stacks. Simulation platform for VANET is still an open problem for the research community. There are two types of simulators available (i) Network Simulators (ii) Traffic Simulators. Both types of simulators should be used simultaneously to get better results. There are a few tools available for VANET simulation but most of them have the problem of improper ‘interaction’. Thus the proper selection of a simulator is a big question that must be properly analyzed and considered before starting any research. MOVE, Trans, VanetMobiSim, NCTUns, NS2, NS3, and NCTUns are the examples of VANET simulations. Following is a table of the comparisons between the various VANET architectures, achieved through simulations.

Table 1. Comparison between C2C-CC, CALM and WAVE

Parameters/Protocol Stacks	C2C-CC	CALM	WAVE
Promotion Identity	Industry consortium of car manufacturers	Standard body	Standard body
Focused on	Car to car multi-hop and geo networking.	Multiple media (802.11p, DSRC, W-LAN etc)	Only 802.11p at MAC layer for purely emergency messages
PHY Layer	DSRC and other WLAN standards	Combination of different technologies	DSRC only
Wireless Technology	Support for Media Dependent and Media Independent Part	Interface Abstraction	Only PHY layers specific to 802.11p
Target Applications	Safety	Non safety and Critical	Safety
Support for Application Types	Active Safety, Traffic Efficiency, Infotainment	Non-IP CALM Aware, IPV6 CALM Aware, IPV6 Legacy	Safety Non-IP, Non Safety IPV6
Addressing Schemes	Geo-routing	Mainly IP addressing	IP addressing
Routing Schemes	Based on MAC protocol (Receiver based) plus IPV6	Mobile IPV6	Different channel allocation plus IPV6
Security issues	Different procedures adopted like certificate etc	Not very clearly defined and addressed	Defined procedures like certificates and signatures
Number of Hops	Single Hop & Multi-hop	Single hop	Single Hop
Communication Mode	Uni-cast, Broad-cast, Geo-uni-cast, Geo-broadcast	Uni-cast, Broad-cast	Uni-cast
Group Addressing	Geo-Addressing	Via Service Initialization	Via WBSS
Flexibility	Flexible and ideal to adapt	Very flexible but has lot of confusions	Not flexible , restricted to single MAC layer
Simulators	GeoNet claimed to have patch in NCTUNs but do not have any open source simulator	No open source simulator	Open source simulator i.e. NCTUNs

The table shows the comparison of the different VANETs protocol stacks. SAFESPOT project documentation provided a lot of material for this comparison. The comparisons show that for short range communication, the two better options are WAVE and C2C-CC. In C2C-CC, there is a provision for single and multi hop as well as Geo-hop communications so we need more resources for multi hop communications in terms of buffer space and processing to process each packet. The main difference is the flexibility parameter. It can be concluded here that C2C-CC provides flexibility to use multiple interfaces and multiple MAC and network layer protocols. It is also fast and reliable due to its geographical routing concept. So it will be suitable to adapt for vehicular communication. However, still there is no open source simulator for C2C-CC, which is a big hurdle for the research community.

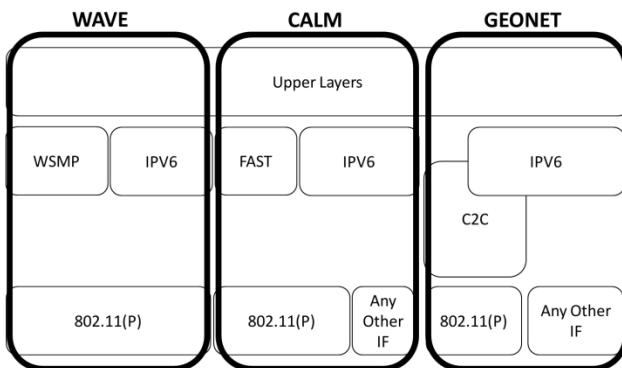


Fig. 4. VANET Architecture Comparison

WAVE is not flexible in terms of MAC and PHY. IEEE 802.11p can be used for scenarios where the focus is on short range communication. Also, every vehicle must have an IEEE 802.11p interface. However, it is very suitable for safety messages as it uses the concept of the dedicated CCH through which urgent traffic can be prioritized. WAVE is also popular among researchers as this is the only architecture with full simulation support [6], [7], [8], [9]. NCTUNs supports the protocol stack of IEEE 802.11p/WAVE (but it does not provide support for IPV6 and WSMP). For C2C-CC there is still no open source simulator.

ISO CALM is a combination of different technologies from 802.11p to UMTS, so its implementation with interface handover is a tough job. No open source simulator supports the complete CALM protocol stack.

In our opinion C2C-CC is flexible and more suitable for adaptation. As in the architecture of C2C-CC at MAC and PHY we can use different WLAN options like IEEE 802.11/a/b/g/p. So if our selection criteria of a protocol stack is flexibility then we should select C2C-CC else it should be WAVE being less complex. At same time, C2C-CC has adequate security mechanisms, which have been well defined. Traffic management applications create attraction for C2C-CC in the scenarios where we want to manage traffic load by providing different information to the drivers. For such applications in C2C-CC, C2CNet (Non-IP for urgent messages) and traditional

TCP/UDP and IP protocol stack can be used. Additionally C2C-CC clearly discusses its suitability for different applications like forward collision avoidance, crossing traffic, red light violation, local danger warning, electronic toll collection, emergency vehicles warning etc whereas in WAVE and CALM no such detail is available.

4 Conclusion and Future Work

In this paper, a review of available architectures for VANET was provided based on our survey. WAVE, C2C-CC and CALM were found to be the dominating protocol stacks. A comparative analysis of these protocol stacks was performed and C2C-CC was found to be a suitable candidate for VANET communication due to its flexibility in terms of protocol options, media access and security. Although simulation tools for C2C-CC are not available as open source, but hopefully, a few will be available in near future.

In future, we intend to compare these protocol stacks on basis of simulations performed under a single platform.

Acknowledgement

The authors want to thank Ms. Sana Ajmal for her valuable contribution in reviewing this paper.

References

1. Car to Car Communication Consortium Manifesto, v1.1 (August 2007)
2. Olivia, B., Martin, K., et al.: A Network Centric Simulation Environment for CALM-based Cooperative Vehicular Systems: SIMUTools. In: Proceedings of the 3rd International ICST; Conference on Simulation Tools and Techniques, Torremolinos, Malaga, Spain, March 15–19 (2010), ISBN 78-963-9799-87-5
3. Task Group p, IEEE P802.11p: Wireless Access in Vehicular Environments (WAVE). draft standard ed. IEEE Computer Society, Los Alamitos (2006)
4. Wang, S.Y., Lin, C.C., et al.: On Multi-hop Forwarding over WBSS-based IEEE 802.11(p)/1609 Networks. In: IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, Tokyo, September 13-16 (2009), E-ISSN 978-1-4244-5123-4
5. Farah, A.E., Bertrand, D.: A Light Architecture for Opportunistic Vehicle-to-Infrastructure Communications. In: Proceedings of the 8th ACM International Workshop on Mobility Management and Wireless Access, MobiWac 2010, Bodrum, Turkey, October 17-18 (2010)
6. Yi, W., Akram, A., et al.: IEEE 802.11p Performance Evaluation and Protocol Enhancement. In: Proceedings of the IEEE International Conference on Vehicular Electronics and Safety, Columbus, OH, USA, September 22-24 (2008)
7. Todd, M., Tammy, M., et al.: Measuring the Performance of IEEE 802.11p Using ns-2 Simulator for Vehicular Networks. In: IEEE International Conference on Electro/Information Technology (EIT), Ames IA (2008)

8. Stephan, E.: Performance Evaluation of the IEEE 802.11p WAVE Communication Standard. In: 66th IEEE International Conference on Vehicular Technology, VTC (2007)
9. Kang, L.C., Huang, L.C.: Development of Telematics Communication System with WAVE DSRC. In: Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA (October 2009)
10. European ITS Communication Architecture: Overall Framework, Proof of Concept and Implementation, Draft Version 2.0, COMeSafety Specific Support Action (October 2008)
11. Final GeoNet Specification, GeoNet Deliverable D2.2 (January 2010)
12. Manabu, T., Ines, J.B., et al.: Experimental Evaluation for IPv6 over VANET Geographic Routing. In: Proceedings of The International Wireless Communications and Mobile Computing Conference, IWCMC 2010, Caen, France, June 28-July 2 (2010)
13. Roberto, B., Andreas, F., et al.: A MANET-centric Solution for the Application of NEMO in VANET using Geographic Routing. In: TridentCom 2008, Innsbruck, Austria, March 18-20 (2008)

Behavior Specification of a Red-Light Violation Warning Application – An Approach for Specifying Reactive Vehicle-2-X Communication Applications

Sebastian Röglinger and Christian Facchi

Institute for Applied Research, Ingolstadt University of Applied Sciences
Esplanade 10, 85049 Ingolstadt, Germany

{sebastian.roeglinger, christian.facchi}@haw-ingolstadt.de
<http://www.haw-ingolstadt.de/>

Abstract. Current intelligent transportation systems are based on Vehicle-2-X communication. They accomplish the next level of cooperative advanced driver assistance systems. Since these Vehicle-2-X communication applications are often located in the safety critical areas, high quality demands arise. One part of these requirements is the correct functional behavior. For its validation, a formal specification is required. In this paper, the *Red-Light Violation Warning* application is used. To specify all the functionality, *Continuous-Time State Machines* (CTSM) are introduced. Such, a formal specification of a Vehicle-2-X application is given and, consequently, no unintended ambiguities exist furthermore.

Keywords: Modeling Reactive System, Vehicle-2-X Communication, Finite State Machine, Red-Light Violation Warning.

1 Introduction

Current *Vehicle-2-X* (V2X) communication applications are often targeted to improve traffic safety and, thus, high-quality requirements have to be met. To be able to achieve this quality demands, it is necessary to use unambiguous specifications of the applications. But, unambiguous specifications require the usage of well-suited specification techniques. Hence, it is worth to spend efforts in selecting the right specification technique.

There are currently several V2X applications in the focus of research activities. Those V2X applications can, according to [7][10][13], be grouped into: safety, transport efficiency, and infotainment/other applications. A *Red-Light Violation Warning* (RVW) application is assigned to the safety application group. It aims to warn the driver if he or she will most probably violate a red-light sign. A typical scenario is that a driver approaches at an intersection and does not recognize that he or she has to stop because the corresponding traffic light shows a red sign. In that case, the RVW application should warn the driver in sufficient time. Typically, those warnings follow a double-staged approach.

Since RVW applications require information about surrounding traffic lights, an information exchange between traffic lights and the vehicle is required. According to [79], at least in Europe, the communication will be based on *Dedicated Short Range Communication* (DSRC) (e.g., WiFi using IEEE 802.11 p [1] or ETSI ES 202 663 [1]) and cellular mobile communication (e.g., UMTS and beyond). However, the communication itself and resulting requirements are not part of this paper.

In this paper, a behavioral specification of a Red-Light Violation Warning application is given. The specification contains functional aspects, but does not contain any implementation details. Since the specification covers only the RVW application from the user's (driver's) point of view, any realization demands (e.g., maximum allowed package loss rate) are omitted. To specify all the functionality, *Continuous-Time State Machines* (CTSM) are introduced and used. So, a formal specification of a Vehicle-2-X application is given and, consequently, no unintended ambiguities exist furthermore.

The remainder of this paper is structured as follows. Section 2 states the requirements for the specification technique. In Section 3, related work towards behavior specification with state machines is discussed. Section 4 introduces CTSMs in combination with its semantics. Next, Section 5 presents a behavior specification of a Red-Light Violation Warning application by using a CTSM. Finally, conclusions and the future work are stated in Section 6.

2 Requirements for the Specification Technique

A specification technique for a RVW application has to cover specific requirements. During the work in our project, the following points raised:

Transitions Dependent on Multiple Continuous-Time Input Signals

RVW application implementations have to react on several physical values like the ego-vehicle speed or the distance between the vehicle and the stop line. Those physical values are fed into ports of RVW applications as signals. Since such physical values change continuously, specification techniques, which deal with real world scenarios, have to be capable to handle continuously changing input signals.

Transitions Dependent on Timing Behavior Specifications

Specifications of RVW applications might contain delay times before the system should give warnings. During delay times, the driver has the opportunity to leave the critical region and, thus, get not annoyed by a too responsive system. Generally, time is important for reactive systems and, consequently, transitions have also to be able to fire dependent on timing constraints.

Output Events

RVW applications interact with the vehicle's driver via warning signals. So, the used specification technique has also to be able to deal with output events.

Variables

Specifications of RVW applications might require to store information, e.g. the point in time where a high-possible red-light violation was detected the first time. Generally, variables provide the possibility to store and recall information which is necessary if a reactive system has not only to be able to use the current input values but also preceding data.

To satisfy these points, a state machine called *Continuous-Time State Machine* (CTSM) is introduced in Section 4. With an application specification using a CTSM, it is feasible to specify the behavior of a Red-Light Violation Warning application. Nevertheless, CTSM models stay manageable, and clear. So, it is possible to get a fast overview of modeled applications and avoid modeling mistakes.

3 Related Work

Using state machines for specification purposes is a widespread technique. Hence, there are many different types of state machines described in the literature. In the following, a selection of relevant state machines is introduced. Those state machines are discussed regarding their capability to be used for specifying a RVW application by fulfilling the requirements stated in Section 2.

Labelled Transition Systems (LTS) [16] are enhanced *Finite State Machines* (FSM) [14] whose sets of states and alphabets may be infinite. The alphabet is a subset of actions¹, without distinguishing between input and output actions. Indeed, actions improve expressiveness of state machines, but, it is still not possible to handle variables or time. At least for the last point, *Timed Automata* (TA) [16] gives a solution by introducing the concept of clock variables which increase monotonically. Clock variables can be reset by firing transitions. This can be used to specify clock invariants assigned to states and to set clock constraints on transitions.

As described in [5], LTS are enhanced by *Input-Output Automata* (IOA) [18], *Input Output State Machines* (IOSM) [19], and *Input Output Transition Systems* (IOTS) [20] by making a distinction between input and output actions. These automata mainly differ in dealing with locally controlled vs. not locally controlled actions. Since these automata react only on actions, it is not comfortable to model the behavior based on multiple continuous-time input signals which have to be treated as continuous flows of actions in that case. Moreover, these automata cannot handle time.

Extended Finite State Machines (EFSM) as described in [8][17][22] provide transitions labeled with guard functions which facilitate EFSM to group internal

¹ *Actions*: In terms of automata, *events* are often called actions. Such actions are used for event-based internal and external communication. Output and internal actions are also called *locally controlled* because they are performed by the automata but independent of the environment. Instead, input actions are not locally controlled because they are performed by the environment. [5]

states into visible state². Moreover, EFSM provide update functions to change variables at each transition. But, EFSM can only change their current state if an input symbol appears. So, multiple continuous-time input signals could not be handled comfortable because every state would require extra transitions for each allowed combination of valuations of the multiple input signals, known as transition explosion problem. In addition, EFSMs are not able to deal with time. E.g., it is not possible to stress a transition to fire if a timer runs out, as done by TA. Furthermore, EFSM semantics require discrete input actions to fire a transition.

Beyond these state machine types, also OMG's UML state machines provide comprehensive expressiveness. The OMG defines the UML semantics just in a semi-formal natural language [2]. But, there are also several approaches to define formal semantics for UML state machines, like [15,20]. However, the semantics of UML generally is still undergoing extensive investigation and the approaches often use implicit assumptions, which make it hard to achieve a comprehensive overview of the available formal UML semantics definitions [12]. Moreover, the fact that there are several semantics established by different parties requires to choose one of these and keep track of it. In addition, like for EFSM, most of the UML state machine semantics approaches require discrete input actions to fire a transition.

In [3], R. Alur et al. introduced a state machine called *Hybrid Systems* (HS). HS state machines comprise many concepts which are adapted by the definition of CTSMs. E.g., transitions which fire automatically instantaneous if the valuation of a variable reaches a specified threshold. However, HS state machines model the whole hybrid system in one single model and assign a set of functions (called activities) to each state to describe the combined behavior of the environment and the controlling system. Moreover, HS state machines do not generate output events if a transition fires.

However, as the best of our knowledge, there is no state machine which can handle multiple continuous-time signals on input ports, uses internal variables, and generate output events on firing transitions in combination. But, during the work in our project, it turned out that the used specification technique has to be capable to handle these points. Therefore, CTSM are introduced in the next section.

4 Continuous-Time State Machine

The definition of Continuous-Time State Machines is given in this section. Figure 1 shows the basic concept of the interaction between CTSMs and their environment. The environmental physical values $f_{i_1}(t), \dots, f_{i_n}(t)$ are fed into a CTSM as continuous-time signals applied to ports i_1, \dots, i_n . Based on the port's values, the transitions a_1, \dots, a_3 fire and the corresponding outputs are given.

² *States of EFSM:* Visible states are an abstraction of internal states. This is required because every suitable combination of internal variable valuations causes an extra internal state, known as state explosion problem. [22]

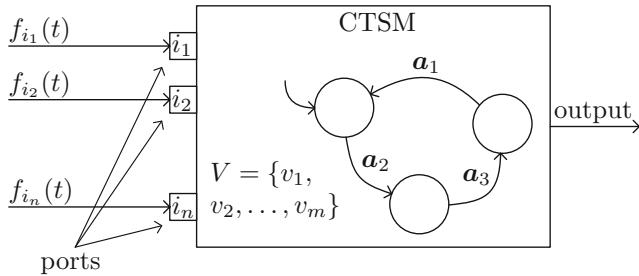


Fig. 1. Example of a CTSM with ports i_1, \dots, i_n , continuous-time input signals $f_{i_1}(t), \dots, f_{i_n}(t)$, and variables v_1, \dots, v_m

4.1 Syntax

In the following, relevant parts of the used notation are stated:

- The apostrophe (or prime) notation indicates the post-firing valuation of a variable symbol $v \in V$ of the state space D .
 - The n -th post valuations could be notated by n apostrophes or the variable symbol to the power of n in brackets, i.e., $\nu(a'') \equiv \nu(a^{(3)})$.
 - The apostrophe notation is only permitted in context of firing transitions. Thus, the valuations of different numbers of apostrophes of one variable symbol are all related to the same point in time.
 - The apostrophe notation is only permitted for variable symbols and is not available for input port symbols and the time symbol.
- The angle brackets $\langle \dots \rangle$ denote a tuple. The symbols, which represent tuples are boldfaced, e.g., $\mathbf{a} = \langle 1, 2, 3 \rangle$.
- The round brackets (\dots) denote either:
 - the application of a function, e.g., $a(x_1, x_2, \dots, x_n)$, or
 - the valuation of symbols of a map $\nu : D \rightarrow \mathbb{R}$, e.g., $\nu(a)$ with ν is a map and a is the symbol.
- The square brackets $[\dots]$ denote the parallel assignment of valuations to symbols of a map.
 - $\nu[a' := 5]$ denotes that the variable symbol a of the map ν is set to 5; afterwards the valuation of the symbol $\nu(a)$ is 5.
 - $\nu[a' := b, b' := a]$ denotes the switch of the values of the symbols a and b .
- $s \xrightarrow{\mathbf{a}} s'$ denotes that $(s_a == s \wedge s'_a == s') == \text{True}$ with $\mathbf{a} = \langle s_a, g_a, u_a, o_a, s'_a \rangle \in A$ and $s, s' \in S$
- $s \xrightarrow{\boldsymbol{\sigma}} s'$ denotes that $(s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} \dots \xrightarrow{a_{n-1}} s_n) == \text{True}$ with $\boldsymbol{\sigma} \in A^k$, $k \in \mathbb{N} \setminus \{0\}$, $\boldsymbol{\sigma} = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle$, $s_1 = s$, $s_n = s'$ and $s, s' \in S$

The CTSM is defined by the following 7 - tuple : $CTSM = \langle S, s_0, D, G, U, O, A \rangle$ with:

- S : finite set of states
- $s_0 \in S$: initial state
- $D = V \cup I \cup \{\tau\}$: finite set of state space symbols
 - V : finite set of symbols of variables; $V = \{v_1, \dots, v_m\}$
 - I : finite set of symbols of input ports; $I = \{i_1, \dots, i_n\}$
 - τ : time symbol
 - V, I , and $\{\tau\}$ are disjoint; $V \cap I \cap \{\tau\} = \emptyset$
- G : finite set of transition guards; $g \in G : \mathbb{R}^{|D|} \longrightarrow \mathbb{B}$
- U : finite set of update functions $u \in U : V \times (\mathbb{R}^{|D|}) \longrightarrow \mathbb{R}$
- O : finite set of output symbols
- $A \subseteq S \times G \times U \times O^* \times S$: finite set of labeled transitions. The first element of the tuple $a = \langle s, g, u, o, s' \rangle \in A$ is the origin state of the transition; whereas, the last element s' is the head state. The remaining elements g , u , and o are the transition's labels and contain the transition's guard, update function, and tuple of output symbols.

The transition guards $g \in G$ are boolean operations of predicates of relations over valuations of the symbols in D according to the inductive definition:

1. *True* and *false* are terms.
2. Predicates of relations over real numbers \mathbb{R} are terms.
3. If α is a term, then $\neg\alpha$ is also a term.
4. If α and β are terms, then $\alpha \wedge \beta$ and $\alpha \vee \beta$ are also terms.
5. Terms consist only of the rules 1. to 4.

E.g., examples could be seen in Equation (7), whereas, simplified notations were used if they could be syntactically transformed to a representation, which fulfills the inductive definition. Since the guards include only valuations of the symbols, it is allowed to use the symbols themselves instead of the projections of the symbols $\pi_d(\mathbf{d})$ with $d \in D$ and $\mathbf{d} \in \mathbb{R}^{|D|}$ for reasons of clarity. For instance, the following two equations (see also Equation (16)) should be treated as equal except for the used notation:

$$isInRange(\mathbf{d}) := d_{SL_min} \leq \pi_{d_{SL}}(\mathbf{d}) \leq d_{SL_max} \quad (1)$$

$$isInRange(\mathbf{d}) := d_{SL_min} \leq d_{SL} \leq d_{SL_max} \quad (2)$$

The update functions $u \in U$ are based on arithmetic terms, which use the operations: addition, subtraction, multiplication, division, and others to calculate the new value for the symbol given as the first parameter of the update function. So, a two-ary update function $u : V \times (\mathbb{R}^{|D|}) \longrightarrow \mathbb{R}$ with $r_v = u(v, \mathbf{d})$ is built, where v is the symbol whose valuation should be updated, \mathbf{d} is the current state space vector, and r_v is the new value for the symbol v could be seen as a unary function (currying) $u : V \longrightarrow \tilde{U}$ with $\tilde{u}_v \in \tilde{U} : \mathbb{R}^{|D|} \longrightarrow \mathbb{R}$ such that $\tilde{u}_v = u(v)$ and $r_v = \tilde{u}_v(\mathbf{d})$. This simplifies the notation of the update functions as it could

be seen exemplarily in the Equations (11), (12), and (14) because it provides the possibility to use a separate update function \tilde{u}_v for every $v \in V$. Additionally, the notation might also be simplified by omitting the projection notation $\pi_d(\mathbf{d})$ with $d \in D$ and $\mathbf{d} \in \mathbb{R}^{|D|}$ by using the symbols $d \in D$ as a replacement as for the guard functions.

4.2 Semantics

State Space. All elements of the state space D are based on real numbers \mathbb{R} and are represented by the variable symbols $v \in V$, the input port symbols $i \in I$, and the time symbol τ as illustrated in Figure 1. Valuations of the symbols are realized by applying the mapping $\nu : D \longrightarrow \mathbb{R}$.

Variables: The valuations of the variables $\nu(v)$ with $v \in V$ stay unchanged within one state until a transition fires. The update functions $u \in U$ overwrite the valuations of the variables when they are applied by firing transitions. The updates are applied in parallel. So, the valuation of the variables do not suffer from the execution order.

Input Ports: The valuations of the input port symbols $\nu(i)$ with $i \in I$ are permanently updated by the environment input signals and valuate to the current values at the corresponding input ports.

Time: The valuation of the time symbol $\nu(t)$ is monotonically updated and results in the point in time where the initial state s_0 was entered plus the length of time t_{run} since entering the initial state s_0 .

Transitions $a \in A$. The firing of transitions is deterministic. Thus, the semantics of firing transitions is defined as follows:

- $\mathbf{c} = \langle s_c, \mathbf{d}_c \rangle \in C$ with $C = S \times (\mathbb{R}^{|D|})$ is called configuration of a *CTSM* with $\mathbf{d}_c = \langle \nu(v_1), \dots, \nu(v_m), \nu(i_1), \dots, \nu(i_n), \nu(t) \rangle$ and s_c is the current state.
- A transition $a = \langle s_a, g_a, u_a, o_a, s'_a \rangle \in A$ fires immediately iff for the current configuration $\mathbf{c} \in C$ it holds that $s_a == s_c \wedge g_a(\mathbf{d}_c) == \text{True}$. In that case, all variables are updated by $\nu[v'_1 := u_a(v_1, \mathbf{d}_c), \dots, v'_m := u_a(v_m, \mathbf{d}_c)]$ in parallel. Thus, the configuration changes to $\mathbf{c}' = \langle s'_a, \langle \nu(v'_1), \dots, \nu(v'_m), \nu(i_1), \dots, \nu(i_n), \nu(t) \rangle \rangle$ and the machine gives the output o_a .
- To ensure the determinism of CTSMs, it has to hold that $\forall \mathbf{c} \in C : |\{a \in A | s_c == s_a \wedge g_a(\mathbf{d}_c) == \text{True}\}| \in \{0, 1\}$. With this property, it is guaranteed that no parallel firing transitions or non-deterministic behavior are feasible.

5 Specifying the Red-Light Violation Warning Application

In this section, the application of CTSMs is shown by specifying the behavior of a Red-Light Violation Warning application. According to [9], the aim of a RVW

implementation is: “Warns drivers when they are going to violate a red traffic light signal.”. Typically, the warnings occur in a double-staged approach. E.g., in the first stage the driver is informed by a visual or audible signaling; whereas, the second stage tries to wake up the driver by a short jerky braking maneuver if the driver shows no reaction on the first signaling.

Figure 2 presents the typical scenario. The shown vehicle contains several *Electronic Control Units* (ECU) and a *ITS Vehicle Station* (IVS). The traffic light is placed exactly at the lane’s stop line and broadcasts periodically among others the information about the signal phase timing: time to the next red-to-green switch t_{r2g} , and time to the next green-to-red switch t_{g2r} . If $t_{g2r} > t_{r2g}$, then the traffic light is currently red. Additionally, intermediate signal phases between red and green are regarded as green. Figure 2 also shows the other input ports: ego-vehicle speed v_{ego} , its derivative (acceleration) $a_{ego} = \dot{v}_{ego}$, and the distance between the vehicle and the lane’s stop line d_{SL} .

The example of this section is based on the requirements listed in the following. Certainly, this list of requirements is not complete, but adequate to show the concept of modeling reactive systems by the use of CTSMs.

- Warnings must not appear if the distance d_{SL} from the vehicle to the stop line is greater than d_{SL_max} or lower than d_{SL_min} .
- The prediction of the point in time where the stop line is crossed has to assume that the vehicle’s acceleration a_{ego} remains constant in the future.
- The first warning w_1 has to appear if the uninterrupted time where the prediction results in a red-light violation exceeds a waiting time. This waiting time has to be adapted to the situation. Since the adaption has to be determined experimentally, for this example, the adjustment is substituted by *sensitivity(d)* · 1.0 s.
- The second warning w_2 has to be given if the prediction results in red-light violation for a longer uninterrupted time than $t_{thres,w2}$ after the first warning w_1 appeared.
- Both warning states have to disappear (output event ‘0’) after the time t_{reset} elapsed and if the prediction results in no red-light violation.

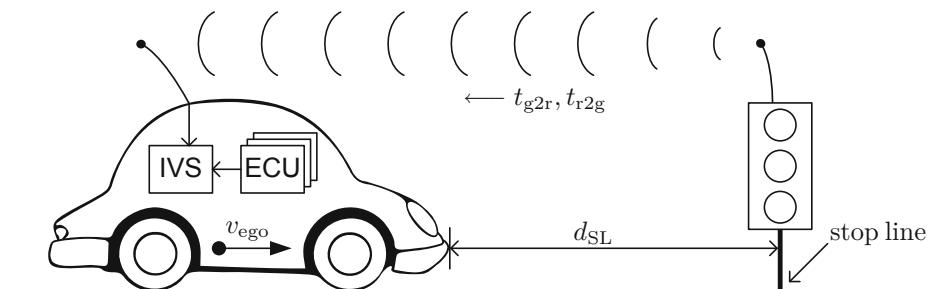


Fig. 2. Example scenario of the Red-Light Violation Warning application

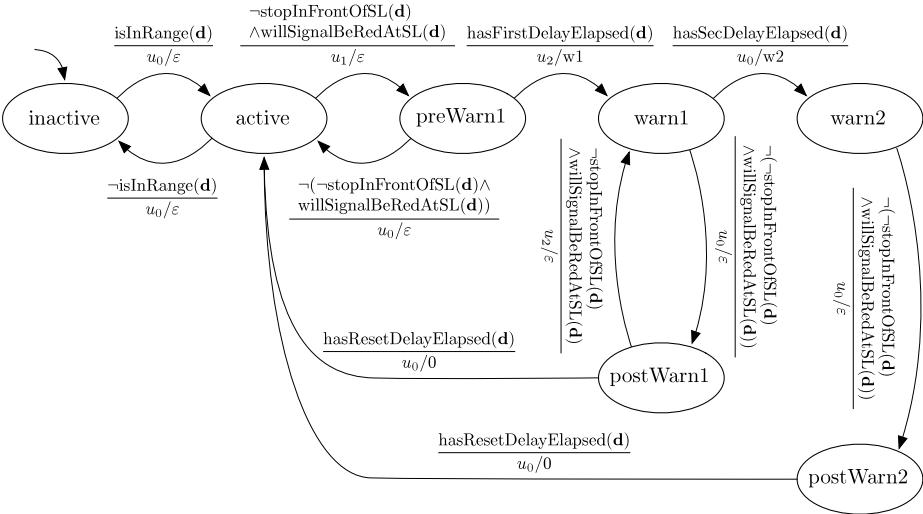


Fig. 3. Red-Light Violation Warning application specified by a CTSM

The following equations define the CTSM for the RVW. Whereas, Figure 3 depicts the appropriate state machine. For reasons of clarity, the guard functions in Equation (7) are built of supporting functions, defined in Equation (16) to (21). Similarly, the update functions in Equation (8) are also defined separately in Equation (11) to (15). The update function u_0 is the identity update function and, thus, does not change any variable. Instead, the update functions u_1 and u_2 modify at least one variable. Beyond this, the output symbols (see Equation (9)) include a empty element ε that is used if the CTSM should not generate any output.

The CTSM for the presented example is defined as follows:

$$S := \{\text{inactive}, \text{active}, \text{preWarn1}, \text{warn1}, \text{warn2}, \text{postWarn1}, \text{postWarn2}\} \quad (3)$$

$$s_0 := \text{inactive} \quad (4)$$

$$V := \{t_{\text{preWarn1}}, t_{\text{warn1}}\} \quad (5)$$

$$I := \{a_{\text{ego}}, v_{\text{ego}}, d_{\text{SL}}, t_{\text{g2r}}, t_{\text{r2g}}\} \quad (6)$$

$$G := \{\text{isInRange}(\mathbf{d}), \neg \text{isInRange}(\mathbf{d}), \neg \text{stopInFrontOfSL}(\mathbf{d}) \quad (7)$$

$$\begin{aligned} & \wedge \text{willSignalBeRedAtSL}(\mathbf{d}), \neg (\neg \text{isInRange}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d})), \\ & \text{hasFirstDelayElapsed}(\mathbf{d}), \text{hasSecDelayElapsed}(\mathbf{d}), \\ & \text{hasResetDelayElapsed}(\mathbf{d})\} \end{aligned}$$

$$U := \{u_0, u_1, u_2\} \quad (8)$$

$$O := \{\varepsilon, w_1, w_2, 0\} \quad (9)$$

$$A := \{\langle \text{inactive}, \text{isInRange}(\mathbf{d}), u_0, \varepsilon, \text{active} \rangle, \quad (10)$$

$$\langle \text{active}, \neg \text{isInRange}(\mathbf{d}), u_0, \varepsilon, \text{inactive} \rangle,$$

$$\begin{aligned}
& \langle \text{active}, \neg \text{stopInFrontOfSL}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d}), u_1, \varepsilon, \text{preWarn1} \rangle, \\
& \langle \text{preWarn1}, \neg (\neg \text{isInRange}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d})), u_0, \varepsilon, \text{active} \rangle, \\
& \langle \text{preWarn1}, \text{hasFirstDelayElapsed}(\mathbf{d}), u_2, \text{w1}, \text{warn1} \rangle, \\
& \langle \text{warn1}, \text{hasResetDelayElapsed}(\mathbf{d}), u_0, \text{w2}, \text{warn2} \rangle, \\
& \langle \text{warn1}, \neg (\neg \text{stopInFrontOfSL}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d})), \\
& \quad u_0, \varepsilon, \text{postWarn1} \rangle, \\
& \langle \text{postWarn1}, \neg \text{stopInFrontOfSL}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d}), \\
& \quad u_2, \varepsilon, \text{warn1} \rangle, \\
& \langle \text{warn2}, \neg (\neg \text{stopInFrontOfSL}(\mathbf{d}) \wedge \text{willSignalBeRedAtSL}(\mathbf{d})), \\
& \quad u_0, \varepsilon, \text{postWarn2} \rangle, \\
& \langle \text{postWarn1}, \text{hasSecDelayElapsed}(\mathbf{d}), u_0, 0, \text{active} \rangle, \\
& \langle \text{postWarn2}, \text{hasSecDelayElapsed}(\mathbf{d}), u_0, 0, \text{active} \rangle \}
\end{aligned}$$

The update functions are defined as follows:

$$u_0(v, \mathbf{d}) = \tilde{u}_{0,v}(\mathbf{d}) := \pi_v(\mathbf{d}) \quad \text{with } v \in V \quad (11)$$

$$u_1(t_{\text{preWarn1}}, \mathbf{d}) = \tilde{u}_{1,t_{\text{preWarn1}}}(\mathbf{d}) := \pi_\tau(\mathbf{d}) \quad (12)$$

$$u_1(v, \mathbf{d}) = \tilde{u}_{1,v}(\mathbf{d}) := \pi_v(\mathbf{d}) \quad \text{with } v \in V \setminus \{t_{\text{preWarn1}}\} \quad (13)$$

$$u_2(t_{\text{warn1}}, \mathbf{d}) = \tilde{u}_{2,t_{\text{warn1}}}(\mathbf{d}) := \pi_\tau(\mathbf{d}) \quad (14)$$

$$u_2(v, \mathbf{d}) = \tilde{u}_{2,v}(\mathbf{d}) := \pi_v(\mathbf{d}) \quad \text{with } v \in V \setminus \{t_{\text{warn1}}\} \quad (15)$$

The supporting guard functions are defined as follows:

- Test if the vehicle is in the required range:

$$\text{isInRange}(\mathbf{d}) := d_{\text{SL_min}} \leq d_{\text{SL}} \leq d_{\text{SL_max}} \quad (16)$$

- Test if the vehicle will stop in-front of the stop line:

$$\text{stopInFrontOfSL}(\mathbf{d}) := \begin{cases} d_{\text{SL}} < -\frac{v_{\text{ego}}^2}{2a_{\text{ego}}} & a_{\text{ego}} < 0 \text{ m/s}^2 \\ \text{False} & \text{else} \end{cases} \quad (17)$$

- Test if the traffic light will show a red-light sign at the point in time where the vehicle will cross the stop line. $t_{\text{SL}}(\mathbf{d})$ denotes the required time for the vehicle to reach the stop line and is defined in Equation (22). The upper line includes the test if the traffic light is currently red; whereas, the lower line includes the test if the traffic light is currently green:

$$\text{willSignalBeRedAtSL}(\mathbf{d}) := \begin{cases} t_{\text{SL}}(\mathbf{d}) < t_{\text{r2g}} & \text{for } t_{\text{g2r}} > t_{\text{r2g}} \\ t_{\text{SL}}(\mathbf{d}) > t_{\text{g2r}} & \text{else} \end{cases} \quad (18)$$

- Test if the delay time $\text{sensitivity}(\mathbf{d}) \cdot 1.0 \text{ s}$ has elapsed:

$$\text{hasFirstDelayElapsed}(\mathbf{d}) := \tau - t_{\text{preWarn1}} \geq \text{sensitivity}(\mathbf{d}) \cdot 1.0 \text{ s} \quad (19)$$

- Test if the delay time between the first and second warning level has elapsed:

$$\text{hasSecDelayElapsed}(\mathbf{d}) := \tau - t_{\text{warn1}} \geq t_{\text{thres_w2}} \quad (20)$$

- Test if the reset delay time has elapsed:

$$\text{hasResetDelayElapsed}(\mathbf{d}) := \tau - t_{\text{warn1}} \geq t_{\text{reset}} \quad (21)$$

The function to calculate the required time for the vehicle to reach the stop line is defined as follows:

$$t_{\text{SL}}(\mathbf{d}) := \begin{cases} \frac{-v_{\text{ego}} + \sqrt{2a_{\text{ego}}d_{\text{ego}} + v_{\text{ego}}^2}}{a_{\text{ego}}} & \text{for } a_{\text{ego}} \neq 0 \text{ m/s} \wedge v_{\text{ego}} > 0 \text{ m/s} \wedge d_{\text{SL}} > 0 \text{ m} \\ & \wedge 2a_{\text{ego}}d_{\text{ego}} + v_{\text{ego}}^2 \geq 0 \text{ m}^2/\text{s}^2 \\ \frac{d_{\text{ego}}}{v_{\text{ego}}} & \text{for } a_{\text{ego}} = 0 \text{ m/s}^2 \wedge v_{\text{ego}} > 0 \text{ m/s} \wedge d_{\text{SL}} > 0 \text{ m} \\ \infty & \text{else} \end{cases} \quad (22)$$

The first line of Equation (22) includes the calculation of $t_{\text{SL}}(\mathbf{d})$ in the case of the vehicle approaches at the stop line and accelerates or decelerates. It is the result from solving the equation:

$$d_{\text{SL}} = \int_0^{t_{\text{SL}}} v_{\text{ego}} + a_{\text{ego}} \cdot t \, dt \quad (23)$$

for t_{SL} by using the conditions appropriate to the first line of Equation (22). The second line of Equation (22) includes the calculation of $t_{\text{SL}}(\mathbf{d})$ in the case of the vehicle approaches at the stop line with constant speed. It is the result from solving the equation:

$$d_{\text{SL}} = \int_0^{t_{\text{SL}}} v_{\text{ego}} \, dt \quad (24)$$

for t_{SL} . Finally, the third line of Equation (22) catches the calculation if the vehicle will not reach the stop line because the deceleration results in a stop of the vehicle earlier.

In short, the state machine (Figure B) keeps a manageable size and the guard functions defined in Equation (16) to (22) are also easy to specify. Thus, a useful specification with clear semantics is available. Since, the specification is unambiguous and from the user's point of view, it assists test designers when creating test cases on system test level. Finally, this comprehensible specification also supports the communication between different parties of RVW development projects at any time during the project and beyond. So, pure textual descriptions of the intended behavior, which is not always clear, could be omitted.

6 Conclusion and Future Work

In the presented paper, a new way for specifying reactive systems has been shown. Therefore, *Continuous-Time State Machines* (CTSM) have been introduced. The presented specification mechanism is based on finite state automata enriched by following features:

- Transitions based on multiple input ports
- Transitions based on timing behavior without necessary event triggers
- Continuous-time
- Modeling physical behavior equations
- Output events
- Use of internal variables

The use of CTSMS has been shown on the *Red-Light Violation Warning* application. Such, the functional behavior of this application has been specified in a formal way. Based on this specification, some properties of the specification can be shown. Even a validation is possible. However, there is some more work to be finished:

- It has to be checked, whether the used specification mechanism is usable also for other V2X-applications. However, as far as it can be seen now the mechanism can be used on other applications.
- A methodology for validation of an application's requirements based on CTSMSs has to be developed.
- A methodology to create test-cases has to be defined. So, based on the described specifications, a way in the proper definition of test-cases has to be found. Therefore, a high test-coverage as well as a proper use of test resources has to be considered.
- Methodologies for checking CTSMSs upon infinite loops and contradictions have to be investigated.
- Further, details of the formal semantics of CTSMS have to be elaborated.
- The modeling capabilities have to be extended to cover message based communication protocols.

In this paper, a mechanism for specifying reactive systems has been introduced and its application has been demonstrated by an example. However, the work is only in an initial stage, but first results are promising.

Acknowledgments. The authors acknowledge the support for this research which has been granted by the research program at Bavarian, Germany universities of applied sciences and is supported by the AUDI AG Ingolstadt, Germany. In addition, grateful acknowledgement is made to Peter Trapp and Markus Meyer for proofreading and providing valuable comments.

References

1. IEEE Draft Standard, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment: Wireless Access in Vehicular Environments (2010)
2. OMG Unified Modeling Language (UML) Infrastructure Version 2.3 (2010), <http://www.omg.org/>

3. Alur, R., Courcoubetis, C., Halbwachs, N., Henzinger, T.A., Ho, P.-H., Nicollin, X., Olivero, A., Sifakis, J., Yovine, S.: The Algorithmic Analysis of Hybrid Systems. Elsevier: Theoretical Computer Science 138(1), 3–34 (1995)
4. Alur, R., Dill, D.L.: A Theory of Timed Automata. Theoretical Computer Science 126(2), 183–235 (1994)
5. van der Bijl, M., Peureux, F.: 7. I/O-automata Based Testing. In: Broy, M., Jonsson, B., Katoen, J.-P., Leucker, M., Pretschner, A. (eds.) Model-Based Testing of Reactive Systems. LNCS, vol. 3472, pp. 173–200. Springer, Heidelberg (2005)
6. Brandán Briones, L., Röhrl, M.: 8. Test Derivation from Timed Automata. In: Broy, M., Jonsson, B., Katoen, J.-P., Leucker, M., Pretschner, A. (eds.) Model-Based Testing of Reactive Systems. LNCS, vol. 3472, pp. 201–231. Springer, Heidelberg (2005)
7. Car 2 Car Communication Consortium: Manifesto: Overview of the C2C-CC System (2007), <http://www.car-to-car.org/>
8. Cheng, K.T., Krishnakumar, A.S.: Automatic Functional Test Generation Using the Extended Finite State Machine Model. In: DAC 1993: Proceedings of the 30th International Design Automation Conference, pp. 86–91. ACM, New York (1993)
9. COMeSafety: Deliverable D31 European ITS, Communication Architecture, Overall Framework, Proof of Concept Implementation, Version 3.0 (2010), <http://www.comesafety.org/>
10. ETSI: ETSI TR 102 638, Basic Set of Applications, Definitions, V1.1.1 (2009)
11. ETSI: ETSI ES 202 663, European profile standard for the physical and medium access control layer of Intelligent Transport Systems operating in the 5 GHz frequency band, V1.1.0 (2010)
12. Harel, D., Rumpe, B.: Modeling Languages: Syntax, Semantics and All That Stuff Part I: The Basic Stuff. Tech. rep. (2000)
13. Hartenstein, H., Laberteaux, K.P.: A Tutorial Survey on Vehicular Ad Hoc Networks. IEEE Communications Magazine, 164–171 (2008)
14. Jonsson, B.: 21. Finite State Machines. In: Broy, M., Jonsson, B., Katoen, J.-P., Leucker, M., Pretschner, A. (eds.) Model-Based Testing of Reactive Systems. LNCS, vol. 3472, pp. 610–614. Springer, Heidelberg (2005)
15. Jürjens, J.: A UML Statecharts Semantics with Message-Passing. In: Symposium of Applied Computing, pp. 1009–1013 (2002)
16. Katoen, J.-P.: 22. Labelled Transition Systems. In: Broy, M., Jonsson, B., Katoen, J.-P., Leucker, M., Pretschner, A. (eds.) Model-Based Testing of Reactive Systems. LNCS, vol. 3472, pp. 615–616. Springer, Heidelberg (2005)
17. Krishnakumar, A.: 16. Reachability and Recurrence in Extended Finite State Machines: Modular Vector Addition Systems. In: Courcoubetis, C. (ed.) CAV 1993. LNCS, vol. 697, pp. 110–122. Springer, Heidelberg (1993)
18. Lynch, N.A., Tuttle, M.R.: Hierarchical Correctness Proofs for Distributed Algorithms. In: PODC 1987: Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing, pp. 137–151 (1987)
19. Phalippou, M.: Relations d'Implantation et Hypothèses de Test sur des Automates à Entrées et Sorties. Ph.D. thesis, L'Université de Bordeaux I, France (1994)
20. Seifert, D.: An Executable Formal Semantics for a UML State Machine Kernel Considering Complex Structured Data. Research Report, Université Henri Poincaré - Nancy I - Université Nancy II - Institut National Polytechnique de Lorraine (2008)
21. Tretmans, G.J.: Test Generation with Inputs, Outputs and Repetitive Quiescence. Software – Concepts and Tools 3(TR-CTIT-96-26) (1996)
22. Utting, M., Legeard, B.: Practical Model-Based Testing: A Tools Approach. Morgan Kaufmann, San Francisco (2007)

Wireless Protocol Design for a Cooperative Pedestrian Protection System

Dirk Lill¹, Manuel Schappacher¹, Shahidul Islam¹, and Axel Sikora²

¹ Steinbeis Innovation Center Embedded Design and Networking, Baden-Wuerttemberg Cooperative State University Loerrach, Poststrasse 35, D79423 Heitersheim, Germany

² Department of Information Technology, Baden-Wuerttemberg Cooperative State University Loerrach, Hangstrasse 46-50, D79539 Loerrach, Germany
{dirk.lill,manuel.schappacher,shahidul.islam}@stzedn.de,
sikora@dhw-loerrach.de

Abstract. In the research project Ko-TAG, which is part of the research initiative Ko-FAS, cooperative sensor technology is developed and its benefit for traffic safety applications is evaluated. As a result the new sensor concept shall provide essential input data for next generation driver assistant systems. A secondary radar principle based on communication signals enables localization of objects with simultaneous data transmission. It mainly concentrates on pedestrian and other vulnerable road user (VRU) detection.

This paper describes the architectural approach of the system design, as well as the main characteristics of the physical and data link layers of the proposed wireless protocol. The protocol is designed in such a way that it can be used as an extension to IEEE802.11p/WAVE protocol. The complete protocol is modeled in OPNET network simulator.

Keywords: VRU eSafety, localization, secondary surveillance radar, time of flight, angle of arrival, network simulation.

1 Introduction

Safety is amongst the strategic objectives of today's automotive industry. Although the absolute number of accidents and fatalities in traffic is continuously decreasing (at least in Europe), the relative risk for vulnerable road users (VRUs) increases, i.e. for pedestrians, cyclists, powered two wheelers, etc. [1].

Today's advanced driver assistance systems (ADAS) are based on on-board environment perception sensors. Car2X communication technology enables a number of new use cases in order to improve driving safety or traffic efficiency and provide information or entertainment to the driver [2].

It is only shortly that predictive pedestrian protection systems (PPPS) can be used. These predictive systems use contact- or non-cooperative perception sensors systems for the detection of a pedestrian in the vicinity. They are limited by uncertainties in target classification. Moreover, these systems offer no benefit in case of fully or partially hidden pedestrians like e.g. children hidden by cars parked at the roadside.

Cooperative pedestrian protection systems (CPPS) that follow the communication model of secondary surveillance radar from air traffic control overcome these weaknesses. Cooperative systems may be used alone or in conjunction with image-based systems. They may be seamlessly integrated into other upcoming car2X-communication systems.

Promising results of two predecessor projects AMULETT [5] and WATCH-OVER [6] led to the current activities within the research initiative Ko-FAS from the German Ministry of Economics and Technologies (BMWi) [3]. The sub-project Ko-Tag concentrates on the development of a pedestrian tag that can communicate with a vehicle and can be localized by the vehicle. Within the Ko-Tag project, it is the task of the authors to provide a stable, secure and scalable network architecture between vehicles and VRUs that integrates into the upcoming C2C- and C2I-networks.

2 Technological Background

The IEEE802.11p standard has developed as the most popular basis for C2X-communication.

For the USA, the “IEEE 1609 Family of Standards for Wireless Access in Vehicular Environment (WAVE)” [4] defines an architecture and a complementary, standardized set of services and interfaces that collectively enable secure car-to-car (C2C) and car-to-infrastructure (C2I) wireless communications.

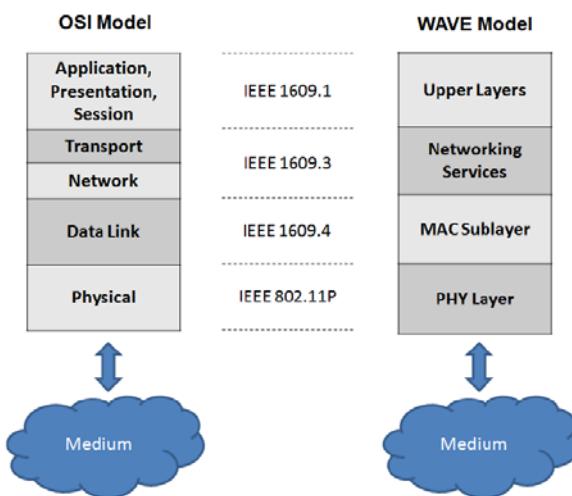


Fig. 1. Protocol stack of IEEE802.11p & IEEE1609 WAVE

The European activities are also using Wireless LAN and a frequency spectrum in the 5.9 GHz range that has been allocated on a harmonized basis in Europe in line with similar allocations in USA. The higher layers are defined by Car2Car Communication Consortium [2].

ETSI as European regulatory body has opened ten channels in the 70 MHz broad frequency range between 5.855 and 5.925 GHz in ETSI EN 302 571 [7]. This allocation is listed in Table 1. It should be highlighted that the seven 10 MHz-channels are non-overlapping and can be used independently. Alternatively, the bandwidth can be split into two 20 MHz and one 30 MHz channel. IEEE802.11p uses these channels with an OFDM PHY.

Table 1. Available channels in the 5.9 GHz band as defined by ETSI EN 203 571 [7]

Channel number	Carrier centre frequency [MHz]	Maximum channel bandwidth [MHz]
1	5860	10 (5855-5865)
2	5865	20 (5855-5875)
3	5870	10 (5865-5875)
4	5880	10 (5875-5885)
5	5890	10 (5885-5895)
6	5900	10 (5895-5905)
7	5910	10 (5905-5915)
8	5915	20 (5905-5925)
9	5920	10 (5915-5925)
10	5890	30 (5875-5905)

3 Architectural Design

3.1 Topologies

Cooperative sensors for VRU detection must integrate into the overall system of the upcoming traffic communication applications. Fig. 2 shows that an additional communication layer is added to the existing system of infrastructure (I) and vehicles (V). The architectural design is in line with the requirements of the project. It was already described in [8].

It is anticipated that VRUs may communicate with V and with I nodes. This latter may be of importance, e.g. at signalized intersections.

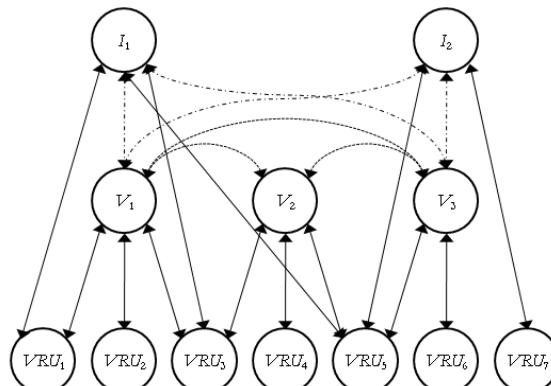


Fig. 2. Integration of VRU communication into the overall communication architecture

With regard to the drawing in Fig. 2, some of the topology requirements can be derived:

- *relationship models*: One VRU may communicate with no (VRU₇), one (VRU₄) or more than one (VRU₃) vehicle nodes. One VRU may communicate with no (VRU₆), one (VRU₇) or more than one (VRU₅) infrastructure nodes.

As a result, we have (at minimum) a two-layer relationship with an extremely large number of mutual permutations.

- *overlaid stars*: In this topology, we see only a topology of overlaid stars. Mesh topologies or hierarchical trees provide major overhead to the network characterized by its extreme mobility and dynamics.

- *flatness*: Whereas we see a hierarchical communication model through the three layer *I-V-VRU*, all nodes on one layer are equal and form a flat topology, i.e. there is no master vehicle from a communication point of view. This approach impedes centralized management or data streaming solutions.

- *range*: Analyses of accidents show that the precise localization deadline is at a distance of around 50 m under the assumption of intra-urban vehicle speeds of 50 km/h. This deadline leads to the range requirement of around 75 – 100m for communication setup and first measurements.

It should be highlighted that this range should be achievable not only under free-space conditions, but in real environments, additionally taking into account multi-path propagation, attenuation (i.e. human bodies), diffusion, and refraction.

- *density*: The density of vehicles may be high, however is limited through the dimensions of a car. Consequently, for car2car-communication, a maximum density of 1 node per 20 m² (traffic jam) or per 60m² (city traffic) is typically anticipated. VRU nodes may be much denser. This especially holds true for intra-urban situations at intersections, where waiting pedestrians may reach a density of 2 to 3 nodes per m².

- *scalability*: The required range of (for example 75 m) leads to a maximum area of $\text{Pi}*(75 \text{ m})^2 = 17,662 \text{ m}^2$, which on one hand, together with the anticipated density would lead to maximum number of up to 50.000 VRU nodes within the range of one vehicle node. Of course, one might argue that no-one would crash incidentally into a crowd of 50.000 people. Nevertheless, the system should remain stable also in such use-cases.

On the other hand, one could also think of sparsely populated topologies, where only one vehicle and VRU are within the communication range.

- *dynamics*: Topologies may change fast. Vehicles move at 50 ... 70 km/h in towns. No ex-ante network management seems to be possible, as the dynamics of node topologies can hardly be predicted – without prior communication channels.

In comparison with car2car-communication, the topologies may change at approximately the same speed, when VRUs on two-wheelers are taken into account.

3.2 System Design

The cooperative sensor system developed in the project Ko-TAG consists of localization units and transponders called SafeTAGs. The sensor concept allows for a selective localization of SafeTAGs and a selective communication link between localization units and SafeTAGs for an efficient use of bandwidth (cf. Fig. 3).

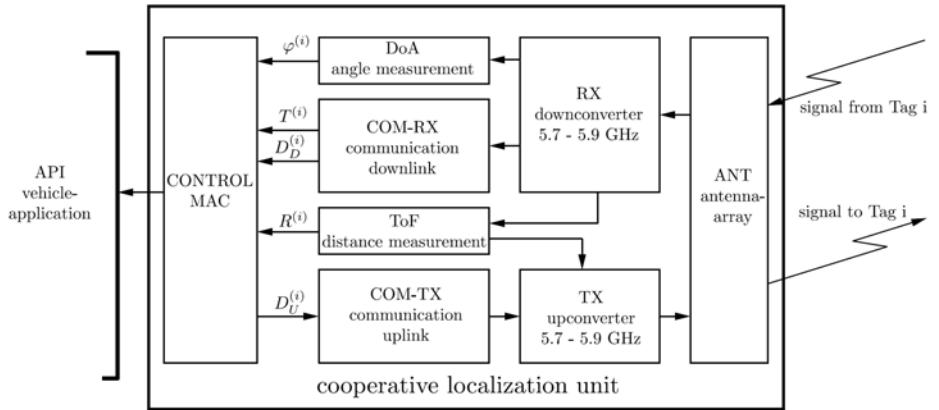


Fig. 3. Block diagram of a cooperative localization unit, data of SafeTAG i : $\varphi^{(i)}$ azimuth angle, $R^{(i)}$ radial distance, $T^{(i)}$ type of transponder, $D_D^{(i)}$ data transmitted from SafeTAG, $D_U^{(i)}$ data transmitted to SafeTAG

3.2 Protocol Design

As described above, no central node shall be available. Therefore, some kind of distributed coordination function (DCF) on the Medium Access Control (MAC) layer is required. Of course, an enormous variety of MAC protocols has been proposed for dynamic and distributed use cases [9]. However, a variety of fundamental design decisions are to be taken into account before the detailed protocol specification can be started. Those are:

- One fundamental design decision envisages asymmetry of the communication partners. VRU tags must be extremely low cost and ultra low energy, whereas the corresponding units at the vehicles may come at a moderate cost and energy consumption.
- It is anticipated that for the short reaction time, a real-time system would be the best fit. However, real-time communication systems require either a centralized master-slave system or a distributed coordination function (DCF) of a known topology. Both approaches are not practical, because neither a master vehicle shall exist nor topologies shall be restricted to static cases.
- Therefore, a DCF with a best effort approach remains as the only solution for the MAC function. As such, these functions are subject to conflicting objectives, again, and must maximize the probability of successful transmission while balancing the probability of frame collisions and the times, when no station accesses the channel. In order to reduce the collision probability, a time-slot based approach is pursued, following the immemorial argumentation of slotted vs. pure Aloha [10].
 - An extended literature scouting has been performed to verify if there was a MAC protocol suitable for direct use. Elements of DPSM (Dynamic Power-Saving Mechanism) [11], S-MAC (Sensor-MAC) [12], T-MAC (Timeout MAC) [13], and

Wise-MAC (Wireless Sensor MAC) [14] become part of the newly developed KoTAG protocol. Also, previous work [17] could be used.

- Another scouting activity was oriented towards the broad variety of networked time synchronisation protocols, also trying to find optimum working points for wireless sensor networks [15] [16]. However, all of those approaches either require a quite static topology or allow a limited accuracy only. Therefore, we anticipate a time-slot based approach with side channel synchronization, e.g. from a satellite clock (GPS, Galileo). This synchronization can be limited to the vehicle side only. If all vehicles work in a synchronized fashion, VRUs can be synchronized to all spatially distributed networks within a certain precision, once they are synchronized to one vehicle.

- The communication is organized in super frames, in which communication and localization between various pairs are organized.

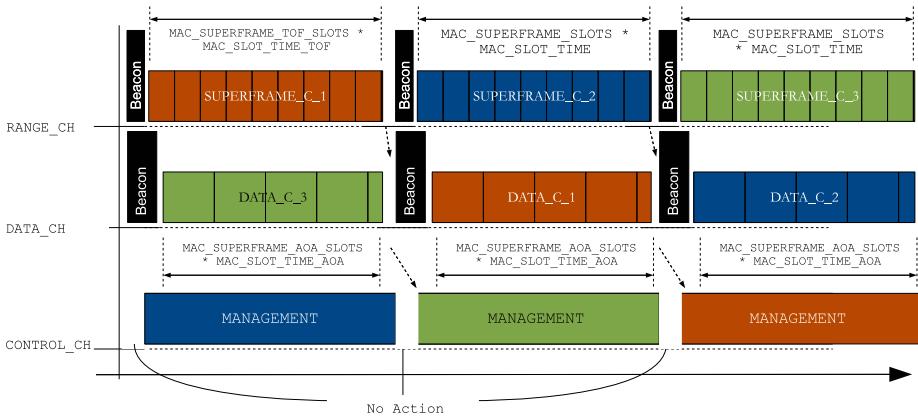


Fig. 4. Time- and frequency multiplexing in the proposed Multi-Channel Splitted Localization (MCSL) approach

- The channel access divides into three phases, which can be multiplexed in various dimensions, i.e. time and frequency (cf. Fig. 4):
 - a detection and network management phase, when VRU tags can register to vehicles.
 - a Time of flight (ToF) measurement phase, which is extremely time critical, as clock drifts between measurement partners lead to localization errors.
 - a Direction of Arrival (DoA) measurement phase, which can be used in parallel for data communication.

Thus, the channel access combines contention periods for arbitrary and management traffic, and non-contention periods (reserved time-slots) for time- and collision-critical operations.

The physical breakup into the different frequencies helps to solve the conflict of objectives with regard to bandwidth requirements.

- According to the radar theorem, the accuracy (resolution) of ToF-measurements grows with increasing bandwidth. Therefore, it is highly preferable to use a channel with maximum frequency bandwidth.

- On the other hand, the precision of DoA is best with a single carrier, i.e. with minimum bandwidth.

Basically, the management functionality can be merged with any of the two other channels. From practical point of view (ease of coding and modulation), a merger between DoA and management phases seems to be easiest.

The described breakup is generally compatible with the frequency band allocated for car-to-x-applications in the US, in Japan, and recently in Europe. It is also possible to adapt the KoTAG-physical layer completely to the IEEE802.11p standard. However, some aspects and especially the increased overhead on the physical frame format may reduce the system performance.

4 Simulation

4.1 Topologies

In parallel with the design of the protocol, a simulation environment is prepared in order to verify the macroscopic soundness of the chosen concept. OPNET models are currently being designed and some first simulation results evaluated.

4.2 Model Description

The OPNET models developed for the simulation are built up on different stages with an increasing complexity, starting from graphical modeling up to real source code implementations:

- *processes*: At the first stage the models are arranged using so called processes. Those represent the models' layers according to the OSI model, which in our case results in three different layers as shown in Fig.5.

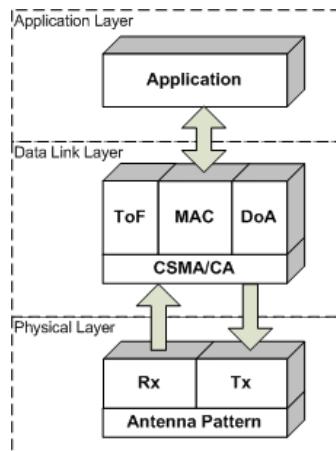


Fig. 5. Process view of the OPNET models

The bottom layer defines the characteristics of the three physical channels. Parameters such as the available data rate, the center frequency or the modulation scheme can be applied here.

The Data Link Layer (DLL) splits up into four sub processes to get a structural division for the tasks of the medium access control (MAC). Thereby the MAC process acts as a dispatcher that observes the different phases of the protocol and passes control to the specified process according to the active phase. The management phases are handled by the MAC process itself.

The CSMA/CA process coordinates the channel access for the participating nodes during phases that do not use time slots (e.g. management phases).

The process of the application layer differs for VRUs, which are equipped with a so called SafeTAGs and vehicles, which have an On-Board Unit (OBU). In the simulation model for the SafeTAG this process feeds the underlying MAC process with dummy data representing the future sensor values that are transmitted to the connected vehicles within DoA phases.

The application process for the vehicle includes a simplified calculation for the connected VRUs endangerment levels depending on the received measure values and sensor data.

- *state machines:* The processes consist of graphically modeled state machines. Fig. 6 shows the state machine of the modeled MAC process.

The management state handles the management phases, e.g. to establish connections to newly appearing network nodes. As soon as one of the other phases of the protocol is activated, the state machine switches to the specified channel and passes control to the corresponding state, which is implemented as separate sub state machine.

Since a time slotted channel access was defined for the ToF and DoA phases the specified sub state machines collect the measure values and data for the connected

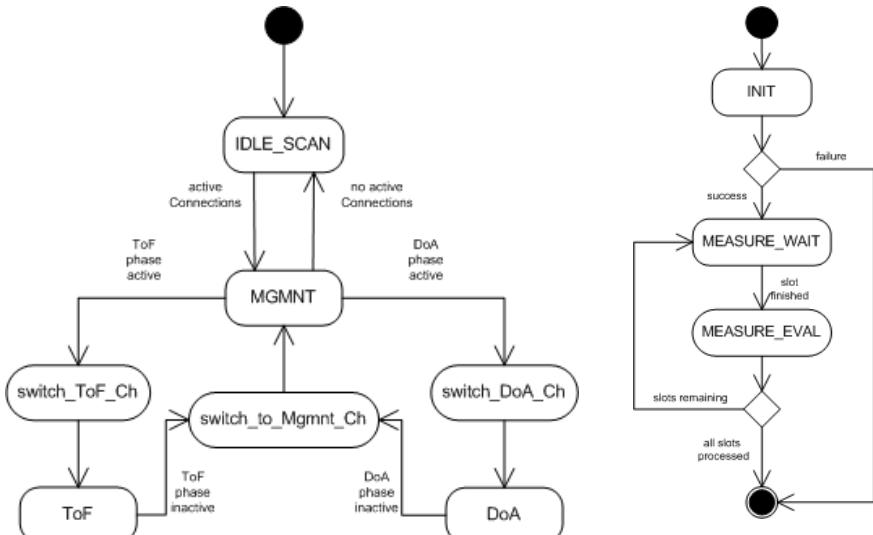


Fig. 6. State machine of the model's MAC process

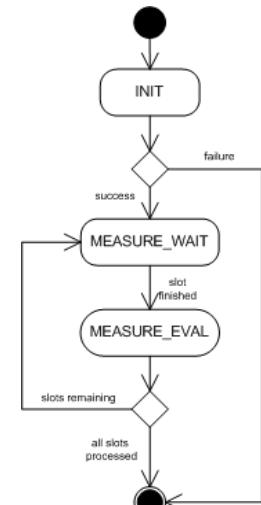


Fig. 7. Measuring state machine

devices in their exclusively assigned time slots (vehicle model) or vice versa send their replies with data and measurement results (SafeTAG model).

In its initial state the vehicle's measurement sub state machine tries to send a beacon frame to identify itself to the listening VRU nodes. After the successful identification the VRUs will send their measurement data in their assigned time slots.

The vehicle will receive and evaluate the replies slot by slot within its MEASURE_* states until the complete super frame was processed. Finally the sub state machine will be left and the control gets back to the main state machine.

- *implementation:* Finally the real functionality of the single states is implemented as C/C++ source code to describe the detailed behavior of the network nodes.

- *parameters & statistics:* Additionally every model includes a specified set of parameters that can be modified for different simulation runs. Combined with the collection of statistical values the best configuration of the protocol can be evaluated this way.

Another advantage is the given debug mechanisms that provides a much more comfortable and faster development process than on the real target hardware. A further benefit of this approach is the high portability to other systems (e.g. microprocessors) since the use of an additional abstraction layer allows to reuse most parts of the models' source code.

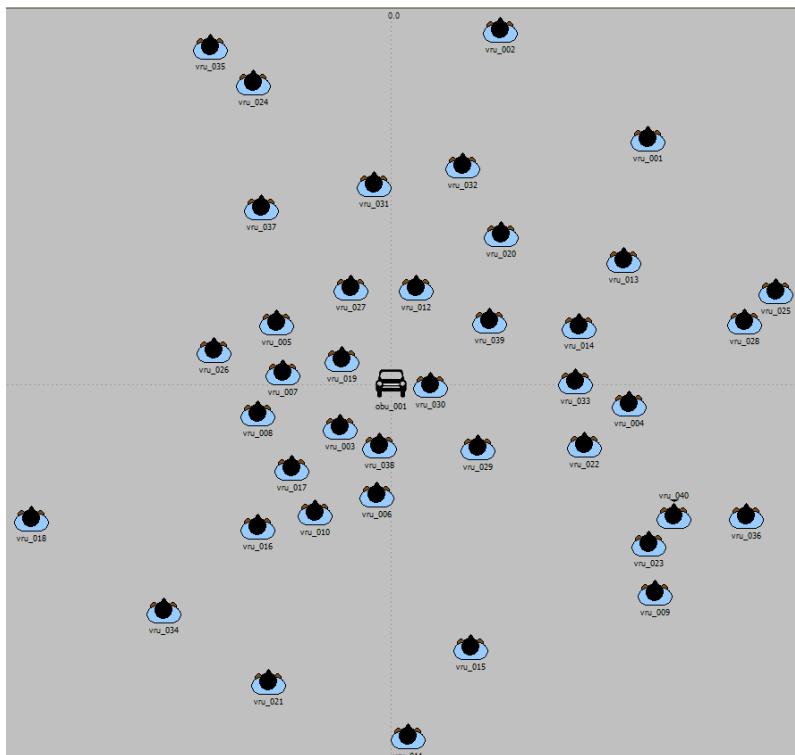


Fig. 8. Scenario with one vehicle and forty VRUs in OPNET model

4.3 Scenario

In the first step, static scenarios with one vehicle and multiple VRUs are considered. The vehicle is placed in the centre position (0, 0) of the scenario. The VRUs are placed within 150 m from the OBU. The distances and angles between an OBU and a VRU are following a uniform distribution (cf. Fig. 8).

The parameters for these simulation runs are shown in tables 2, 3, and 4. These parameters were selected as initial values in an iterative approach and were derived from some macroscopic estimation. However, the results of the simulations show that the initial guess was quite realistic.

Table 2. Basic Simulation Parameters

Parameter	On-Board Unit		VRU Unit	
	Control Channel	ToF Channel	Control Channel	Unit
Inter frame spacing (IFS)		2	2	2 μ s
Decrement of the contention window		1	1	1 μ s
Maximum contention window size		15	20	15 μ s
Minimum contention window size		2	2	2 μ s

Table 3. Basic Simulation Parameters

Parameter		Unit
Number of TOF slots per super frame	50, 75, 100, 125	-
Duration of a TOF slot including the guard time	12 μ s	
Duration of a AOA slot including the guard time	75 μ s	
Data rate of the CTRL CH	6 Mbps	
Data rate of the TOF CH	30 Mbps	
Data rate of the DATA CH	6 Mbps	

Table 4. Derived parameters for the OBUs and the VRUs

TOF slots/ super frame	Duration of super frame (μ s)	AOA slots/ super frame	TOF beacon duration = size/data rate (μ s)	AOA beacon duration = size/data rate (μ s)	TOF slots / beacon
50	$50 \times 12 = 600$	$600/75 = 8$	$128/30 = 4.3$	$447/6 = 75$	$75/12 = 7$
75	$75 \times 12 = 900$	$700/75 = 12$	$128/30 = 4.3$	$479/6 = 80$	$80/12 = 7$
100	$100 \times 12 = 1200$	$1200/75 = 16$	$128/30 = 4.3$	$511/6 = 85$	$85/12 = 8$
125	$125 \times 12 = 1500$	$1500/75 = 20$	$128/30 = 4.3$	$543/6 = 91$	$91/12 = 8$

4.4 Results

For the statistics shown in fig. 9, the number of connected VRUs with a Vehicle was measured in dependence of the length of a network management phase and for a specified number ToF slots per ToF measurement phase. In each group, the left bar represents the number of VRUs in a scenario, the other four bars from left to right

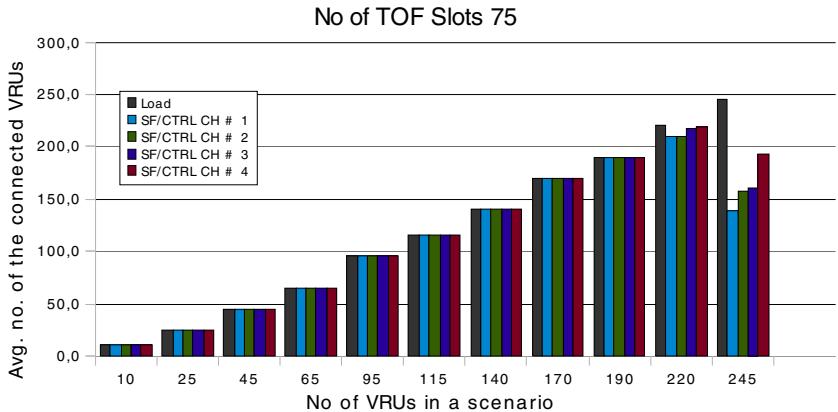


Fig. 9. Number of Connected VRUs in dependence of Number of VRUs in the scenario and the number of control channels per super frame considering 75 ToF slots per superframe

represent the number of connected VRUs. The measurements have been collected selecting the length of the management phases for the different runs as multiples from 1 to 4 of the length of an ToF and DoA super frame.

For a new connection establishment the underlying protocol defines the exchange of three packets. The SafeTAG discloses itself using an announcement (ANN) awaiting a following ASSOCIATION (ASS) response from a vehicle. A terminal acknowledge (ACK) from the device ends the procedure. Since the times to transmit the single packets add up to 61 µs for an ANN, 72 µs for an ASS and 70 µs for an ACK the minimum time needed to establish a connection totals to 203 µs excluding the random offsets caused by the CSMA/CA algorithm.

By using different values for the number of ToF time slots in a ToF super frame, and for the duration of a management phase in relation to the duration of the ToF phase we can vary the time window that allows connection establishments (c.f. tables 3 and 4). Therefore the results of Fig. 9 confirm that a longer management phase leads to an increasing number of successfully connected SafeTAGs at the end of the simulation run. This effect is brought out especially with an increasing number of nodes followed by an increasing rate of packet collisions.

5 Outlook

As soon as further simulation results will be available, the protocol parameters will be dimensioned. A detailed specification will be elaborated and integrated into the sensor system of a vehicle.

Acknowledgement

This work results from the joint project Ko-TAG, which is part of the project initiative Ko-FAS, and is partially funded by the German Ministry of Economics and Technologies (BMWi) under contract number 19S9011. The usage of OPNET network simulator was made possible through the support of OPNET Technologies, Inc. The authors are grateful for this support.

References

1. European Commission, Towards a European road safety area: policy orientations on road safety 2011-2020, SEC(2010) 903, (20.7.2010),
http://ec.europa.eu/transport/road_safety/pdf/com_20072010_en.pdf
2. Car2Car Communication Consortium, Manifesto Overview of the C2C-CC System (2007)
3. <http://www.kofas.de/>
4. U.S. Department of Transportation, IEEE 1609 - Family of Standards for Wireless Access in Vehicular Environments (WAVE), Intelligent Transportation Systems Standards Fact Sheet (September 25, 2009),
http://www.standards.its.dot.gov/fact_sheet.asp?f=80
5. Rasshofer, R., Schwarz, D., Gresser, K., Biebl, E.M.: Fußgängerschutz mittels kooperativer Sensorik. In: 5. Workshop FAS, Walting (2008)
6. Andreone, L., Visintainer, F.: Prevention of road accidents involving Vulnerable Road Users: the outcomes of the WATCH-OVER European project. In: ITS 2009, Stockholm (2009)
7. Harmonized European Standard (Telecommunication series), ETSI EN 302 571 V1.1.1: Intelligent Transport Systems (ITS); Radiocommunications equipment operating in the 5 855 MHz to 5 925 MHz frequency band; Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive,
http://pda.etsi.org/exchangefolder/en_302571v010101p.pdf
8. Lill, D., Schappacher, M., Gutjahr, A., Sikora, A.: Development of a Wireless Communication and Localization System for VRU eSafety. In: 2nd Int'l Workshop on Communication Technologies for Vehicles (Nets4Cars 2010) at 7th IEEE CSNDSP 2010, Newcastle, UK, July 21-23 (2010)
9. Karl, H., Willig, A.: Protocols and Architectures for Wireless Sensor Networks. J. Wiley & Sons, Chichester (2007)
10. Tanenbaum, A.S.: Computer Networks. Pearson Education, London (2002)
11. Jung, E.-S., Vaidya, N.H.: An energy efficient MAC protocol for wireless LANs. In: Proc. of the IEEE INFOCOM (June 2002)
12. Ye, W., Heidemann, J., Estrin, D.: An energy-efficient MAC protocol for wireless sensor networks. In: INFOCOM (2002)
13. van Dam, T., Langendoen, K.: An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks. In: Proc. SenSys 2003, Los Angeles (2003)
14. El-Hoiydi, A., Decotignie, J.-D.: WiseMAC: An Ultra Low Power MAC Protocol for the Downlink of Infrastructure Wireless Sensor Networks. In: Proc. ISCC 2004, Alexandria, Egypt (June 2004)
15. Sadler, B.M., Swami, A.: Synchronization in Sensor Networks: an Overview. In: MILCOM (2006)
16. Schindelhauer, C.: Lecture Notes: Wireless Sensor Networks. University of Freiburg (Winter 2006/2007)
17. Sikora, A.: MAC-layer design for co-operative and infrastructureless Wireless Sensor Networks for Relative Localization. In: GMA/ITG-Fachtagung Sensoren und Messsysteme 2008, Ludwigsburg, März 11./12. (2008)

A Vehicular Mobility Model Based on Real Traffic Counting Data

Yoann Pigne¹, Grégoire Danoy², and Pascal Bouvry²

¹ SnT, University of Luxembourg,

6 Rue Richard Coudehnove-Kalergi, L-1359 Luxembourg

yoann.pigne@uni.lu

² ILIAS, University of Luxembourg,

6 Rue Richard Coudehnove-Kalergi, L-1359 Luxembourg

{gregoire.danoy,pascal.bouvry}@uni.lu

Abstract. This paper proposes VehILux, a new vehicular mobility model based on real traffic counting data. It relies on two freely available sources of real information for the country of Luxembourg. The first source is traffic data collected by counting devices located on the Luxembourgian road network, while the second is geographical information about different types of areas: residential, industrial, commercial and other services. VehILux models vehicles commuting around the city of Luxembourg by considering two types of traffic, outer traffic with vehicles entering in the defined geographical area and inner traffic starting from residential zones located inside the geographical area. One part of the collected traffic data is used as input traffic, while another part is used to control the produced traffic and to fine-tune the model. VehILux is coupled with the microscopic road traffic simulator SUMO to produce realistic vehicular traces.

Keywords: Mobility Models, Traffic Management, Vehicular Ad Hoc Networks, Simulations, Intelligent Transportation Systems.

1 Introduction

With the increasing vehicles traffic and its environmental impact, road safety and traffic management has become one of the biggest concerns of most national and regional transport departments. Solutions have emerged with the recent vehicular wireless communication capabilities, which constitute a key component of Intelligent Transportation Systems (ITS). Such Vehicular Ad-hoc Networks (VANETs) have rapidly emerged and raised novel research challenges such as networking protocols adapted to high speed environments.

Due to logistic, environmental and economical reasons, simulation is a must in the validation process of any of such solution. However, the relevance of a simulation platform depends on its ability to accurately consider the road environment and the vehicles motion. This implies the development of realistic models for the signal physical characteristics (attenuation, reflection, diffraction, scattering

[22]), the road network (speed limits, number of lanes, traffic light logic, etc.), the mobility at a macroscopic level (flows of vehicles, planned routes) or at a microscopic level (acceleration/deceleration, car following each other, slow down anticipations [11]).

The simulation of the mobility of vehicles can be achieved with many different models. Some models, almost uncorrelated with real movements, yet theoretical, carry some properties that can be reproduced. Such models can be scaled to produce traffic of any volume. On the opposite mobility models made of real mobility pattern are realistic but they are also bound to the size of their dataset of traces and usually cannot be scaled and produce simulations for a larger number of entities.

In this paper we propose a mobility model that is both realistic (based on real data) and able to produce routes for a wide area (the order of a city, or a region). For this we rely on two kinds of real road data freely available for the country of Luxembourg. The first one consists of the vehicles flows provided by traffic counting devices. This data is provided by the Luxembourg Ministry of Transport [4]. The second source of information is the OpenStreetMap project [6] that provides detailed information about the road network and the zones of activity in the country. One of the main contributions of this work is thus the possibility to produce routes out of a set of flow information and localized land use information. We then rely on a microscopic traffic simulation tool, SUMO [17], to generate realistic traffic that can then be used by other tools such as VANET simulators [20][21]. Another important point is that a part of the input data is only used to measure the accuracy of the generated traffic.

This paper is organized as follows. Next section draws a state-of-the-art of mobility models available for road purposes. In Sect. 3 after explaining details about the nature of the data set used, we show how the proposed model creates routes from traffic counting information and how destinations of created routes are computed thanks to the input data. Section 4 explains details of the Luxembourgian setup and analyses the produced experimental results. Finally, Sect. 5 concludes the paper.

2 State of the Art

The key influence of mobility models on mobile ad hoc networks connectivity results has been emphasized in many works. The necessity of realistic vehicles motion in VANETs is even more crucial as it implies a more constrained environment (road network, vehicular networks research has already produced a wide range of mobility models, typically classified as macro-mobility models (vehicular density, velocity) and micro-mobility models (acceleration/deceleration, lane change, etc.) [15] for which an exhaustive state-of-the-art can be found in [13]. The contribution proposed in this paper lies in this first category. Mobility models classification can be refined by using the categories proposed by Härry in [12], which includes random models, flow models, traffic models, behavioral models and trace-based models.

Random models have been adapted to vehicular mobility, such as the Manhattan model [9], but still remain too simplistic. Flow models, are mathematical models based on flow theory which have been developed by traffic engineers to reproduce vehicles particular movements through interactions between themselves and with the environment. However, one of the main drawbacks of flow models is their incapacity to take into account the human behavior, i.e. the influence of other factors such as perturbations, external stimuli or social parameters. As an answer, behavioral models, based on behavioral theory, are a recent and promising answer to the lack of human influence on vehicles mobility [18].

Our contribution lies at the border of the two remaining models classes, i.e. traffic and trace-based, which we describe in more details in the following.

As mentioned by Härri, traffic models can be subdivided in two motion patterns, trip planning, which first defines origin and destinations, and path planning which then fixes the sequence of directions to be taken by each vehicle. Two random trip planning approaches exist, random trip which randomly chooses an origin and a destination, and stochastic turn which chooses a random direction at each intersection, which are the most commonly implemented in existing vehicular mobility models. The last trip planning approach brings more realism by defining a set of departure and destination points, also referred to as *attraction points*, either user-defined as in NCTUNS [23] or based on a sequence of activities (e.g. hotel, restaurant, etc.) extracted from realistic maps, as in the mobility traces generator CanuMobiSim [2], its extension VanetMobiSim [14] and the traffic simulator SUMO [17]. The path computation is then ensured between those attraction points using in most cases a Dijkstra-type graph algorithm, to find the shortest path, the fastest path or the less crowded path.

Trace-based and survey-based models are an alternative to complex mobility models development and calibration using real traffic data. They instead allow to define generic mobility patterns from real vehicular traces. The University of Delaware (UDel) mobility model [7] is based on time-used studies by different research areas (US Bureau of Labor and Statistics, business research communities, etc.) from which daily vehicles dynamics are derived. Another example is the computationally demanding Multi-agent Microscopic Traffic Model (MMTS) [19] from ETH Zurich, which generates highly realistic traffic for a 24 hours period. Finer-grained information can be obtained from traces obtained by GPS tracking of vehicles, from which models can be extrapolated. However, such traces have a limited availability and are limited to the type of tracked vehicles, like mobility models based on taxis in San Francisco Bay area [11] or Shanghai [16], campus vehicles in ETH Zurich [8] or buses in Chicago [10].

The mobility model proposed in this paper combines both traffic and survey-based advantages, by respectively defining routes destinations by exploiting geographical information from open-source maps and by using real traffic counting data to define and assess the validity of the generated vehicles mobility. The next section provides a detailed description of the proposed mobility model.

3 Model

This paper introduces a novel realistic vehicular macro-mobility model which allows the simulation of vehicles flows at a city or region scale by combining features from survey and traffic based mobility models.

On the one hand it uses information gathered from traffic counting devices (induction loops) that provide precise information about vehicles flows at precise geographical points. Although a lot of information is captured in terms of traffic volume, no information about the destinations of vehicles is captured. Therefore, on the other hand, a path planning heuristic is also included such that the selection of destinations is not random but according to the time-dependent attractiveness of geographical zones. As an example, destinations at commuting hours are certainly more concentrated on work places than on residential places. Moreover limiting the source of traffic to the counting loops positions would not be realistic either. As a solution, this mobility model includes a comprehensive and tunable model for the selection of destinations based on geographical attraction points (work places, home places, etc.) extracted from real data maps. Additional traffic is also generated from other sources than from the simple counting loops. Finally we keep some of the flow information for comparison purpose. The volume of traffic we produce is compared with some of the real counting data when available.

This section is organized as follows. After explaining the detailed nature of the input data in Sect. 3.1, a description of the destinations distribution model used to generate routes is given Sect. 3.2. Finally, Sect. 3.3 explains our model for creating end to end routes from these flows and destinations.

3.1 Input Data

In order to generate realistic routes we rely on two sources of information. The first one, maps from OpenStreetMap, provides detailed road network and land use information. The second source, vehicle counting data based on induction loops, is given by the Luxembourg Ministry of Transport.

Detailed Maps. OpenStreetMap (OSM)⁶ provides collaboratively constructed maps of the world made available online and copyrighted under a Creative Commons³ license. Two sources of information are gathered¹, first on road network (road type, speed limit on each lane, traffic lights, etc.) and second on the classification of geographical zones by type. Four different types are proposed, *landuse* for general activities zones (industrial, commercial, etc.), *shop* refining commercial zones (clothes, florist, etc.), *amenity* for first necessity zones (pharmacy, banks, etc.) and finally *building* or *office* for affecting an activity to a given building/office.

For the purpose of our proposed mobility model, we used a coarser-grained classification with 3 main types of activities that are: (1) *commercial* which

¹ Map data ©OpenStreetMap contributors, CC-BY-SA.



Fig. 1. On the Left, an original map from OpenStreetMap (©OpenStreetMap contributors, CC-BY-SA), on the right the resulting gathered information with 3 types of zones (commercial, industrial,residential)

gathers landuse commercial and retail zones, plus the shop and amenity; (2) *industrial* and (3) *residential* corresponding to the respective landuse zones.

Finally, the surface of each zone is also computed as it will be a parameter of its attractiveness. Figure 1 illustrates the 3 types of zones gathered from OSM data of the Kirchberg area of the city of Luxembourg. Residential zones are represented in light blue, commercial zones in purple and industrial in black.

Traffic Counting Data. The Luxembourg Ministry of Transport through the *Ponts et Chaussées* administration carries on a systematical and automatic counting of the vehicle traffic through a set of counting devices (induction loops and radars) installed all over the country, on main roads and motorways. As of 2010, traffic is counted in both directions at 126 different locations. The resulting data is made available on a dedicated website [4]. Results can be retrieved on different time scales (hour, day, week, month, year), for different date ranges. A difference can be made between truck and car, and between work days and weekends. Means are computed if needed. Figure 2 is an example of data retrieved of the *Ponts et Chaussées* administration's website providing hourly traffic for trucks (in blue) and cars (in red) on a 24 hours time slot.

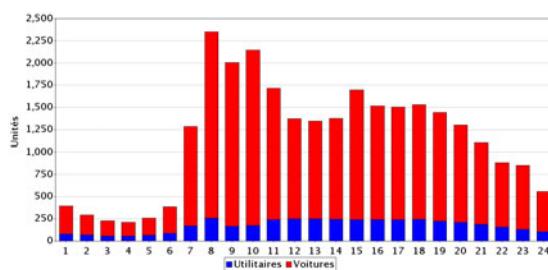


Fig. 2. Example output graphic produced by the *Ponts et Chaussées* (©Administration des Ponts et Chaussées, Division Informatique et Gestion)

3.2 Destination Distribution Model

After defining the two sources of data exploited by our mobility model, we now describes how destinations are chosen. As mentioned in Sect. 3.1, geographical zones' *surface* and *type* (i.e., commercial, industrial or residential) are extracted from OSM data. Based on this information, the probability for choosing a zone is influenced by three parameters:

- the weight of its zone type,
- the weight of its attractivity area (or 1 if it does not belong to such an area),
- its own surface.

The attractiveness of a general zone type can first be adjusted by using a *weight* factor. Additionally, any given zone attraction is also proportional to its surface. However the attractiveness of a zone may not only depend on its surface and zone type anywhere in the simulation area. For instance two commercial zones of equal surface may not be equally attractive if one of them is located in a big city center and the other is located on a country side area with low population-density. For this reason we define some extra geo-localized weight that apply to the attractiveness of geo-locally close zones. These geo-localized weights are called *attractivity areas*, which are disks with a given position, radius and weight (the weight is an integer value). A zone belonging to the radius is given the weight of its attractivity area. A zone outside of any area is applied a unitary weight. Attractivity areas are define by hand or tuned to fit some defined function of objective. This question is discussed in Sect. 3.3.

Those 3 parameters can be seen as 3 events in a probability space (Ω, \mathcal{A}, P) : A “choose a type”, B “choose an attractivity area”, and C “choose a zone”. Those 3 events are not independent one from the other. For instance, the probability that B occurs ($P(B)$) will be different if we know that A already occurred: an attractive area is linked to a zone type, so if a zone type is already chosen (event A) then the probability of choosing an attractivity area ($P(B)$) is influenced. For the 3 events A , B and C , the probability that C occurs given events A and B already occurred is the *conditional probability* of C : $P(C|B \cap A) = P(C|(A \cap B))P(B|A)P(A)$.

Let us consider $A_i, i \in I$ and $B_j, j \in J$ being 2 partitions of Ω , and each event A_i and B_j having a probability $P(A_i) \neq 0$ and $P(B_j) \neq 0$, then for any event C of the same probability space, the *law of total probability* states that

$$P(C) = \sum_{i \in I} \sum_{j \in J} P(C \cap B_j \cap A_i) = \sum_{i \in I} \sum_{j \in J} P(C|B_j \cap A_i)P(B_j|A_i)P(A_i). \quad (1)$$

Back to our problem, the event C ”choose a zone” is constituted of only one element of ω , it is an elementary event. By definition of what a zone is, each of the solutions of C (each zone) can only be associated with one and only one couple (zone type, attractive area). Thus, given one element of C , only one couple (i,j) in Π can lead to a product $P(C|B_j \cap A_i)P(B_j|A_i)P(A_i) \neq 0$. The equation is then simplified and is equal to the *conditional probability* of C knowing B and A :

$$P(C) = P(C \cap B \cap A) = P(C|(A \cap B))P(B|A)P(A). \quad (2)$$

$P(C)$ is the product of the 3 following probabilities:

- $P(A)$ is the probability of choosing a zone type which is computed from the normalized weight of each zone type. $zt_i, i \in I = |zt|$ ($|zt|$ being the cardinal of zt), is a zone among all the I possible zone types.

$$P(zt_i) = \frac{weight(zt_i)}{\sum_{j \in I} weight(zt_j)}. \quad (3)$$

- $P(B|A)$ is the probability of choosing an attractivity area, knowing the chosen zone type, is computed according to the weight given to each attractivity area of the same zone type. $aa_{i,j}, j \in J = |aa_i| + 1$, is an attractive area chosen among all $J - 1$ possible areas (plus the possibility of choosing no area: $\overline{aa_i} = aa_{i,J}$) knowing zt_i . Note that the weight of $aa_{i,J}$ is 1.

$$P(aa_{i,j}) = \frac{weight(aa_{i,j})}{\sum_{k \in J} weight(aa_{i,k})}. \quad (4)$$

- $P(C|(A \cap B))$ is the probability of choosing a zone according to its surface knowing a given attractivity area and thus a given zone type. $z_{i,j,k}, k \in K = |z_{i,j}|$, is a zone chosen among all K possible zones that belong to $aa_{i,j}$ and zt_i .

$$P(z_{i,j,k}) = \frac{surface(z_{i,j,k})}{\sum_{l \in K} surface(z_{i,j,l})}. \quad (5)$$

3.3 From Flows to Routes

The work presented in this paper focuses on the simulation of vehicles mobility during commuting hours in the city of Luxembourg and surroundings, when a huge part of the daily traffic appears, in direction to the city. Therefore, we only focus on the traffic entering in this area and not on the traffic leaving this area.

Information about the source of traffic can be extracted from the traffic counting data, however only considering vehicles entering in the simulation area from a few geographical points is not realistic. Based on this hypothesis, we classify the traffic entering the simulation area in two classes based on its origin: traffic originated outside the area (*outer traffic*) and traffic originated inside the area (*inner traffic*). Both traffic models use a route planning model between a selected source and destination based on the shortest paths in time, computed thanks to OSM data. We also provide a way to control and adjust the produced traffic according to the real flow data available.

Outer traffic. As for any survey-based mobility model, the objective is to generate end to end routes based on traffic counting data. To this end a traffic model is created, which uses this data as geographical sources of traffic and chooses destinations in the considered simulation area using the model defined in [B.2](#). Since these flows are directed we only select flows toward Luxembourg city. Figure [B.3](#) illustrates the considered simulation area, where black arrows represent the position of the real traffic counting devices. Those devices located around the city of Luxembourg are considered as the *outer traffic* input sources.

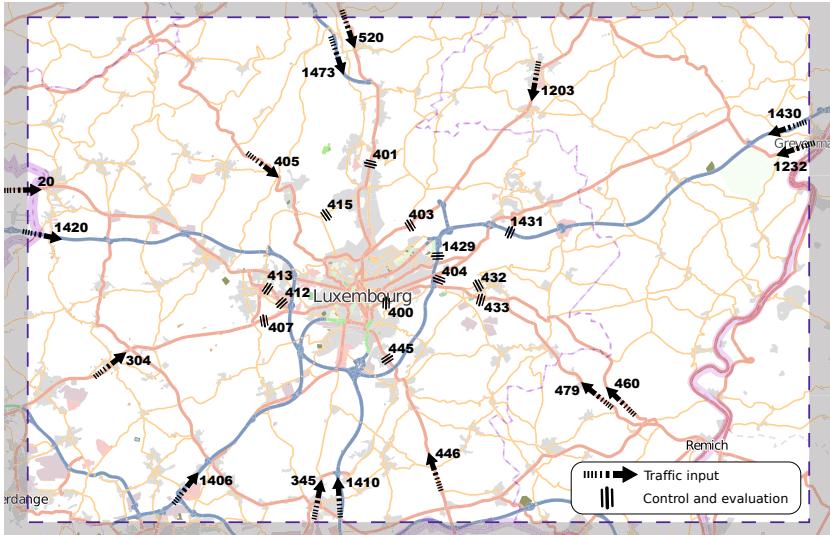


Fig. 3. Simulation area and position of traffic counting devices (named with the classification given by the Ministry of Transport)

Inner traffic. This complementary traffic is originated from inside the simulation area. No real information is available about such sources of traffic, however, in such a commuting traffic simulation, most of the traffic will originate from residential zones (from OSM data). We propose to generate this inner traffic by randomly choosing sources among the residential zones and destinations using the described model. In addition, no information is available about the corresponding flows either. However, since the evolution of the outer traffic is known, one simple assumption is that the inner traffic will be a function of it. We propose to define the overall hour by hour volume of inner traffic proportionally to the outer traffic, according to a defined ratio. This ratio is latter referred to as the *inner traffic ratio*.

Traffic control. As it can be observed Fig. [B.3](#), some traffic counting data is also available inside the considered area. The real data collected in these locations

is first used to compare the generated traffic with those real information. In a second step, the data can be used to adjust the parameters that rule the model:

- The *inner traffic ratio* that gives the hour by hour amount of traffic generated inside the area proportionally to the outer traffic, is adjusted experimentally to fit real values.
- The weight value of the three geographical *zones types*.
- The *Attractivity areas* defined in the destination model (Sect. 3.2) that set the importance of designated areas of the simulation space. Although it might sound obvious that the city center is more attractive than other low density areas, these zones have to be manually set.

The parameter tuning of this model could be optimized by using the real traffic data as a reference global objective. This part is beyond the scope of this paper but belongs to the perspectives considered by the authors.

4 Experiments

The mobility model proposed in this paper has been developed and fine tuned on a Luxembourgian testbed, using real geographical and traffic counting data. In this section, we first provide details on the Luxembourgian testbed, followed by a presentation of the micro-simulation tool used, i.e. SUMO. Finally experimental results are presented and discussed.

4.1 The Luxembourgian Testbed

The simulation area around the city of Luxembourg is selected in order to maximize the number of information about traffic counting. Figure 2 shows this 1700 square kilometers area (47×36 kilometers). 28 traffic counting devices are considered, 15 for input traffic and 13 for comparisons.

According to the defined 3 zone types, this area has 304 residential, 772 commercial, 47 industrial zones of different surfaces. Table 1 gives more detailed information about these zones.

Table 1. Zones of interest in the selected Luxembourg scenario

Zone type	#zones	Σ surface (km^2)	Mean surface(m^2)
Commercial	772	8.82	11425
Industrial	47	14.20	302215
Residential	304	82.58	271645

To reflect the attractiveness of these areas during commuting hours, we experimentally fixed the relative weight of commercial, industrial and residential zones to respectively 80%, 15%, and 5%.

A set of *attractivity areas* is experimentally proposed as an attempt to model the real attractiveness of specific zones. The latter comprise three commercial areas, one over the city center, one over the financial district, and one over a set of large mall centers respectively weighted 10, 15, and 5. Another industrial area is weighted 10 and a residential one is weighted 5, both centered around the city center.

For this testbed, values for *attractivity areas* and the attractiveness of zones are only proposed by the authors and are not based on any real data. However this framework allows to refer to real data. Future works will consider the usage of data about densities of populations and other zones of activities in Luxembourg gathered in projects like MOBILLUX [5].

The traffic generation, is set to produce an equal volume of inner traffic and outer traffic (the *inner traffic ratio* is set to 50%).

Finally the output of the mobility model is a set of vehicles with their route given as a list of edges from the OSM map. This output is ready to be used as an input for micro-simulation.

4.2 Microsimulation with SUMO

The micro-simulation is achieved by the SUMO [17] traffic simulator. It handles the micro-mobility level, including road interactions such as car following models, traffic light logic, or overtaking models. It is an open source project, actively developed by a team from DLR (the German space agency).

For the set up of this scenario, OSM maps are converted into SUMO's network format. The position of real traffic counting devices is used as measuring points so as the simulated traffic is counted at the same position the real devices are.

4.3 Results Analysis

The simulation runs from midnight to 11am. During this period approximatively 150,000 vehicles are launched in the simulator. After the execution, the experimentally obtained traffic counting values can be compared to the real ones. Since our mobility model constructs routes made of lists of edges, we can also count the number of vehicles passing by edges having a traffic counting device. As a consequence three values can be shown in the results, i.e. real data, mobility-model data and micro-simulation data. Figure 4 provides charts for those three values obtained by each counting devices available in the simulation area (see Fig. 3). For a visibility purpose, in the figure, only 3 out of 13 available data charts are displayed. The evolution of the traffic in flows of vehicles is shown, hour by hour, from midnight to 11 in the morning.

By analyzing the results, it clearly appears that for devices 445, 415 and 432 (only 445 is shown of Fig. 4) that the generated traffic is almost adequate to the real data. On some other roads the model generates too much traffic in comparison to real data, just like on devices 1431, 1429, 433 and 400 (only device 1429 is shown on Fig. 4). On many others like 404, 407, 401, 403, 420

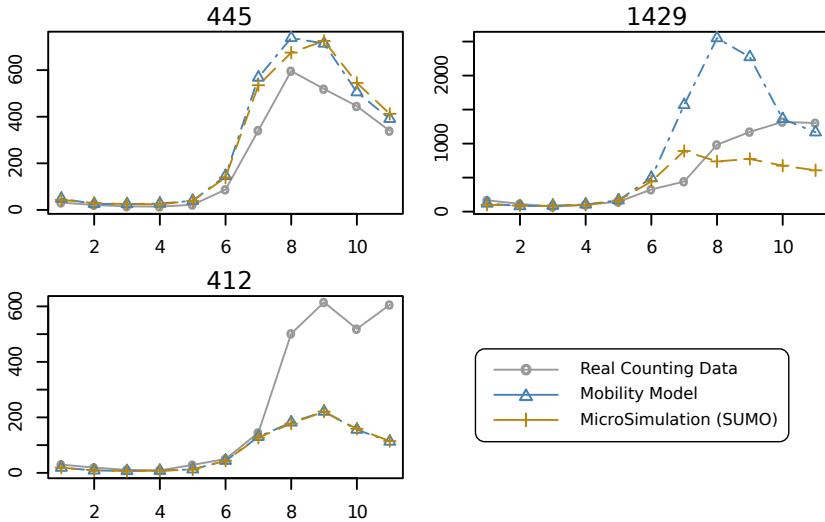


Fig. 4. Traffic counts for real data, modeled mobility and simulation. The y-axes is the number of vehicles per hour, the x-axes is the time of the day in hour.

and 412 (only 412 is given on Fig. 4), real data shows more traffic than the generated one. This may reflect the need for optimizing the zone type weights and the attractivity areas, but it may also be a consequence of using shortest paths in terms of journey duration which is a theoretical value computed with the maximum allowed speed. Volumes of traffic from the mobility model are almost always above the volumes observed in SUMO. This is normal since the mobility model is the objective traffic SUMO is trying to generate. On very high density traffic roads (device 1429, 1431 and 400) the simulator acts differently compared to the model, as if it could not generate such traffic.

5 Conclusion

In this paper we have introduced a novel vehicular macro-mobility model, using both survey-based and traffic-based models' features. Relying on SUMO for microscopic traffic simulation, it provides a tool that generates vehicular mobility traces based on real traffic counting data. A parametrized destination model based on real data from OpenStreetMap allows to tune the model. A set of traffic counting data is used to compare the generated traffic and offers a framework for optimizing the parameters of the model. This model and its output may be used as is in other projects in the fields of VANET simulation or for other traffic management simulation tools. Our future work focuses on proposing optimization schemes for the tuning of the mobility model to fit more adequately the real data.

References

1. The cabspotting project, <http://cabspotting.org>
2. The canu project, <http://canu.informatik.uni-stuttgart.de/mobisim/>
3. Creative commons, <http://creativecommons.org/licenses/by-sa/2.0/>
4. Luxembourg ministry of transport - traffic data, <http://www.pch.public.lu/trafic/comptage/>
5. Mobillux, <http://mobilluxweb.ceps.lu/>
6. Openstreetmap, <http://www.openstreetmap.org/>
7. Udel mobility model, <http://www.udelmodels.eecis.udel.edu/>
8. A mobility model based on WLAN traces and its validation, vol. 1 (2005)
9. Bai, F., Sadagopan, N., Helmy, A.: Important: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks. In: INFOCOM (2003)
10. Doering, M., Pögel, T., Pöttner, W.B., Wolf, L.: A new mobility trace for realistic large-scale simulation of bus-based dtns. In: Proceedings of the 5th ACM workshop on Challenged networks, CHANTS 2010, pp. 71–74. ACM, New York (2010)
11. Gipps, P.: A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological 15(2), 105–111 (1981)
12. Härri, J.: Vehicular Mobility Modeling for VANET, pp. 107–156 (2009)
13. Härri, J., Filali, F., Bonnet, C.: Mobility models for vehicular ad hoc networks: a survey and taxonomy. IEEE Communications Surveys and Tutorials 11(4) (2009)
14. Härri, J., Fiore, M., Filali, F., Bonnet, C.: Demo: simulating realistic mobility patterns for vehicular networks with vanetmobisim. In: WiVeC 2007: 1st IEEE International Symposium on Wireless Vehicular Communications (2007)
15. Helbing, D.: Traffic and related self-driven many-particle systems. Reviews of Modern Physics 73(4), 1067–1141 (2001)
16. Huang, H., Zhu, Y., Li, X., Li, M., Wu, M.Y.: Meta: A mobility model of metropolitan taxis extracted from gps traces. In: WCNC, pp. 1–6. IEEE, Los Alamitos (2010)
17. Krajzewicz, D., Bonert, M., Wagner, P.: The open source traffic simulation package SUMO. In: RoboCup 2006 Infrastructure Simulation Competition (2006)
18. Musolesi, M., Mascolo, C.: A community based mobility model for ad hoc network research. In: Proceedings of the 2nd International Workshop on Multi-hop ad hoc Networks, REALMAN 2006, pp. 31–38. ACM, New York (2006)
19. Naumov, V., Baumann, R., Gross, T.: An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces. In: Proceedings 7th ACM International Symposium on Mobile ad hoc Networking and Computing, pp. 108–119. ACM, New York (2006)
20. Pigné, Y., Danoy, G., Bouvry, P.: A platform for realistic online vehicular network management. In: MENS 2010 in conjunction with IEEE GLOBECOM 2010. IEEE Computer Society, Los Alamitos (2010)
21. Sommer, C., German, R., Dressler, F.: Bidirectionally coupled network and road traffic simulation for improved ivc analysis. IEEE Transactions on Mobile Computing 99(RapidPosts) (2010)
22. Tan, I., Bahai, A.: Physical Layer Considerations for Vehicular Communications. John Wiley & Sons, Ltd, Chichester (2009), <http://dx.doi.org/10.1002/9780470740637.ch6>
23. Wang, S.Y., Chou, C.L.: Nctuns tool for wireless vehicular communication network researches. Simulation Modelling Practice and Theory 17(7), 1211–1226 (2009)

Driver-Centric VANET Simulation

Pedro Gomes¹, Cristina Olaverri-Monreal¹, Michel Ferreira¹, and Luís Damas²

¹ Instituto de Telecomunicações, Departamento de Ciência de Computadores,
Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 1021/1055,
4169-007, Porto, Portugal

{prg,cristina.olaverri,michel}@dcc.fc.up.pt

² LIACC and Departamento de Ciência de Computadores,
Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 1021/1055,
4169-007, Porto, Portugal
luis@dcc.fc.up.pt

Abstract. Inter-vehicle communication is becoming increasingly relevant in the research and development of novel, innovative vehicular applications. To support the driver in his/her primary driving task in an effective non distracting way, these applications need to be evaluated in a realistic context from a driver's perspective of the VANET environment. In this paper we propose an innovative driver-centric simulation tool that integrates a VANET simulator with a driving simulator using communication technologies to relay information about the vehicle to the VANET environment and vice versa. The driver behavior is reflected in the VANET simulation system affecting the mobility of the cars in the vicinity and providing the intelligent driving model with new realistic features.

Keywords: Driving Simulation, VANET Simulation, VANET Applications, Driving Model.

1 Introduction

Existing VANET simulators provide a broad perspective of the network, aiming essentially connectivity studies and protocol design, disregarding a driver viewpoint of the VANET environment. In cases where e.g., accidents, road congestion or other type of information that can influence the driving performance, are transmitted over the VANET, both road traffic simulation and network traffic simulation concepts need to be combined [17]. Since the performance of field tests to evaluate innovative VANET applications could lead to dangerous traffic situations, it is not a viable test procedure in many cases. A simulated system like the one proposed in this paper can reduce these difficulties and provide the right framework to test and evaluate advanced VANET functions. Our approach bases on a novel driver-centric simulator that permits the evaluation of a number of VANET enabled Driver Assistance and Driver Information Systems. The driving is simulated in a VANET environment allowing the driver to interact with devices and applications that are communicating with the cars in his virtual

vicinity. Our simulator builds on top of the open source micro-traffic simulator DIVERT [8]. DIVERT represents single vehicles' mobility in a detailed map of a real city. Since it provides a number of different traffic situations such as high traffic density it allows the evaluation of novel ideas permitted in the inter-vehicle communication context in a very realistic manner. For our simulation we use the city of Porto, in Portugal. The driver-centric simulation tool can be considered an advance in the existing current evaluation tools since the driver is immersed in a VANET environment. It aims to fill the gap between existing VANET simulators, and driving simulation platforms, providing a driver-centric perspective of VANET-enabled applications and becoming a fundamental tool for the research, validation and evaluation of Advanced Driver Assistance and Information Systems.

The remainder of this paper is organized as follows. In the next section we revise related work in the area of VANET simulation tools. Section 3 presents a detailed description of the simulator architecture. Section 4 describes the mobility model of the simulator, Section 5 reports on applications that can be tested with the simulator. Finally, Section 6 concludes the paper.

2 Related Work

Since it is crucial to test and evaluate protocol implementations on a large-scale realistic environment, studies of vehicular communication protocols in the VANET context are typically based on simulation models [16]. While a variety of simulation tools have been developed to analyze transportation scenarios at the micro- and macro-scale levels and in the last years the number of tools for integration in traffic and network simulators has rapidly increased [15], little effort has been devoted to the integration of communication techniques and scenarios in a realistic transportation simulation environment [11]. In [17], the need for bidirectional coupling of network simulation and road traffic micro simulation was discussed and the hybrid simulation framework Veins (Vehicles in Network Simulation) was explained.

VANET research has also made available tools to facilitate the rapid generation of realistic mobility models for simulations [11] like in modular integrated traffic and network simulators [12] and microscopic traffic simulators on the real maps of cities like Porto [9]. In addition, other open-source simulation environments have been introduced for proper evaluation of protocols for Vehicular Ad Hoc Networks (VANET) [15] and new scenarios like online virtual worlds have been suggested to conduct driving simulations [4].

Research has shown that driving simulators are proven to be excellent practical tools to test Advanced Driver Assistance Systems (ADAS) or Driver Information Systems (DIS) [3]. There has been a lot of research and an equally large amount of efforts in modeling concepts and techniques for improving behavioral intelligence and realism in driving simulation scenarios. For example, the authors in [6] developed Neural Driver Agents to learn and successfully replicate human lane changing behavior based on data collected from a car simulator [7]. Other

techniques like artificial neural network (ANN) models have also been used in the development of simulators: data collected from highways in Saudi Arabia was used for estimating headways in vehicles [2]. These models are capable of learning from training examples and demonstrating learned behavior in imaginary situations. In spite of all the new features that have been incorporated to the vehicles in the last years and the high realism reached in the current simulators on the market, the interaction between the driver and the vehicle has not changed significantly.

Our Driver-Centric VANET simulation intends to improve this interaction reflecting the driver behavior in the VANET simulation system and relaying information about the vehicle operated by the human driver to the VANET environment and vice versa.

3 Architecture

The implementation of a realistic driver-centric simulation tool that allows the validation and evaluation of VANET applications demands the design of an interface for coupling a driving simulator with a VANET simulator. The proposed architecture uses a Transmission Control Protocol (TCP) connection to link the DIVERT simulator with the driving simulator.

The diagram in Figure 1 illustrates the coupling architecture between the large-scale VANET simulator DIVERT 2.0 and the driving simulator.

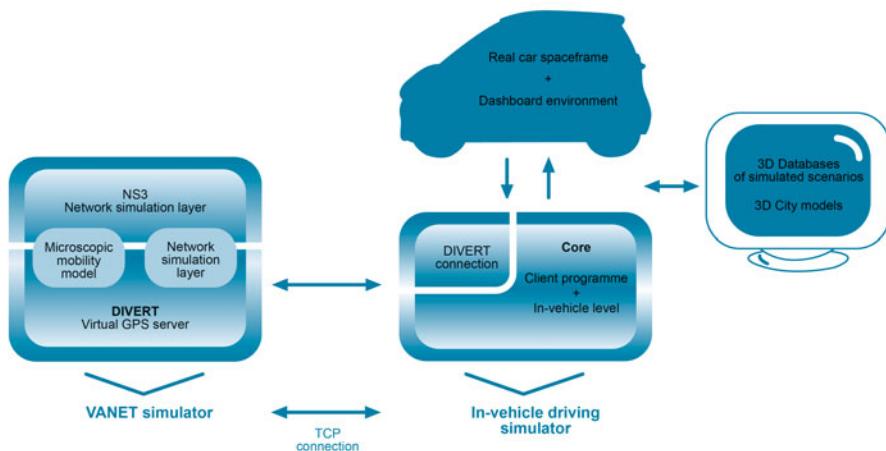


Fig. 1. Coupling architecture between Divert 2.0 and the driving simulator

3.1 Components

The simulator consists mainly of three components: the DIVERT simulator; an in-vehicle realistic driving simulator and an interface able to couple both.

DIVERT Simulator: DIVERT provides the variety of traffic scenarios in respect to the microscopic mobility characterization of simulated vehicles, together with a network simulation layer that models the wireless channel and ensures that the information is routed towards the location where it is most useful. In addition a TCP server allows the integration of simple modules to communicate with external applications. DIVERT includes a complex editor of traffic entities where road segments can be defined. It describes a detailed connectivity at intersections, traffic-lights interplay and several components that affect the vehicles' behavior (e.g. braking and acceleration patterns, etc). By using geographic information to partition the road network layer in asynchronous regions, DIVERT takes advantage of multi-threading and parallelism and defines an aimed number of vehicles in the city.

The microscopic realistic traffic simulator is capable of capturing the complexity and detail of the vehicle' trajectories that suit a wide scope of perspectives, based on models such as the Intelligent Driver Model (IDM) [18]. To enhance and refine the traffic flow modeling much more detailed properties are required, that are able to provide a realistic feeling for a human driver guiding a vehicle through the client program. Thus, the Driver-Centric VANET simulation tool aims to significantly enhance the mobility patterns of the individual vehicles from the DIVERT simulator by taking in consideration the unpredictability of the driver.

In-Vehicle Driving Simulator: A key element of the Driver-Centric VANET simulation architecture is an in-vehicle environment that is fully equipped with driver interfaces able to convey information provided by VANET technology. This environment guarantees a realistic interaction between the driver and the system. We use a SMART car to build the physical simulator platform that consists of the following components:

- Cockpit with all controls;
- Steering wheel;
- Acceleration and braking pedals;
- Driver's car seat;
- Dashboard with instruments;
- Windshield for projecting information;
- A portable display system for easy assembly at any location;
- Sound system.

In addition we incorporate to the platform head up displays (HUD) to represent advanced in-vehicle signs. The multimodal interaction with the system through manual controls, touch screen, voice, maps, lights, etc. ensures a natural control and efficient use of the integrated applications.

We based the development of the in-vehicle driving simulation tool on the OpenSceneGraph [4] library, since it provides all the support for different scenario visual perspectives, in particular from the inside vehicle viewpoint. Along with the OpenSceneGraph, we have used the Bullet physics library [5] to provide the realistic physics associated with the driving task and the simulated scenarios.

A desktop computer allows operating the simulated vehicle using a simple client program. The client component includes 3D databases of the simulated scenarios. We designed an accurate 3D city model of the city of Porto that supports simulated driving experiments and allows to improve the mobility patterns from the vehicles simulated by DIVERT. Figure 2 illustrates the simulated urban scenario from the city of Porto populated from the DIVERT environment. The research in-vehicle driving simulator is shown in Figure 3.



Fig. 2. Simulated urban environment from the city of Porto populated from the DIVERT road scenario

Coupling Interface: A simulated scenario supported by one base map allows a feasible coupling of both, the driving simulator and DIVERT, since both applications share the same coordinate system. Therefore an accurate car trajectory can be assured. The scenario elements, namely, the roads, traffic signs, buildings, trees, etc are populated based on their geographical position from a 3D objects database.

The coupling architecture itself is a typical TCP client-server architecture, where DIVERT acts as a virtual GPS server, providing rich data sets gathered by GPS receivers, such as vehicle's position, heading and speed information of each vehicle being simulated. Since DIVERT provides a TCP server thread for external applications and the in-vehicle driving simulator is implemented as an autonomous external application, both can be easily connected. In this context the driving simulator acts as a client which connects to DIVERT to get information related to each simulated surrounding car. Any action performed by the driver in the driving simulator affects all his neighbors in DIVERT. The next section explains in detail the data exchange protocol that makes possible the simulated VANET environment from a driver-centric perspective.



Fig. 3. Research Driving Simulator of the Computer Science Department of the University of Porto

3.2 Data Exchange Protocol

The data exchange protocol for the coupling of the driving simulator and DIVERT consists of the following parts:

- **Data point creation:** DIVERT receives the information related to the vehicle operated by the human driver in the driving simulator, namely its current position, heading and speed. This information allows us to fit the data points into DIVERT locating and creating a new vehicle in the corresponding map coordinates. To locate the vehicle in the correspondent coordinates, we obtain the nearest road to the point efficiently, relying on spatial information technology for indexing multi-dimensional information. The data structure consists of nodes that have a variable number of entries. These entries store a child node Id and a bounding box of all entries within this child node. Each tree node corresponds to roads containing the information related to the lanes that constitute the road. The vehicles are bounded in the lanes, thus making it possible to select the nearest neighbors of vehicle simulated in the driving simulator.
- **Data point update:** A periodical update of the transmitted data set coming from the vehicle in the driving simulator is performed and sent automatically to DIVERT. These new values are then distributed to the neighbor data points, thus resulting in an update of the relevant parameters: vehicle

Id, position, heading and speed. The new values related to the vicinity nodes of the simulated vehicle are then sent back to the driving simulator.

- **Data point deletion:** In order to be capable of ending the driving simulation the data points related to the vehicle in the driving simulator need to be removed from DIVERT. Therefore the relevant information is sent to DIVERT and the deletion follows, being the vehicle not visible any more. A final confirmation sent to the driving simulator terminates the application.

3D Representation of Vehicles from DIVERT: To fit the data points from DIVERT into the driving simulator the process we follow is equivalent to the previous one. After having updated the vehicle's position from the driving simulator in DIVERT, the data points from the vicinity are received and a new 3D object is created in the in-vehicle driving simulator in the correspondent coordinates. In this way we achieve that each car in the in-vehicle driving simulator maintains the current position and heading from DIVERT. At each step both the in-vehicle driving simulator as well as DIVERT keep synchronized in terms of actual position, heading, and acceleration. The data exchange protocol used in our approach allows to populate the driving simulator with vehicles from DIVERT with the warranty and reliability of the DIVERT vehicles' data. In addition this approach allows us to calculate new data points in a simple and efficient way minimizing the introduction of errors. The next section describes the data synchronization between the coupled simulators in more detail.

3.3 Data Synchronization

In our approach we split the executable program into two running tasks being the instructions set contained on one hand, inside the driving simulator process and on the other inside the DIVERT process. Our approach allows the thread from the driving simulator process to receive continuously the position coordinates and other relevant information of the neighbor's nodes from DIVERT and at the same time to send the location, heading and acceleration of the car operated within the driving simulator to DIVERT. Even if the running tasks are divided into two processes, the data exchange protocol maintains some dependencies between each thread: in particular, when data is sent from the in-vehicle driving simulator to DIVERT or from DIVERT to the in-vehicle driving simulator.

In both cases the TCP client-server threads have to wait for the information to arrive from the TCP client located at the simulator. The TCP server in DIVERT also waits for information from a variety of external applications such as vehicles from the vicinity coming from the in-vehicle driving simulator. This behavior is crucial for the coupling since it is precisely this send-wait-response protocol that causes the thread data synchronization in both threads to occur without any kind of positioning error neither in DIVERT nor in the driving simulator.

Being based on TCP client-server architecture, the data exchange protocol allows having warranties on the delivery of each packet sent by both applications. The process assures synchronism in the simulators' coupling and enables

coherence on the simulated vehicles in both applications. The next section describes the mobility model of the Driver-Centric VANET simulator.

4 Mobility Model

The Driver-Centric VANET simulator mobility model relies on the DIVERT traffic flow mobility model based on the IDM. In the IDM model description the acceleration is expressed as a continuous function of the velocity v_0 , the space between the vehicle α and the vehicle in-front and their approaching rate. Table II describes all the parameters of the IDM model.

$$\dot{v}_\alpha = a^{(\alpha)} \left[1 - \left(\frac{v_\alpha}{v_0^{(\alpha)}} \right)^\delta - \left(\frac{s^*(v_\alpha, \Delta v_\alpha)}{s_\alpha} \right)^2 \right]$$

Parameters such as acceleration depending on the traffic situation can be expressed by the following formula.

$$a_f(v_\alpha) := a^{(\alpha)} [1 - (v_\alpha/v_0^{(\alpha)})^\delta]$$

Deceleration from the vehicle α depends on the required space to the vehicle in front (s^*) and the current space between both vehicles (s_α). This expression describes the relationship between the parameters:

$$s^*(v, \Delta v) = s_0^{(\alpha)} + s_1^{(\alpha)} \sqrt{\frac{v}{v_0^{(\alpha)}}} + T^\alpha v + \frac{v \Delta v}{2\sqrt{a^{(\alpha)} b^{(\alpha)}}}$$

These model expressions denote a pre-defined or randomly generated route for the cars to follow. They determine if the car needs to adapt the velocity to the current road conditions (i.e. encountering an obstacle ahead). The behavior of the vehicles behind is specified by the cars ahead.

Table 1. Description of the IDM parameters

Parameter	Description
v_0	desired velocity
T	safe time headway
a	maximum acceleration
b	desired deceleration
δ	acceleration exponent
s_0, s_1	jam distance
v	velocity
α	vehicle
v_α	vehicle velocity
s^*	vehicle in front
s_α	space between vehicles
a_f	acceleration on a free road

Coupling the vehicles from DIVERT and the driving simulator requires a synchronization of the mobility behavior of both simulators. The driver's performance in the driving simulator affects the vehicles' behavior and consequently the mobility model in DIVERT. The positioning in DIVERT of a new object, in this case, the car operated by the human driver in the driving simulator, alters the original DIVERT car-following model behavior. The simulated vehicles need to adapt their speed to the speed of the new object and thus their mobility patterns. Thus we adjust the vehicle's speed to avoid collisions with the inserted vehicle through additional acceleration and deceleration parameters. The vehicles can then respond to new mobility patterns such as being able to stop when the simulated car from the in-vehicle driving simulator stops. This extension allows us to test traffic jams and VANET connectivity issues that affect applications or devices from a realistic driver's perspective.

5 Simulation Platform Applicability

In this section we present two applications that can be explored within the Driver-Centric VANET simulation framework. Since our approach intended the creation of a simulation tool to reproduce real traffic conditions in a driving scenario, we designed the Driver-Centric VANET simulation according to the usability heuristics for users interfaces suggested by Nielsen [13] and we performed a preliminary usability inspection during the design phase that allowed us to correct possible problems in the user interface to further work on a more detailed usability study in future projects. During the inspection issues related to animation and artwork could be improved. In addition the grade of accuracy of the system in terms of damage, physics, and control was increased. We concluded that the Driver-Centric VANET simulation tool provides the right framework to validate VANET-based applications in the inter-vehicle communication context due to its capacity to model realistic traffic condition scenarios. Two examples for such VANET applications are described below. To test the feasibility of the systems we need to evaluate the wireless network and the communication capabilities between the vehicles involved; in particular the system performance over the transmission process focusing on delay and packet loss as function of the distance.

5.1 Virtual Traffic Lights (VTL)

Virtual Traffic Lights (VTL) will allow the migration of traffic lights as road-based infrastructures to in-vehicle virtual signs supported only by vehicle-to-vehicle communication [10]. The Driver-Centric VANET simulator constitutes the perfect framework to measure the drivers' responsiveness to Virtual Traffic Lights projected on the windshield.

When a crossing conflict is detected a need for a VTL arises. One of the conflicting vehicles acts as an intersection leader creating and controlling the VTL. This leader stops at the intersection and temporarily replaces a road-based traffic light in the control of the intersection. The fact that a person

is operating a car simulated in the VANET environment provides us with the perfect scenario to perform tests related to the intersection lead. In addition our simulator allows us to analyze the VTL performance based on the number of vehicles in each intersection. The VTL approach means a radical change from the traditional way of controlling intersections through physical infrastructures. Thus a progressive process of the VTL deployment is required and a tool for experimenting with a partial deployment of the new technology that do not arise safety issues is crucial.

The Driver-Centric VANET simulation tool described in this paper can evaluate not only scenarios of 100% penetration rate of the technology, but also partial deployment scenarios and interaction with cyclists and pedestrians. Furthermore, several user studies on the interaction between the VTL and the driver can be conducted using this simulation tool.

5.2 The See-Through System (STS)

The See-Through System (STS) is a cooperative Advanced Driver Assistance System (co-ADAS) for the overtaking maneuver [14]. Vision obstructing vehicles clearly reduce the awareness of drivers that travel behind. Thus the overtaking of long, vision-obstructing vehicles on the road, is a difficult and challenging task when there is no overtaking lane other than the one used by vehicles traveling in the opposite direction. As a result, a longer following distance is normally kept by the vehicle traveling behind, to improve the field of vision, as well as to increase the reaction time in the event of a sudden brake or maneuver by the vehicle in front.

The STS uses VANET technology to provide a video-streaming between the vehicle in front and the vehicle behind. VANET allow an automated vehicle-to-vehicle communication carried out through wireless technologies that provide the driver with additional visibility. Since connection issues that can affect the reliability of this information are likely to arise we need to evaluate the wireless network and the communication strength between the vehicles involved. Our Driver-Centric VANET simulator represents the ideal testing platform to evaluate these issues. In addition, like an additional rear-view mirror the STS not only provides additional visibility of the road ahead to the driver but also suffers from blind spots. These blind spots affect the visibility of vehicles coming from the opposite direction leaving the area of sight of the camera located on the vehicle streaming the video. Our Driver-Centric VANET simulator allows detecting the significance of the blind spot issue.

6 Conclusion

In this paper we propose a methodology for a simulation platform that integrates both a VANET simulator and an in-vehicle driving simulator for a realistic simulation of different traffic situations and traffic flow. The platform provides a suitable environment for validating VANET-based applications in the

inter-vehicle communication context in a realistic manner without having safety related issues. The Driver-Centric VANET simulation permits to test the utility of the systems as well as the communication capabilities between the simulated vehicles. The usefulness of the proposed modeling was shown by the analysis of the VANET-based applications See-Through System and Virtual Traffic Lights.

Acknowledgments. The authors gratefully acknowledge the financial support of the DRIVE-IN project “Distributed Routing and Infotainment through VEHicular Inter-Networking” funded by the Portuguese Foundation for Science and Technology (FCT) under the contract number CMUPT/ NGN/0052/2008.

References

1. Openscenegraph (2007), <http://www.openscenegraph.org>
2. Abdennour, A., Al-Ghamdi, A.: Artificial neural networks application to the estimation of vehicle headways in freeway sections. In: Proceedings of the 81st Transportation Research Board Annual Meeting, Washington, DC (2002)
3. Alm, T., Ohlsson, K., Kovordanyi, R.: Glass Cockpit Simulators: Tools for IT-based car systems design and evaluation. In: Proceedings of Driving Simulator Conference (2005)
4. Chu, K., Joseph, S.: Second Life Prototyping of Augmented Automobile Navigation Assistance. In: 11th International IEEE Conference on Intelligent Transportation Systems, ITSC 2008, pp. 389–394 (2008)
5. Coumans, E.: Bullet physics library (2009), <http://www.bulletphysics.com>
6. Dumbuya, A., Booth, A., Reed, N., Kirkham, A., Philpott, T., Zhao, J., Wood, R.: Complexity of Traffic Interactions: Improving Behavioural Intelligence in Driving Simulation Scenarios. Complex Systems and Self-organization Modelling, 201–209 (2009)
7. Dumbuya, A., Wood, R.: Visual perception modelling for intelligent virtual driver agents in synthetic driving simulation. Journal of Experimental & Theoretical Artificial Intelligence 15(1), 73–102 (2003)
8. Fernandes, R., D’Orey, P.M., Ferreira, M.: DIVERT for Realistic Simulation of Heterogeneous Vehicular Networks. In: 2nd IEEE International Workshop on Intelligent Vehicular Networks - InVeNET (2010)
9. Ferreira, M., Conceição, H., Fernandes, R., Tonguz, O.: Stereoscopic aerial photography: an alternative to model-based urban mobility approaches. In: Proceedings of the sixth ACM International Workshop on VehiculAr InterNETworking, pp. 53–62. ACM, New York (2009)
10. Ferreira, M., Fernandes, R., Conceição, H., Viriyasitavat, W., Tonguz, O.K.: Self-Organized Traffic Control. In: 7th ACM International Workshop on Vehicular Inter-Networking-VANET (2010)
11. Karnadi, F., Mo, Z., Lan, K.: Rapid generation of realistic mobility models for vanet. In: IEEE Wireless Communications and Networking Conference, WCNC, pp. 2506–2511 (2007)
12. Mangharam, R., Weller, D., Stancil, D., Rajkumar, R., Parikh, J.: GrooveSim: a topography-accurate simulator for geographic routing in vehicular networks. In: Proceedings of the 2nd ACM International Workshop on Vehicular ad hoc Networks, p. 68. ACM, New York (2005)

13. Nielsen, J.: Usability inspection methods. In: Conference companion on Human Factors in Computing Systems, pp. 413–414. ACM, New York (1994)
14. Olaverri-Monreal, C., Gomes, P., Fernandes, R., Vieira, F., Ferreira, M.: The See-Through System: A VANET-enabled assistant for overtaking maneuvers. In: Intelligent Vehicles Symposium (IV), pp. 123–128. IEEE, Los Alamitos (2010)
15. Piórkowski, M., Raya, M., Lugo, A.L., Papadimitratos, P., Grossglauser, M., Hubaux, J.P.: TraNS: realistic joint traffic and network simulator for VANETs. ACM SIGMOBILE Mobile Computing and Communications Review 12(1), 31–33 (2008)
16. Sommer, C., Dressler, F.: Progressing toward realistic mobility models in VANET simulations. IEEE Communications Magazine 46(11), 132–137 (2008)
17. Sommer, C., Yao, Z., German, R., Dressler, F.: On the need for bidirectional coupling of road traffic microsimulation and network simulation. In: Proceeding of the 1st ACM SIGMOBILE Workshop on Mobility Models, pp. 41–48. ACM, New York (2008)
18. Treiber, M., Hennecke, A., Helbing, D.: Congested traffic states in empirical observations and microscopic simulations. Physical Review E 62(2), 1805–1824 (2000)

Simulative Evaluation of the Potential of Car2X-Communication in Terms of Efficiency

Benno Schweiger¹, Philipp Ehnert¹, and Johann Schlichter²

¹ BMW Group Research and Development
Hanauer Straße 46, 80992 München, Germany
{benno.schweiger,philipp.ehnert}@bmw.de

² Technische Universität München
Institut fuer Informatik
Boltzmannstr. 3
85748 Garching
schlichter@in.tum.de

Abstract. Much effort has been put to evaluate the potential of Intelligent Transportation Systems based on Car2X communication with respect to traffic safety. However, harnessing Car2X communication for efficiency purposes also offers potentials in terms of reducing fuel consumption and CO₂. In this paper we evaluate these potentials by simulating three different traffic situations. We compare an uncontrolled scenario with one that incorporates a control mechanism that uses the possibility of Car2X communication to influence a single car directly. Our results show a significant potential for a reduction of fuel consumption for some of the traffic scenarios, motivating further research into Car2X communication based traffic control systems.

Keywords: Car2X, Car2Car, Car2Infrastructure, Efficiency, Fuel Consumption, CO₂.

1 Introduction

Intelligent Transportation Systems (ITS) are an area that is currently under wide research because of their promise to enhance various aspects of transportation. Communication systems between vehicles (Car2Car) and between vehicles and infrastructure elements such as traffic lights (Car2Infrastructure), together known as Car2X communication, are one example for ITS. Many projects have dealt and are still dealing with using Car2X communication to enhance traffic flow and increase traffic safety. Examples are state funded projects like simTD [1], Aktiv [2] and CVIS [3]. On the other hand, Car2X communication also offers the possibility to reduce fuel consumption. By being able to address a single car directly, precise driving instructions can be given. An adequate system (be it centralized or decentralized) can thus control traffic in a manner so that fuel consumption is minimized.

In this paper we want to motivate research into these systems by evaluating the potential on fuel consumption reduction they show. For this purpose, we

assume a system that is able to influence the driving patterns of single cars. We also assume it has total information about current traffic.

For evaluation purposes we select three common traffic situations. We simulate these scenarios using the microscopic traffic simulator SUMO [4]. The evaluation is done by comparing simulation runs where cars drive without any additional control with runs where we control the traffic so that traffic flow is optimized.

The remainder of this paper is structured as follows: After a review of related Work in section 2 we describe the traffic scenarios we consider for our work in section 3. Section 4 contains the simulation setup and execution followed by an evaluation of the results in section 5. Section 6 concludes the paper and gives an outlook on future work.

2 Related Work

In the context of using Car2X communication for efficiency purposes only few scenarios have yet been researched. Most research has been in the field of intersections, both with and without traffic lights. [5] shows a 12% to 14% reduction in fuel consumption for an arterial road passing several traffic lights. In [6] the authors even identified reductions in fuel consumption of up to 47% for a similar setting. [7] proposes a decentralized unmanaged system for intersection control that reduces delay times at intersection, which should reduce fuel consumption as well.

[8] and [9] propose the use of communication technologies to collect information about surrounding traffic to further optimize the operating strategy of the drive train of hybrid electric vehicles. [10] mentions the use of probe vehicles to gather real time traffic information to reduce fuel consumption by finding alternative routes to avoid areas with traffic disturbances.

In this paper however, we do not analyze fuel saving potentials by changing the route to the destination or by adapting the operating strategy of the vehicle but instead focus on increasing efficiency by influencing the manner of driving. In this area, reduction of fuel consumption can be accomplished by reducing the dynamic of speed and idling time.

3 Traffic Scenarios

In order to be able to perform microscopic traffic simulations in a timely manner, we needed to create small road networks that focus on the desired effects. Thus, the first step in our work was to select the traffic scenarios we want to simulate.

3.1 Selection

To get a good overview of the efficiency potential we wanted to pick traffic scenarios from inner city traffic as well as grade separated and at grade highway traffic.

The first step in the selection process was to compile a catalogue of possible traffic scenarios, divided into the three areas mentioned above. The result of this step is shown in table 1. Some scenarios show up more than once because they can happen in more than one area.

Table 1. Catalogue of traffic situations

Inner city traffic	Highway traffic	Freeway traffic
Turns	Turns	Speed recommendation
Approaching a traffic light	Approaching a traffic light	Approaching an obstacle
Speed recommendation	Speed recommendation	Grade separated intersections
Approaching an obstacle	Approaching an obstacle	Merging lanes
Roundabout	Roundabout	Changing of lanes
Looking for parking space	Grade separated intersections	Weather warnings
Grade separated intersections	Merging lanes	Circumnavigating traffic jams
Merging lanes	Changing of lanes	
Changing of lanes	Circumnavigating traffic jams	
Circumnavigating traffic jams	Passing without dedicated lane	
Unmanaged intersections	Unmanaged intersections	
Emergency vehicles	Weather warnings	

In the second step we used the Analytical Hierarchy Process (AHP) [11] to prioritize the situations. As criteria we used the amount of dynamic inherent in the situation, frequency of the situation, feasibility of optimizing driving behavior in the situation, influence on traffic safety and potential to unburden the driver. The top five results of the AHP are shown in table 2.

Table 2. Top five results of the AHP

Situation	Result
Approaching an obstacle	19.49%
Approaching a traffic light	13.26%
Unmanaged intersections	12.79%
Roundabout	11.79%
Turns	11.56%
...	...

As mentioned before, traffic lights and unmanaged intersections have already been dealt with. We didn't think we could add new findings to the work already done, so we decided to put our focus towards other situations. The situations we selected are: Inner city roundabout, left turn on a single lane highway and approaching an obstacle in the form of a congestion on a freeway.

3.2 Roundabout

Roundabouts are a self regulating form of intersection. They offer a better traffic flow resulting in less delay than other forms of intersection [12]. At first glance, the design of a roundabout seems to offer an already high level of optimization. However, unnecessary stopping and idling time can occur: Vehicle A may force vehicle B to stop, if vehicle A is on the ring at the position, where vehicle B wants to enter the ring. This can be avoided, if vehicle B arrives a few seconds earlier or later so vehicle A doesn't block vehicle B. In the controlled scenario we achieve this behavior by adding a timing to the vehicles in the approaches, so that all vehicles arrive at the circle at the same time. This way, vehicles on all four approaches can enter the ring without affecting each other. This should reduce the need to stop and thus enable a smoother traffic which in turn causes a reduction in fuel consumption.

3.3 Left Turn on Single Lane Road

On a single lane at grade highway, a vehicle that wants to take a left turn¹ has to yield to oncoming traffic. While waiting, it blocks the road, making it impossible for the following vehicles to pass. In case of high traffic volume, this can lead to long delays and accordingly long idling times causing increased fuel consumption. This can be avoided, if the oncoming traffic is influenced in a way, that a vehicle leaves a gap to its leading car that is sufficiently large for the turning vehicle to perform its turn.

In the controlled scenario this is realized by influencing the oncoming traffic so that gaps of a sufficient size are formed. At the same time, the arrival of the turning vehicles is timed so they arrive at the beginning of a gap.

3.4 Approaching Congestion on Freeway

Despite the best control systems, congestions e.g. as a result of accidents or construction work cannot be avoided. When approaching a congestion or a traffic jam on a freeway, a vehicle has to reduce its speed to a stop. The speed at the beginning of the deceleration is so high, that a reduction by shearing or coasting is often not possible because the driver would have to start coasting or shearing before he is aware of the congestion. Instead the vehicle drives towards the congestion at cruising speed and decelerates so late that, depending on the line of sight conditions, the necessary braking may even be safety critical. This way, the vehicle loses the chance to utilize its kinetic energy and thus an opportunity to increase fuel efficiency. To eliminate the need for hard braking, we apply the idea of a virtual speed funnel [13] in our controlled scenario. Figure 1 shows exemplary speed profiles for the controlled and uncontrolled approaching of a congestion.

¹ For left hand traffic: right turn.

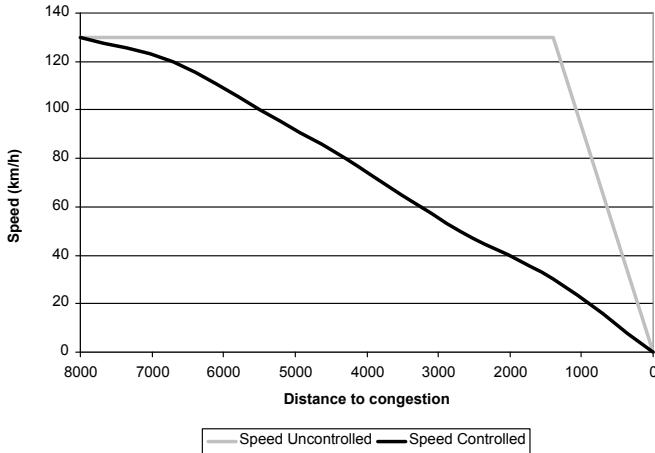


Fig. 1. Comparison of speed profiles for approaching of a congestion (Simulation result)

4 Simulation

To be able to evaluate the effects on single cars, we needed to use a microscopic traffic simulation. We chose SUMO [4] as simulation tool. Our simulations were performed using versions 0.11.1 and 0.12.0. SUMO uses the Krauß traffic model [14] that takes into account the distance to the leading vehicle, the velocities of the leading and ego vehicle, acceleration and deceleration profiles. It also includes a component to allow for human error. In order to get realistic traffic behavior, all scenarios are simulated as part of a small road network.

The values for fuel consumption are calculated using the SUMO HBEFA-Based-PollutantEmission model [15]. This model uses a slightly modified version of the Handbook Emission Factors for Road Transport (HBEFA) [16] to calculate the values for pollutant emissions.

4.1 Simulation Setup

Roundabout. We simulated a typical urban single lane roundabout [2] with a diameter of 35m. It has four single lane approaches end egresses (see figure 2). The traffic volume is Poisson distributed on all approaches with an average of 300 vehicles per hour per approach. The egress is equally distributed on the three remaining roads, vehicles doing a full circle are not included. The speed limit for the entire scenario is 50 km/h.

Left Turn on Single Lane Highway. Our simulation model for the left turn scenario consists of a single lane highway intersecting with an also single lane minor road. The speed limit on the highway is 80 km/h, on the minor road 50 km/h. The traffic volume is Poisson distributed with an average of 800 vehicles

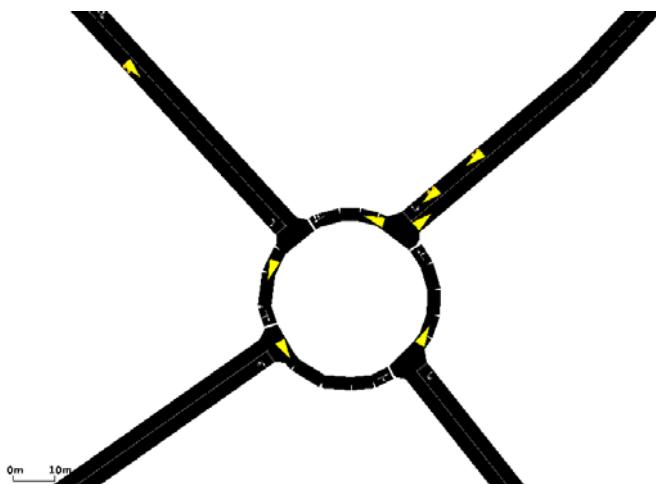


Fig. 2. SUMO representation of roundabout scenario

per hour in each direction of the highway. Traffic originating on the minor road is not included in the simulation. **3** shows an illustration of the intersection.



Fig. 3. SUMO representation of left turn scenario

Approaching Congestion on Freeway. For the congested freeway scenario we simulated a three lane freeway without any on-ramps or exits. Traffic volume is Poisson distributed with an average of 4000 vehicles per hour combined for all three lanes. The congestion is provoked by a complete lockdown of all lanes for one direction. Figure **4** illustrates the simulated freeway.

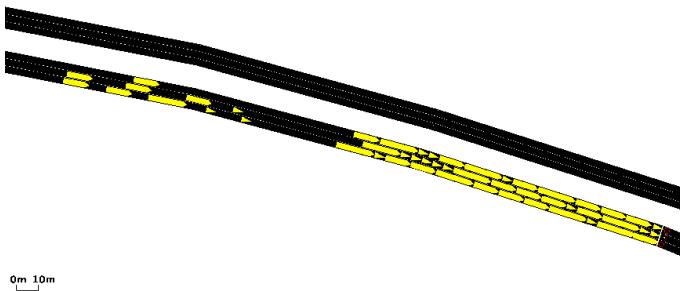


Fig. 4. SUMO representation of freeway congestion scenario

4.2 Simulation Execution

For every traffic scenario, two simulation runs were performed: The first one represents the uncontrolled version of the scenario: All vehicles were using standard behavior models, no external control was present. For the second run, we implemented the control mechanisms as described in section 3. For the purpose of this potential evaluation, the control mechanisms were implemented by manually assigning individual driving profiles to each vehicle. For each scenario, the driving profiles were hand tuned using the visual 2D output of SUMO. Once the desired effect was achieved, we performed the complete simulation run and recorded the raw log data of all vehicles.

5 Results and Evaluation

In our evaluation of the simulation results, we concentrated on the parameters travel time, waiting time and fuel consumption. Since we are interested in realistic data rather than best cases, we calculated the average value over all vehicles in the respective scenario for these parameters. Travel time is defined as the time, the vehicle travels between start and destination. Waiting time is the time, a vehicle spend at a standstill. Fuel consumption is the accumulated amount of fuel the vehicle burned while traveling from start to destination.

5.1 Roundabout

The results for the roundabout scenario are shown in table 3. At first glance the relative reduction of the waiting time of over 30% is significant. However, due to the small absolute value, this doesn't reflect in comparably significant reductions in travel time and fuel consumption. Nevertheless, a reduction of 1.13% in average over all vehicles is noticeable. The effect is caused by the more fluent traffic movement which eliminates the need to brake and accelerate again.

Table 3. Results of scenario “Roundabout”

	Uncontrolled	Controlled	Difference
Avg. Travel Time (s)	501.04	494.30	-1.35%
Waiting Time (s)	7.52	5.21	-30.7%
Fuel Consumption (ml)	443	438	-1.13%

5.2 Left Turn on Single Lane Road

For the left turn scenario, the goal was to reduce the waiting time by allowing vehicles to cross oncoming traffic quicker. Our approach proved successful: The average waiting time per car was reduced by over 90%. The effect is augmented by the fact that our scenario took place on a single lane highway: While the left turning car has to wait, no other vehicles are able to pass. So by allowing the turning car to perform the turn quickly, the waiting time for all following vehicles is minimized. We are aware of the fact, that this augmentation can be weakened by introducing dedicated left turn lanes enabling the turning car to wait without blocking the following traffic. However, it is not totally eliminated. In heavy traffic conditions dedicated left turn lanes often prove ineffective. Because of their limited size, they can simply overflow which forces additional turning cars to block the following traffic again.

The reduction in fuel consumption seems comparably small. However, the reduction of 2.18% is an average of all vehicles in the scenario includes those in the opposite direction of the left turning vehicle. These vehicles do not profit from the traffic flow enhancement. They have the right of way, so their travel time is quite the same with and without control. Because of that, they do not add to the reduction potential making the entire saving caused by only half the vehicles.

Table 4. Results of scenario “Left Turn”

	Uncontrolled	Controlled	Difference
Avg. Travel Time (s)	216.91	205.85	-5, 10%
Waiting Time (s)	102.95	9.52	-90.75
Fuel Consumption (ml)	1649	1613	-2.18%

5.3 Approaching Congestion on Freeway

When approaching the congestion, our system starts influencing traffic about 6700m ahead of the lockdown. This results in a very smooth deceleration compared to the standard behavior. That smooth deceleration means, that vehicles can make use of their kinetic energy instead of releasing it into the air while braking. This is reflected in the reduction in fuel consumption our simulation shows: In average consumption is reduced by as much as 13.9%. Smoother traffic

also helps in dissolving the congestion once the lockdown is released. This shows in the average travel time being shortened by 12.26%.

Table 5. Results of scenario “Approaching Congestion”

	Uncontrolled	Controlled	Difference
Avg. Travel Time (s)	1346.75	1181.62	-12.26%
Waiting Time (s)	23.82	17.79	-25.31%
Fuel Consumption (ml)	3509	3021	-13.90%

6 Conclusion and Outlook

In conclusion, our simulations show a significant potential for Car2X based traffic control systems in term of reducing fuel consumption and travel time in some scenarios. In other cases, the savings possible are rather small. Figure 5 illustrates this.

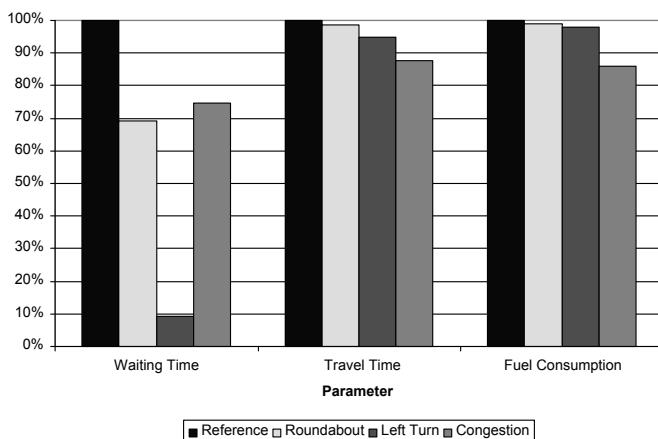


Fig. 5. Comparison of potentials for the regarded scenarios

We are well aware that the hand crafted driving strategies we implemented in our simulations are in no way to be regarded as optimal solutions or generally applicable algorithms. However, the potentials we identified represent a lower boundary for the theoretical potential of Car2X communication based traffic control systems. As such they show that additional research into this topic is merited.

An aspect we left completely out of scope for now is how the control systems we emulated by manually influencing single vehicles can be realized. In future work

we plan to analyze exactly what information would be necessary for systems like the ones we described to optimize traffic flow. Based on a comparison of these information need with the capabilities of currently discussed communication and measurement systems, the feasibility of the respective control system can be determined.

References

1. simTD Consortium: simTD Website (November 2010), <http://www.simtd.org/>
2. aktiv Forschungsiniziativte: aktiv Website (November 2010), <http://www.aktiv-online.org/>
3. CVIS Consortium: CVIS Website (November 2010), <http://www.cvisproject.org/>
4. Krajzewicz, D., Hertkorn, G., Wagner, P., Rössel, C.: SUMO (Simulation of Urban MObility) An open-source traffic simulation Car-Driver Model
5. Mandava, S., Boriboonsomsin, K., Barth, M.: Arterial Velocity Planning based on Traffic Signal Information under Light Traffic Conditions. In: 12th International IEEE Conference on Intelligent Transportation Systems, St. Louis, Mo, USA, pp. 160–165 (2009)
6. Asadi, B., Vahidi, A.: Predictive Cruise Control: Utilizing Upcoming Traffic Signal Information for Improving Fuel Economy and Reducing Trip Time. IEEE Transactions on Control Systems Technology (2010)
7. Vanmiddlesworth, M., Dresner, K., Stone, P.: Replacing the Stop Sign: Unmanaged Intersection Control for Autonomous Vehicles. In: Padgham, L., Parkes, D., Müller, J., Parsons, S. (eds.) Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008), Estoril, Portugal, pp. 1413–1416 (2008)
8. Back, M.: Prädiktive Antriebsregelung zum energieoptimalen Betrieb von Hybirdfahrzeugen. PhD thesis, Karlsruhe, Germany (2005)
9. Abdul-Hak, M., Al-Holou, N.: ITS based Predictive Intelligent Battery Management System for plug-in Hybrid and Electric vehicles. In: 2009 IEEE Vehicle Power and Propulsion Conference, pp. 138–144 (2009)
10. Ericsson, E., Larsson, H., Brundell-freij, K.: Optimizing route choice for lowest fuel consumption - Potential effects of a new driver support tool. Transportation Research 14, 369–383 (2006)
11. Saaty, T.L.: The Analytic Hierarchy Process: Planning Setting Priorities. Resource Allocation (1980)
12. Turner Fairbank Highway Research Center: Roundabouts: An Informational Guide (2000)
13. Assemacher, S., Killat, M., Schmidt-Eisenlohr, F., Vortisch, P.: A Simulative Approach For The Identification of Potentials and Impacts of V2X-Communication. In: 15th World Congress on ITS, New York, USA (2008)
14. Krauss, S.: Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics (1998)
15. SUMO Developers: SUMO HBEFABasedPollutantEmission (November 2010), http://sourceforge.net/apps/mediawiki/sumo/index.php?title=SUMO_HBEFABasedPollutantEmission
16. INFRAS: Handbook of Emission Factors for Road Transport (2010)

Performance Study of an In-Car Switched Ethernet Network without Prioritization

Hyung-Taek Lim, Kay Weckemann, and Daniel Herrscher

BMW Group Research and Technology

Munich, Germany

{Hyung-Taek.Lim, Kay.Weckemann, Daniel.Herrscher}@bmw.de

Abstract. This paper presents the current state of our research in real-time communication of an IP-based in-car network. The Internet Protocol (IP) will serve as convergence layer of different specific in-car network protocols and IEEE 802.3 Ethernet will be the basic technology to transport IP. In this work, we evaluate a legacy switched Ethernet network without any Quality of Service (QoS) mechanisms. While there are arguments for not using QoS mechanisms, we give evidence that communication requirements and service constraints of a more and more streaming intensive in-car network cannot be met without. We argue for a setup with different traffic types: CAN and FlexRay like control messages, camera streaming, video and audio streaming, and bulk traffic. We will also argue for a simple double star topology as a valid assumption where the target architecture of the IP-based in-car network is not yet clear. Setup and simulation will serve as framework and motivation for future work: Analyzing IP-based real-time communication using QoS mechanisms - characterizing traffic classes after IEEE 802.1Q and IEEE 802.1 Audio Video Bridging (AVB).

Keywords: IP-Based In-Car Network, Switched Ethernet, Quality of Service Performance Evaluation.

1 Introduction

A current automotive network consists of more than 70 Electronic Control Units (ECUs) which are interconnected by different automotive specific network technologies such as the Controller Area Network (CAN), FlexRay, Media Oriented System Transport (MOST), and Ethernet. Due to the heterogeneous in-car network, different protocols are used and translated by an application layer gateway (see Fig. 1). Current automotive specific network technologies only provide low bandwidth. In the near future, the number of applications with high bandwidth demand will grow and thus imply the introduction of new technologies: Driver assistance camera applications will increase safety and comfort and the users' demand for high quality streaming data in the infotainment domain (e.g. BlueRay, HDTV) and bulk data due to back-end services is increasing. IEEE 802.3 Ethernet is a tested and proven standard that fulfills the high bandwidth

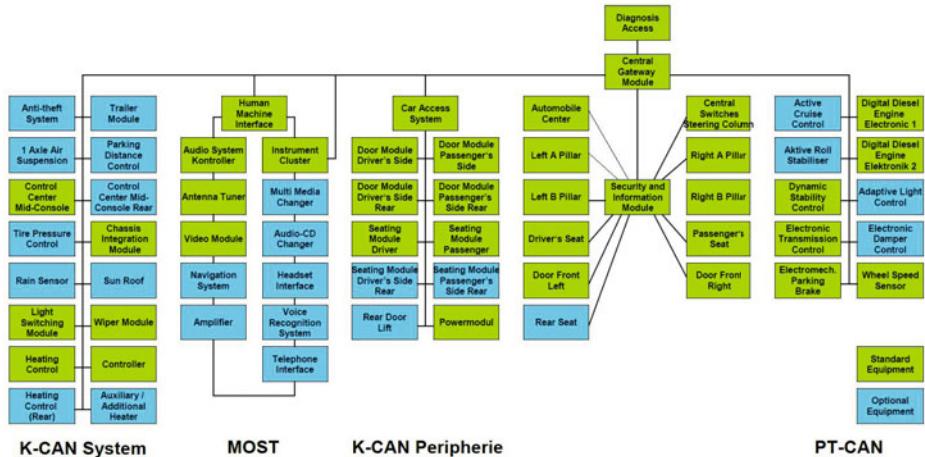


Fig. 1. Automotive network architecture [6]

demand. Several derivates of Ethernet exist that differ in connectors, cables, physical and medium access control layers. The automotive requirements for high bandwidth, low costs and good electromagnetic compatibility (EMC) can be met when the cable harness of Ethernet is realized with a 100 Mbit/s duplex unshielded derivate using a single twisted pair cable with adapted physical layer [1].

In this paper we provide the setup for analyzing the performance of a switched Ethernet based in-car network with different traffic classes and a general topology. This setup is used for simulation and interpretation of traffic characteristics where the IP traffic is modeled by using standard IP frames and packet sizes. The here-presented simulation using Ethernet without prioritization mechanisms shall motivate future work towards finally meeting automotive communication requirements.

The remainder of this paper is structured as follows: In section 2 we present performance studies with similar set-ups and the main differences compared to our work. Section 3 describes the setup by arguing for the modeling of Ethernet and a network topology, analyzing in-car traces of CAN and FlexRay for deriving a real-time communication model, and defining streaming and bulk data traffic models. Section 4 evaluates the performance of a switched Ethernet network without prioritization by simulation, presents the results, and reflects on the chosen setup. Section 5 concludes the paper as we define the tasks which have to be addressed on our research path.

2 Related Work

During the last years, researchers analyzed the applicability of a switched Ethernet network for automotive usage. Daoud et al. analyzed a switched Ethernet

network based on 1000Base-T in a star topology without any QoS mechanisms for an in-car network, where two traffic classes were defined: control data and video streaming data [2]. The authors showed that service constraints are not violated, but they did not analyze the influence of other traffic classes (e.g. bulk data) towards the safety related control data. Furthermore, the authors modeled a gigabit Ethernet network. This is evidently a solution to easily address future bandwidth demand, but not a near-term option due to costs. We consider the usage of 100 Mbit/s Ethernet as a precondition.

A switched Ethernet network could have different topologies, e.g. star, double-star, daisy chain or combinations of different topologies. The effects of different topologies considering several automotive applications were analyzed by Rahmani et al. [3] [4]. The authors gave evidence that a ring based topology provides better performance than a double-star based topology concerning the need for traffic shaping mechanisms for variable bit rate (VBR) data traffics to avoid large bursts and packet losses. However, the proposed ring based topology is configured as a unidirectional ring only. Each connected device and switch has to be modified which objects our precondition to only use standard-conform components.

3 Performance Study Setup

In this section, we describe the setup for our ongoing research on analyzing Ethernet performance. The first part argues for the modeled Ethernet derivate and for not using prioritization mechanisms in the first step. Afterwards, the usage of a double-star topology is motivated and described. The network load has a strong influence on the service behavior, thus we consider different traffic classes for the simulation.

3.1 Usage of Ethernet

Our model uses Ethernet based on the 100Base-TX standard. In today's cars, 100Base-TX Ethernet as a point-to-point connector is used as isolated application for On-Bord Diagnostics (OBD) / Diagnostics over IP (DoIP) [5] where flashing time and available bandwidth are important factors to minimize time at the repair shop. The second use case is transmitting bandwidth-intensive map information for navigation applications.

In case of more than two nodes in a network, switched Ethernet is required as next expansion stage. The prospected Ethernet derivate for in-car communication over a single twisted pair cable with adapted physical layer does not differ from the 100Base-TX from the simulation point of view, as data rate and Medium Access Control (MAC) remain the same. A gigabit Ethernet is not a near-term option, because it is not sufficiently electromagnetic compatible, hardware is expensive, and most ECUs could not handle data at such high rates.

In a switched Ethernet network, the network has to be carefully dimensioned in order to support real-time applications with high service constraints. The

standard though does not enforce Quality of Service (QoS) mechanisms that would deal with packet delays and packet loss at switches. Prioritization needs configuration of end devices and/or switches: The currently built-in Ethernet for Diagnostics over IP and Media-Oriented Bulk Traffic do not implement prioritization for simplicity reasons. Also QoS configuration strategies for switches and end devices are still an unsolved issue. For these reasons we want to analyze the applicability of off-the-shelf Ethernet technology by considering different traffic classes used in a vehicle.

3.2 Topology

The trend towards integration of functionality from several small ECUs onto one combined, powerful device will generate the need for more communication bandwidth. This integration can be motivated by functional similarity, spatial proximity or other means. To prepare the migration, general concepts are needed that will later-on be adopted to a certain migration scenario. The topology in a switched Ethernet in-car network directly influences the service properties. While the future network topology is not foreseeable, the double-star topology we assume for our simulations is suitable for our analysis. Looking at the geometrical arrangement of ECUs in a vehicle, most of them are located either at the front or at the rear, as there is only little installation space inside or under the passenger cabin. Therefore they can easily be mapped to a double-star topology with a front and a rear switch (see Fig. 2). Future detailed consideration may lead to a multi-star topology which we see as a variant only. Star-based topologies lead to a less complicated cable harness compared to e.g. daisy chain topologies, especially if some of the nodes are optional. Furthermore star-based topologies are confirmed with the IEEE 802.3 and IEEE 802.1 Ethernet standards, so that no modifications on the MAC layer are required. In our system model there is

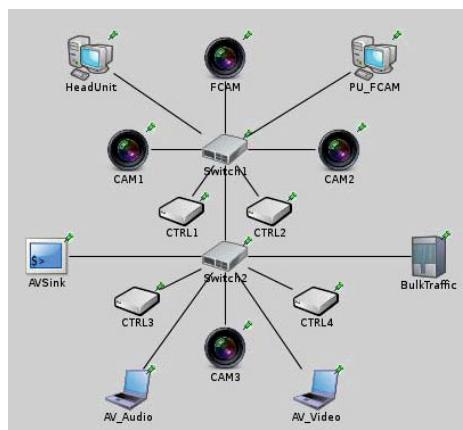


Fig. 2. Reference simulation topology

a processing unit (Head Unit) connected to the front switch which operates as the sink of control data, camera data and bulk data. There are two side-cameras and one rear-camera which generate and transmit the driver assistance related video data to the Head Unit. The Head Unit calculates a bird's eye view from these three video streams. Another camera generates video streaming data for the front view system terminated at a dedicated processing unit ('PU_FCAM'). In the rear part the 'AVSink' represents a rear seat entertainment (RSE) system. Audio and video streaming data from end nodes located in the rear part of the network are transmitted to the RSE. The 'BulkTraffic' node represents open connections to the internet which transmit bulk data to the Head Unit.

3.3 Traffic Characteristics

The simulation scenario contains four different traffic types: control data, driver assistance camera streaming, video and audio streaming, and bulk traffic data. First, an analysis of CAN and FlexRay serves as reference for the modeling of the control data traffic type. Afterwards the other traffic types are discussed.

Control Message Traffic Type - CAN and FlexRay Real-Time Communication. The analysis of real data traffic gives evidence on traffic characteristics of a current in-car network with focus to real-time control communication. Therefore the analysis concentrates on the CAN- and FlexRay bus systems. We used CAN and FlexRay message traces derived from a BMW car.

Packet size. CAN has a maximum message size of 8 byte while FlexRay is able to transmit data with a maximum payload size of 254 byte. The results give an indicator for the traffic characteristics of control and regulation messages in a current automotive network. Figure 3(a) and figure 3(b) show the distribution of CAN and Flexray payload sizes as used for the probability density functions (PDFs).

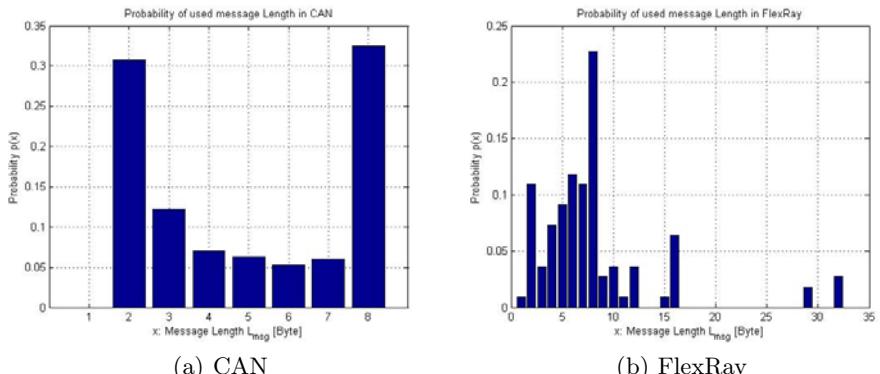


Fig. 3. PDF of data length based on CAN and FlexRay

Cycle Times. The cyclic messages on CAN have higher cycle times than those on FlexRay. ECUs which generate messages with high cycle times are likely on a CAN bus while messages with low cycle times are likely on FlexRay. The cycle times are less than 100ms for more than 80% of the transmitted FlexRay messages and approximately 52% of the CAN messages (see Fig. 4(a)). The situation is quite similar in case of event-based messages, while these are mostly sent on CAN (see Fig. 4(b)).

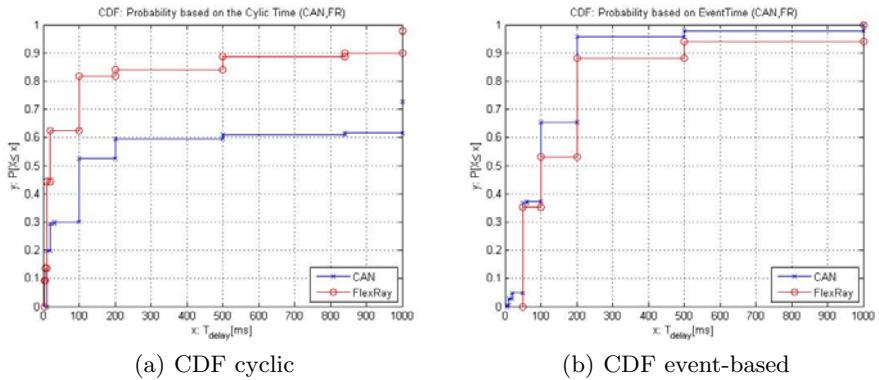


Fig. 4. Cumulative Distribution Function (CDF) of cyclic and event-based messages in CAN and FlexRay

Following this analysis, we chose a service rate based on an uniform distribution from 10ms to 100ms with a UDP payload size of 20 bytes, which is a reasonable size, because it works for virtually all CAN and FlexRay messages - shorter UDP payloads would be padded to at least 18 byte anyway. The control data has the highest latency and service requirement. It is sent to the Head Unit.

More Traffic Types - Driver Assistance Streaming, Infotainment Streaming and Bulk Traffic. In a vehicle, different driver assistance cameras capture high quality video streaming data from the outside of a vehicle. The Head Unit aggregates the different video streams to derive environment information. We assume that the driver assistance camera data is based on MPEG2-TS video streaming data which has a bit rate of approximately 25Mbit/s [1]. This traffic type has the highest bandwidth and a medium end-to-end delay requirement. A single camera generates and transmits UDP frames with a service rate of 0.25ms. The infotainment streaming data is characterized by multimedia data, where a rear seat entertainment (RSE) is simulated. The RSE ('AVSink') node receives video and audio streaming data transported by UDP frames from the video and audio source nodes within a vehicle. The audio streaming data is based on a Stereo Audio CD with a sampling rate of 44.1khz while video streaming data is based on a DVD. Audio and video streaming data have a medium bandwidth and end-to-end delay

requirement. The bulk data is modeled with assumed 15 TCP connections between the source node and the Head Unit, where the source node could, for instance, be an antenna module managing internet connections. Bulk traffic is transmitted using TCP's flow control mechanism to dynamically use the available bandwidth. Table 1 presents a detailed overview of the different kinds of traffic.

Table 1. Traffic characteristics of the simulation

Node Name	UDP/TCP Packet Length [byte]	Service Rate [ms]	Bandwidth	Destination Node
CTRL1 ..	20	uniform(10, 100)	1.6Kbit/s - 16Kbit/s	Head Unit
CTRL4				
CAM1 ..	786	0.25	25.1Mbit/s	Head Unit
CAM3 [1]				
FCAM [1]	786	0.25	25.1Mbit/s	PU_CAM
Audio	1472	8.4	1.4Mbit/s	AVSink
Video	1472	1	11.8Mbit/s	AVSink
Bulk Traffic	1400	uniform(1,10)	1.12Mbit/s - 11.2 Mbit/s	Head Unit

4 Simulation Results and Discussion

4.1 Methodology and Assumptions

The influence of the network load on the availability of application services is evaluated based on a simulation with reduced complexity. We set up the simulation by using certain assumptions and selecting a specific scenario. This specific scenario describes the in-car network situation where ECUs are placed at the front and at the rear of a vehicle, so that the double-star based topology is a first and logical approach to analyze the influence of typical automotive applications. In this work, we use the OMNet++ simulation tool with the INET framework which contains the required networking protocols IP, TCP, UDP and Ethernet [7]. Following assumptions were made for the simulation:

1. Switch processing time: $3\mu s$ [4].
2. MAC transmission Queue size (TXQueue length): 1000 packets.
3. Static IP-address configuration.
4. Ethernet Link Bandwidth: 100Mbit/s.

4.2 Metrics

The network load in a switched Ethernet network has a strong influence to the service behaviors in an in-car network. A packet delay or packet loss of control data would negatively influence driving stability and safety which is not

tolerable. Packet delay and packet loss result from overloaded networks. We define the following metrics in order to evaluate the network, its load situation and the influence of different traffic characteristics.

- The **End-to-End Delay** is calculated by the time difference between the generated packet time at a source node and the receiving packet time in the application layer at the sink node.
- The **Inter-Arrival Times (IATs)** are determined by measuring the arrival time of two subsequent data frames at the sink. It is used to see how the network load influences the packet arrival time at the sink node. Furthermore it indicates a packet loss at the sink node.

Service Constraints. The service constraints are used in order to evaluate the switched Ethernet in-car network and to verify if the services fulfill the requirements in terms of end-to-end delays. For safety-relevant control systems (e.g. engine control system, dynamic stability control), the maximum end-to-end delay is 10ms. This value is currently used for a typical hard real-time communication in a vehicle [3]. The data transmitted by a driver assistance camera must have an end-to-end delay of 45ms [3]. In case of multimedia audio and video streaming data the latency should be less than 150ms [13]. The service constraints of different traffic types are listed in Table 2. In addition to the end-to-end delays, the IATs are compared with the service rate of each traffic class to determine the load situation on the network.

Table 2. Service constraints

Traffic Number	Traffic Type	Max. End-to-End Delay [ms]
1	Control Data	≤ 10 [3]
2	Driver Assistance Camera Data	≤ 45 [3]
3	Multimedia Audio Data	≤ 150 [13]
4	Multimedia Video Data	≤ 150 [13]

4.3 Results

End-to-End Delay. The cumulative distribution function (CDF) of the end-to-end delay by transmitting control data to the Head Unit (HU) is represented in figure 5(a) while figure 5(b) shows the CDF for camera data transmitted to the HU. In both cases, we can observe that the end-to-end delays decrease, if the number of intermediate switches along the transmission path is low. This result is expected, because the end-to-end delay increases with each intermediate switch in the entire path from the source node to the sink node. The end-to-end

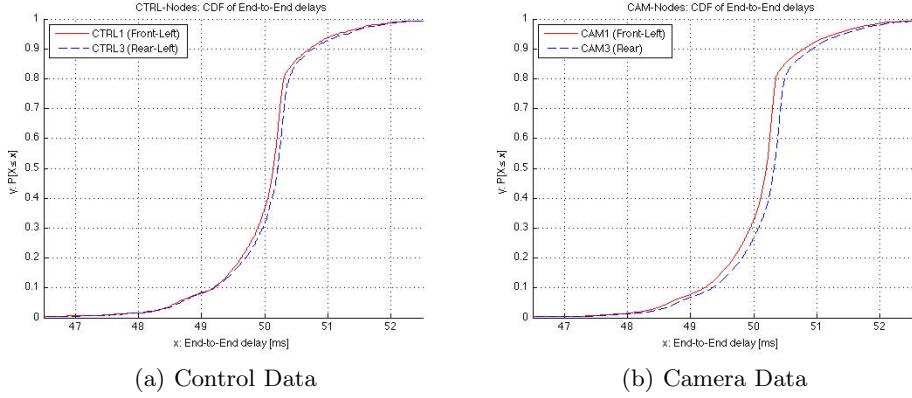


Fig. 5. CDF of End-to-End delays on the HU

delay of camera data and control data have quite similar characteristics where the CDFs vary between 47 ms to 53 ms. For both cases, the service constraints are not fulfilled and the QoS constraints of the applications are violated.

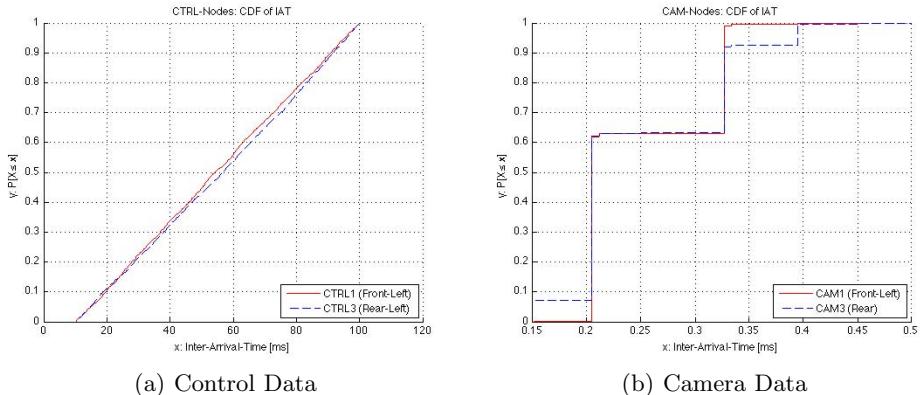


Fig. 6. CDF of IAT on the HU

Inter-Arrival Time (IAT). The service rate of the control nodes is uniformly distributed over 10ms to 100ms and we would expect that the CDF of inter-arrival times reach a probability of 1 for a value of 100ms. Figure 6(a) indicates that all control data packets are transmitted to the Head Unit so that no packet loss is occurred. In case of camera data, the situation is quite different. The camera data which are sent with a service rate of 0.25ms have an IAT of approximately 0.5ms in a worst case (see Fig. 6(b)). This would mean that half of

the transmitted data have a packet loss. Furthermore we can observe that IAT of the front CAM node (CAM1) is lower than of the rear node (CAM3). This is due to the intermediate switch ('Switch2') which has to be passed for packets generated by the rear end node.

4.4 Summary

The simulation based on a double-star topology gives evidence that the service constraints of typical automotive applications are not fulfilled. The end-to-end delay of control data is five times higher than allowed. The inter-arrival time of camera data gives further evidence for an overloaded network where service constraints of driver assistance cameras are violated.

5 Conclusion and Future Work

A switched Ethernet network without any prioritization mechanisms cannot guarantee the service requirements of the presented IP-based in-car communication scenario with a double star topology and several traffic types relevant to future in-car communication. In order to avoid violation of the highest service constraints, services must be classified and prioritized. Our research continues evaluating two different QoS mechanisms to do so: Characterizing traffic classes after IEEE 802.1Q and IEEE 802.1 Audio Video Bridging (AVB).

The IEEE 802.1Q standard defines VLAN tagging within the Ethernet header which provides a prioritization on Layer 2 and a scheduler in a switch to treat different types of packets. Four priority queues per output port per switch are configured for the next analysis: Control Data for Hard Real-Time traffic, Driver Assistance Camera Data for Soft Real-Time traffic, Multimedia Video/Audio Data for MM Streaming traffic, and Bulk data.

The second QoS mechanism for supporting QoS on Layer 2 is the IEEE 802.1 Audio Video Bridging (AVB) draft standard [9]. It specifies a latency less than 2ms over 7 hops for synchronized time sensitive streams [8]. We will look at the following mechanisms of AVB:

- IEEE 802.1AS [10]: Time synchronization of distributed Ethernet nodes
- IEEE 802.1Qat [11]: Stream reservation protocol for AVB streaming data
- IEEE 802.1Qav [12]: Scheduling, Queuing and Forwarding rules on switches

References

1. Bruckmeier, R.: Ethernet for automotive applications, Freescale Technology Forum, Orlando (2010), http://www.freescale.com/files/ftf_2010/Americas/WBNR_FTF10_AUT_F0558.pdf
2. Daoud, R.M., Amer, H.H., Elsayed, H.M., Sallez, Y.: Ethernet-based car control network. In: CCECE. IEEE, Los Alamitos (2006)

3. Rahmani, M., et al.: Performance analysis of different network topologies for in-vehicle audio and video communication. In: 4th International Telecommunication Networking WorkShop on QoS in Multiservice IP Networks (QoS-IP 2008), Venice, Italy (February 2008)
4. Rahmani, M., et al.: Traffic shaping for resource-efficient in-vehicle communication. IEEE Transactions on Industrial Informatics 5(4), 414–428 (2009)
5. International Organization for Standardization (ISO). Iso 13400 - road vehicles - diagnostic communication over internet protocol (doip) - part 1: General information and use case definition,
http://www.iso.org/iso/catalogue_detail.htm?csnumber=53765
6. Freymann, R.: Anforderungen an das automobil der zukunft. The 2nd Mobility Forum, Munich, Germany,
http://www.munichnetwork.com/fileadmin/user_upload/konfernenz/mobilitaetsforum-2/071128MUN_Prof_Freymann_Raymond.pdf
7. OMNet++ Simulation. Version 4.0, <http://www.omnetpp.org/>
8. Teener, M.J.: Worst case latency in 802.1qav ethernet bridges,
<http://www.ieee802.org/1/files/public/av-mjt-max-delay-0408-v2.pdf>
9. IEEE 802.1 AVB TG. Ieee 802.1 audio/video bridging (avb),
<http://www.ieee802.org/1/pages/avbridges.html>
10. IEEE 802.1 AVB TG. Ieee p802.1as/d7.0 - timing and synchronization for time-sensitive applications in bridged local area networks (2009),
<http://www.ieee802.org/1/pages/802.1as.html>
11. IEEE 802.1 AVB TG. Ieee p802.1qat/d6.1 - virtual bridged local area networks - stream reservation protocol (2009),
<http://www.ieee802.org/1/pages/802.1at.html>
12. IEEE 802.1 AVB TG. Ieee p802.1qav/d7.0 - forwarding and queuing enhancements for time-sensitive streams (2009),
<http://www.ieee802.org/1/pages/802.1av.html>
13. Wolf, L.C., Griwodz, C., Steinmetz, R.: Multimedia communication. Proceedings of the IEEE, 1915–1933 (1997)

Degradation of Communication Range in VANETs Caused by Interference 2.0 - Real-World Experiment

Robert K. Schmidt², Bernhard Kloiber¹,
Florian Schüttler², and Thomas Strang¹

¹ Deutsches Zentrum für Luft- und Raumfahrt (DLR),
Münchner Strasse 20, 82234 Wessling, Germany

² DENSO AUTOMOTIVE Deutschland GmbH,
Freisinger Strasse 21, 85386 Eching, Germany

Abstract. High channel load in vehicle-to-vehicle communication leads to a degradation of the vehicles' communication range, due to interference and hence packet loss at larger distances. Packet loss results from two or more concurrent transmissions, colliding at receivers located in-between, which is also known as the hidden station problem. In previous works, our simulation study has shown that this packet loss leads to a degradation of 90% of the communication range. In this paper, we confirm the simulation results by real-world measurements. We present a methodology for transferring the simulation scenario to a real-world measurement scenario, able to evaluate the problem of hidden stations. With three radios applying the IEEE 802.11p standard, we measure the degradation of the communication range under interference. In the measurement, we find a degradation of 50 to 70%. On the one hand, there are less collisions due to only one hidden station. On the other hand, we identify that the receiving vehicle as a shadowing object itself is an additional origin for hiding the other station which slightly increases the number of collisions even at close distances.

Keywords: Congestion, Field test, Interference, High load, Measurement, VANET, V2V.

1 Introduction

Enabling vehicles to communicate with each other spontaneously opens a new field for active safety applications. Vehicles equipped with wireless communication technologies according to IEEE 802.11p may exchange their status and movement. Such information on vehicles is included in Cooperative Awareness Messages (CAMs) which are periodically broadcasted to vehicles in the vicinity. All vehicles store the information in a local dynamic map of their surrounding [12]. Thus, vehicles are able to check for potential dangerous situations like collisions at intersections or an approaching ambulance vehicle.

In high vehicle density like on a multi-lane highway, the periodicity of CAMs causes an overload of the communication channel. There, packet loss occurs due

to a high number of collisions in medium access. The well-known hidden station problem becomes more likely which, even worse, prevents the carrier sensing from working properly. We have analyzed the reasons and consequences of the channel overload by means of simulation. In [1] we found that in extreme situations the communication range under interference is reduced by 90%. Successive packet loss and high load on the medium also causes a significant increase of the neighbor update delay [7].

In this paper, we compare the simulation results with the results from the real-world experiment. We firstly provide a methodology on how to simplify and transfer the simulation scenario to a basic measurement scenario. We secondly conduct the measurement and evaluate the results, followed by a comparison to the results obtained from the simulation study. The presented scenario also forms the basis for testing countermeasures against packet loss.

The remainder of the paper is structured as follows. The next section reviews the literature related to measurements of wireless communication under high channel load. In Sec. 3, the methodology for our measurement is explained. Sec. 4 shows the results of the measurement, followed by the conclusions given in Sec. 5.

2 Related Work

Weinfield describes in [3] a testbed for emulating multiple communicating vehicles. Using this testbed, he conducts experiments at different high channel load in an in-lab test. The testbed comprises 15 units each equipped with two IEEE 802.11p radios. All units are placed in a distance of 10 meters to each other. To emulate higher distances, different attenuators are attached to the radios. Both radios are used in order to emulate higher numbers of (virtual) vehicles, up to 180. In different measurements, the transmission-related parameters are varied, i.e. data rate of 6 and 12 MBit/s, transmit power of 10 and 20 dBm, packet generation rate of 5 and 10 Hz, message sizes of 300, 378 and 464 Bytes. The results show the highest increase of packet loss at the units with higher attenuation, i.e. higher emulated distance. However, in this test, the maximum attenuation is still too low to have a significant number of hidden stations as in our simulation study.

In [6] by Ramachandran et al., a measurement of the performance of many-to-many broadcast applications is conducted. A testbed of 400 small PCs with two transceivers is used for this experiment. The PCs are mounted at the ceiling of a big hall. Out of these 400 PCs, 100 radios are used at the same time to emulate a high dense vehicular network. The authors state that the results are valid only for this snapshot. Locations and hence distances between the radios are fixed. Also, the stations are close to each other so that no hidden stations occur. The experiment considers two workload scenarios: First, all stations saturate the medium by generating as much packets as can be actually transmitted. Second, all stations transmit 10 packets per second. As the authors conclude, with 100 radios and 10 packets per second they do not saturate the channel. Thus, the

Table 1. Comparison of related work

	Weinfield	Ramachadran et al.	Jardosh et al.
Protocol	IEEE 802.11p	IEEE 802.11a	IEEE 802.11b
Units physical/virt.	15 / 180	400 / 100	max. 523 / all
Scenario	Lab-test	Big hall	Multi-room conference
Communication	Broadcast	Broadcast (Ad-hoc)	Unicast (Infrastructure)
Measured values	CBT, good packets	PDR, good packets	No. of frames, acceptance delay
Workload	5 / 10 Hz	10 Hz / Saturation	typical Internet over WLAN
Packet sizes	300, 378, 464 Byte	128 Byte	0 - 1200 Byte
Duration	300s	120s	5:33h & 2:46h
Hidden stations	Few	None	Not stated, but probably many
Spatial separation	Fixed	Fixed	Unknown

packet delivery rate stays above 95 %. However, for the saturated channel, the packet delivery rate goes down to 45 %. A proper countermeasure to mitigate this degradation is to enable the packet capturing feature.

Jardosh et al. [4] investigate nearly a full day of wireless LAN communication at a conference. Various channel loads and packet sizes have been logged and analyzed. They find that under high load, RTS-CTS cannot guarantee fair medium access and should be avoided. Higher data rates should be used to reduce the transmission time. However, it is difficult to compare the results of Jardosh et al. as they use IEEE 802.11b in unicast mode with access point association.

To summarize and compare the related work, Tab. II highlights the differences between these experiments. All experiments mainly focus on in-lab tests and thus non-realistic channel conditions for vehicle-to-vehicle communication. The static experiments highlight the packet loss under high load, mostly using broadcast communication. However, due to the experiment setup only few hidden stations occur. No attention is paid to the variation of the spatial separation of the transceivers and their interrelation.

In our real-world experiment, we explicitly want to have all effects of signal propagation occurring outdoors, as this will be the case in the real-system on the road. We focus on the spatial separation of two stations potentially leading to hidden stations and provoking excessive packet loss at certain distances as found by our simulation study in [1].

3 Methodology

Following, we review the application scenario of the simulation study. Based on the road traffic scenario in the simulation, we explain the procedure to simplify and transfer the simulation scenario to the measurement scenario. Finally, we describe the hardware and software equipment being used for the measurement and its setup.

3.1 Application Scenario

For the simulation study in [1], we designed a scenario which involves many special communication aspects of vehicle-to-vehicle communication: High velocity, quickly changing vehicle density resulting in a severe hidden station scenario.

Such a situation can easily happen when an ambulance vehicle (AV) approaches a slowly moving traffic jam, i.e. an area of high vehicle density. We analyze the packet receive ratio at the very end of the traffic jam, i.e. the tail-end vehicle, with respect to the packets sent by the AV.

All vehicles apply a static (high) CAM rate. Due to the spatial separation, the transmissions of the vehicles within the traffic jam are partially hidden to the AV. Hence, there is a high likeliness of having two successive and interfering transmissions leading packet loss at the tail-end vehicles. In the end, the tail-end vehicles cannot receive the packets from the AV. We will review the simulation results when we compare them with the measurement results in Sec. 4.2.

3.2 Scenario Transfer

As real-world measurements usually involve an uncontrollable environment compared to the simulation, we carefully select the aspects that should be considered for the scenario transfer, guided by the following questions:

1. What are the most significant factors and aspects that cause the degradation and hence need to be transferred?
2. What can be eliminated from the simulation scenario to have a more controllable environment?

Transfer - Most important for the transfer is the variation of the spatial separation of the transmitter to receiver and interferer. We selected a sufficiently large scenario at a private airport next to DLR in Oberpfaffenhofen (Munich). It offers a 2286 x 45 meter runway mainly in free-space environment with nearly no reflecting objects. Most hangars are covered with grass which will scatter the signals rather than reflect them. We transferred the spatial separation as shown in Fig. 1. Similar to the simulation, we fixed the receiving vehicle's location (light grey circles, R3 … R1) where the measurements log files are created. As a high

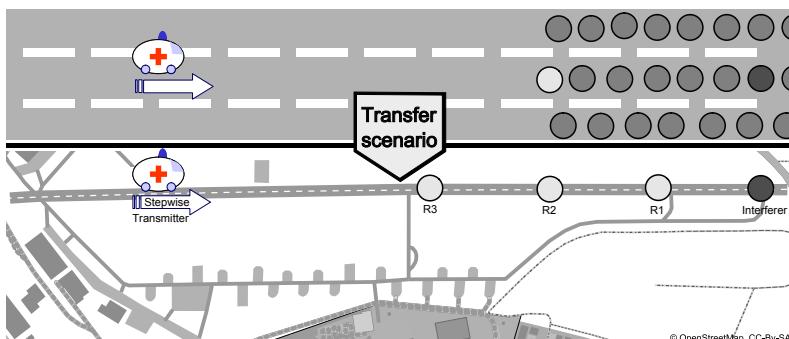


Fig. 1. Sketch of the spatial scenario transfer from simulation to real-world

number of interfering vehicles is very difficult to setup and control, we decided to let one vehicle fully utilize the channel (dark grey circle). This interfering vehicle I generates as much packets as actually can be transmitted with respect to the MAC protocol. With this high load, it is expected to get the same result of communication range degradation as for the highest CAM rate, i.e. 90%.

Compared to the simulation, it is very difficult to achieve the same communication range with different vehicles. Due to roof curvature, antenna type and antenna position, the achieved range can vary significantly. Hence, we conducted a rough calibration of the transmit power prior to the measurement to achieve the same communication range as in the simulation. Details on that will be described in the following subsection.

Elimination - Due to the difficulty to drive vehicles in a reproduceable manner, we decided to avoid movement while measuring. Measurements are done only when all vehicles are standing still. This also avoids different antenna gains while the vehicle is pitching, due to accelerating or decelerating. After each measurement, the transmitting vehicle T is moved stepwise. The interfering vehicle always remains at the end of the runway. The receiver is moved after accomplishing all transmitter position. This is in line with the simulation setup, where no movement-dependent signal propagation effects like doppler shift are modeled. As the simulation results rely on the free-space model, reflections and multi-path effects are also eliminated from the real-world scenario.

3.3 Communication Range Calibration

To be able to compare the results of the measurements with the results of the simulations, the transmission ranges are calibrated to approximately 1000 m as in the simulation. Both, interferer and transmitter, are calibrated against the receiver. The interferer is placed at the end of the runway, the receiver is placed at a distance of 1000 m to the interferer. Next, the transmission power of the interferer is adjusted, so that the receive ratio at the receiver is approximately 90% which we assume as the edge of the communication range. The same procedure is done for transmitter and receiver. As a result, the transmitter applies 18 dBm, the interferer applies 12 dBm. The significant difference between the calibration results is supposed to be due to antenna characteristics, roof curvature and elevation angle differences. Fig. 2 depicts the elevation profile¹ of the runway which poses slight angle differences in the elevation angle.

3.4 Equipment and Setup

Three cars have been used for the experiment, each equipped with the hardware and software as described in the following.

¹ The elevation profile of the runway has been created with a barometer-equipped GPS receiver.

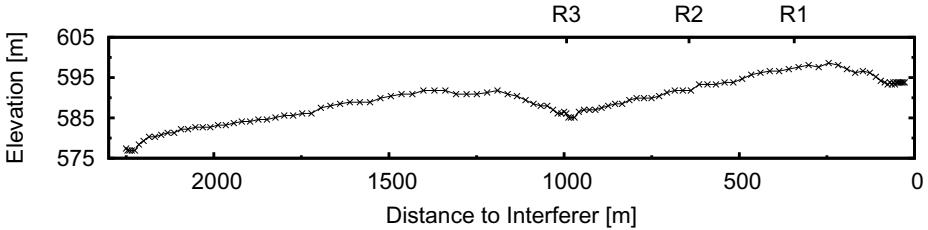


Fig. 2. Elevation profile of the runway

- *Communication unit*: The basic component is the WSU (Wireless Safety Unit) [2], which is an integrated prototype platform implementing the IEEE 802.11p standard [9]. All log data is written to a compact flash card installed at the receiving vehicle.
- *Antenna*: An antenna has been connected to each WSU by using a 3 m long antenna cable. The antennas have been mounted centered on the roof of each car. The transmitting and the interfering car have used a monopole antenna by Nippon Antenna having a gain of 0 dBi in the horizontal and 5 dBi average gain in an elevation of 15°. The receiving car has been equipped with a dipole antenna by Mobile Mark which has a gain of 9 dBi² in omni direction.
- *Positioning*: Each WSU is connected with a Garmin GPS18-LVC receiver. The GPS receivers have been placed outside of line-of-sight to the communication direction, i.e. not in the center line of the roof.
- *Logging*: The test application used for the measurements is the WSU Test Application (WTA) [2]. It enables the user to test the performance of the communication channel and the MAC protocol. Furthermore, it configures the radio parameters, executes packet tests, displays real-time statistics, and logs key metrics. A screenshot of the GUI is shown in Fig. 3.
- *Vehicles*: Transmitter and interferer vehicles are compact cars with a round roof curvature. The receiver vehicle is a big offroad vehicle with a flat roof.

Using the WTA, the radio is configured as summarized in Tab. 2. The data rate is set to 6 Mbit/s which was found to be optimal for VANETs [10]. The transmitter is configured to send 10 packets per second which is a usual message rate. By using the continuous mode of the WTA, the resulting channel busy time at the interferer is approximately 84-86 %.

As shown in Fig. 1, measurements are done at three different receiver locations, i.e. distances $R - I$ of 1000 m (R3), 650 m (R2) and 350 m (R1). For each receiver position, the transmitter starts at a distance of 900 m to the receiver and approaches in steps of 50 m. We did not start the measurement at the calibrated communication range to eliminate the fluctuations of packet loss at the border of the communication range. For each step, the receiver logs the packets for 5

² Due to the flat roof of the receiving vehicle, we used a high gain bar antenna.

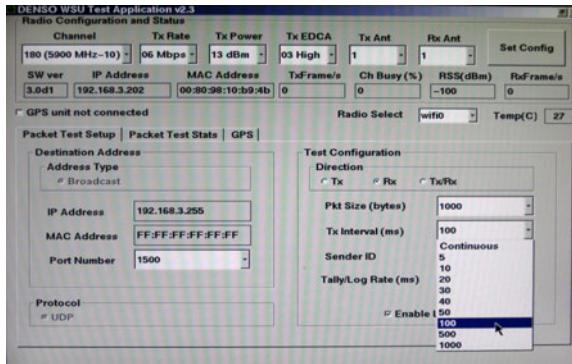


Fig. 3. Screenshot of the WTA showing the configuration of radio and packet test parameters

Table 2. Overview of the measurement parameters

Fixed Parameter	Value
Number of vehicles	3
Channel	180 (5900 GHz)
Receiver diversity	Off
Transmit power (Transmitter/Interferer)	18/12 dBm
Maximum communication range	\approx 1000 m
Data rate	6 MBit/s
Packet length (payload)	1000 Bytes
Packet rate (Transmitter / Interferer)	10 Hz / Continuous
Measurement length	5 seconds
Number of runs	2
Antenna gain - Nippon Antenna ($0^\circ/15^\circ$)	0/5 dBi
Antenna gain - Mobile Mark	9 dBi
Varied Parameters	Values
Interferer enabled	true, false
Distance $R-I$	1000 m, 650 m, 350 m
Distance $T-R$	900, 850, ..., 50 m

seconds. The logging procedure for each step is done twice, once with interferer switched off and once with interferer switched on. After the first run (i.e. one complete approach of the transmitter to the receiver for each receiver position) is finished, a second run is started with the same configuration parameters.

4 Measurement Results and Comparison

The results of the measurements are shown for each receiver location separately and afterwards summarized. In the last subsection, the trends of the communication range degradation are compared between the measurement and the simulation.

4.1 Measurement

In the following, we present the results for the packet receive ratio at the receiver from the transmitter. The average values per run are connected with lines for the sake of readability. Note that the lines do not represent actual measurements. Single crosses or stars indicate single results for the packet receive ratio, i.e. the number of received packets per one second divided by the number of transmitted packets by the transmitter.

It is expected that the number of collisions drops significantly when the transmitter is closer to the receiver than the interferer so that the signal-to-interference-ratio (SIR), i.e. T's signal strength over I's signal strength, is sufficiently high. The packet capturing feature also allows to recover the stronger one from two colliding transmissions, even if it arrived after the reception of the weaker signal has already been started.

Receiver-Interferer distance 1000 meters - R3. When receiver and interferer are at the edge of communication range to each other, the received signal strength of the interferer is quite low. As the transmitter approaches the receiver at a distance of 600 meters, a first increase above one packet per second on average can be seen in Fig. 4. At 500 meters, the packet receive ratio is within a transient area. The SIR of both signals is in the order of the one needed at the receiver. 100% of the transmitted packets are received at the receiver at distance 450 meters. This remains nearly stable as the transmitter gets closer, except for single outliers at 350 and 200 meters.

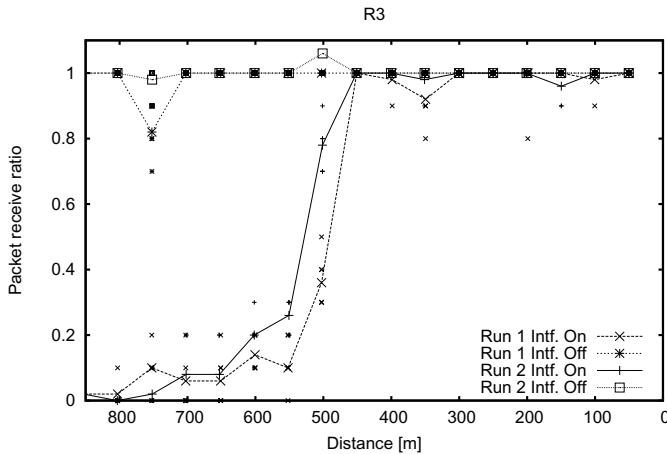


Fig. 4. Packet receive ratio - 1000 meters

Receiver-Interferer distance 650 meters - R2. At R2, receiver and interferer are within communication range to each other. Compared to R3, the results in Fig. 5 present that the transmitter needs to approach the receiver closer to be within communication range under interference. The transient area starts at

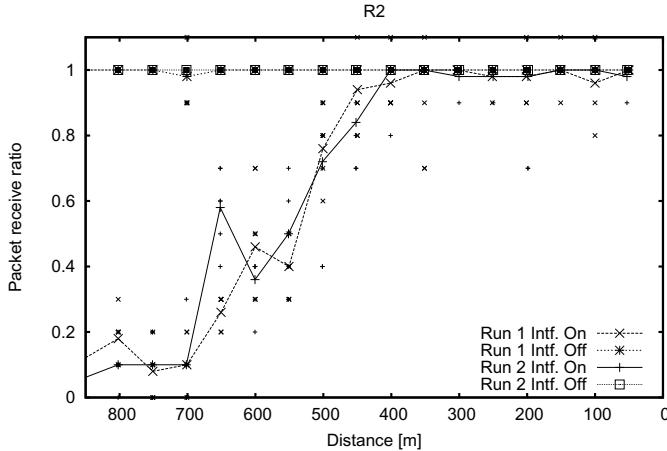


Fig. 5. Packet receive ratio - 650 meters

around 700 meters and stretches till 400 meters. Similar to R3, there are also some (more) outliers located at 350 and 200 meters. As receiver diversity is disabled, these similar results could be due to the ground reflection, canceling the signals at certain distances (i.e. following the two-ray-ground model).

Receiver-Interferer distance 350 meters - R1. The results for R1 depicted in Fig. 6 exhibit a much larger transient area of packet loss. The average value of the receive ratio is fluctuating much more and does never reach 100 %. As the signal strength of the interferer is relatively high at R1 compared to R2 and R3, the impact of the interferer is much more significant. As in R3 and R2, we also see outliers at 350 meters. After reviewing the simulation results in the following subsection, we will summarize the results for all receiver locations and consider the measurement for the channel busy time which will reveal more details about the above mentioned fluctuation.

Due to difficult weather conditions and increased air traffic, the measurements had been stopped after the first run of R1. However, compared to the other receiver locations, we assume to have a valid measurement for the receiver location R1.

4.2 Simulation

Fig. 7 depicts the simulation results published in [1]. All vehicles periodically broadcast packets of size 1000 Bytes. The message rates in the plot are varied 15, 18, 20 and 40 Hz. For the sake of readability, these rates are selected for visualization as they show the low load, medium load and heavy overload on the communication channel. As seen in the figure, the packet receive ratio decreases when the message rate is increased from 15 to 18 Hz, especially at larger

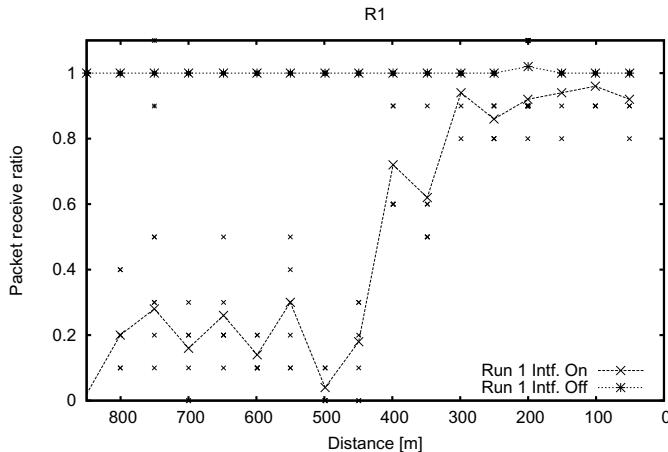


Fig. 6. Packet receive ratio - 350 meters

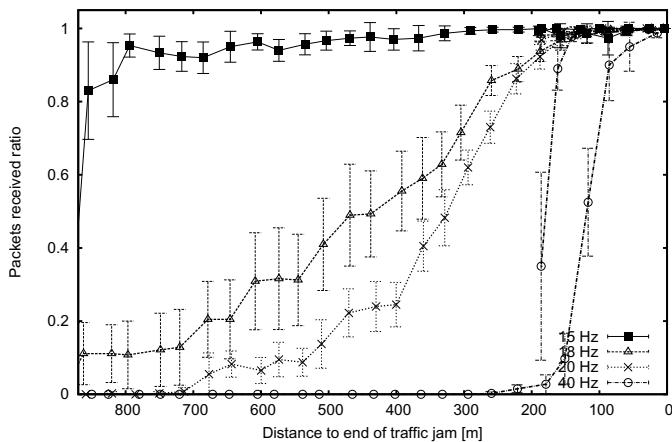


Fig. 7. Simulation result - Packet receive ratio depending on message rate and distance of approaching vehicle and tail-end vehicle

distances between tail-end vehicle and the ambulance vehicle. However, with further increase of the CAM rate (40 Hz) it converges, presenting a steep increase of the packet receive ratio at a distance of 100 meters. Hence, the communication range under interference is about 10% of the original communication range.

Summary and Comparison. At each receiver location, we have seen a significant degradation of the communication range between the transmitting vehicle and the receiving vehicle. To summarize and better understand the results, we also consider the interaction between the transmitting vehicle and the interfering vehicle with respect to the MAC protocol. Assuming roughly symmetric

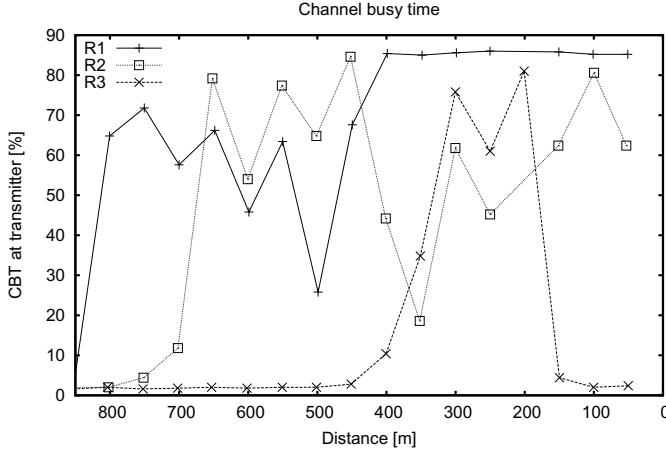


Fig. 8. Channel busy time at transmitter when interferer is flooding the channel

signal propagation between them, we can compare the results with the Channel Busy Time (CBT) at the transmitting vehicle when the interferer is switched on, shown in Fig. 8.

The graphs of the CBT provides two aspects:

1. Synchronization of medium access between transmitter and interferer
2. Shadowing by the receiving vehicle.

The receive ratio at R1 increases at around 400 to 300 meters to the receiver. Adding up the distance between receiver and interferer roughly equals to the calibrated communication range. From this distance on, the MAC protocol has a chance to sense the transmissions of the interferer and avoid from medium access. At the same time, the interferer also senses the transmissions of the transmitter and also refrains from medium access so that the distributed medium access protocol is able to reduce the number of collisions.

Similar trends are observable for R2 and R3. The graph for R2 increases at around 650 to 550 meters, for R3 the increase is at around 800 to 700 meters. However, following the graphs from left to right present significant dips. During the measurements, we log the distance where line-of-sight is significantly blocked by the receiving vehicle. The values are ≈ 150 m for R3, ≈ 300 m, and ≈ 700 m for R1. For R1, this confirms the dips at 600 and 500 meters. Beyond that, the CBT goes to maximum which means that the signal attenuation due to shadowing is not significant. The transmitter is already quite close to the interferer. R2 has an even worse dip at 400 and 350 meters, followed by scattering results around 50 % CBT. The worst shadowing is experienced for R3. The CBT drops much below 10 % due to the shadowing. The signal is already strongly

attenuated due to the large distance between interferer and receiver. Shadowing typically adds up 10 to 20 dB which then makes it impossible to sense any transmission at this distance.

When we now look back to the packet receive ratio, lower ability to sense packets due to shadowing has only few impact at lower distances as the SIR is sufficiently high. The graph in R3 has the same steep increase as in the simulation. For R2, the increase of the receive ratio is less steep. Some single measurements show a loss of 10 % at lower distances (400 to 50 meters). Most of the fluctuation in the results is seen in the graph of R1. In this case, the better signal-to-interference ratio is an advantage for the interferer, especially due to packet capturing, more packets of the transmitter get lost.

The measurement showed a lower degradation than the extreme case in the simulation. Although the channel load is at maximum in both cases, the (reduced) number of interfering vehicles has a strong impact on the degradation. In the simulation, we found that the carrier sensing is not working properly in extreme high load situations. As the measurement comprises only three vehicles, the carrier sensing is mainly interfered by one hidden station at a time. The problem of many hidden stations leading to multiple interfering signals is not accounted in the measurement scenario.

The hidden stations are more wide spread in the simulation while they are isolated in the measurement. Hence the less steep increase in the transient area.

5 Conclusions and Outlook

Active safety applications in VANETs rely on receiving updates from neighboring vehicles frequently. Under high channel load, the packet error rate can significantly increase at higher distances between receiver and transmitter. In our previous work [1], we analyzed and quantified this degraded communication range due to interference.

We confirm in this paper the trend of communication range degradation under interference by a real-world experiment with a basic measurement setup of three vehicles. As a result, we already determine a degradation of 50% to 70%. We further find that due to shadowing of the vehicles itself, additional packet loss occurs. The other transmitter becomes hidden due to significant signal attenuation in the shadow of the bigger vehicle. While in this shadow, another vehicle cannot correctly sense that the channel is busy.

The advantage in signal-to-interference ratio is supported by packet capturing which improves the communication range under interference. All in all, even under high channel load the communication range under interference is sufficiently large for time-critical safety applications operating in a range of 50 - 300 meters. No significant degradation of the packet receive ratio is found in this area.

As a next step, we plan to employ multiple interfering vehicles to be placed in the field, having a high CAM rate instead of flooding the channel. For our future work, we also plan to use the presented scenario as a basic validation scenario for Decentralized Congestion Control [8].

Acknowledgements

The authors would like to thank EDMO-Flugbetrieb GmbH [11] for the opportunity to run the experiments under almost ideal radio propagation conditions and for their safe coordination of accessing the runway alternating with normal air traffic.

References

1. Schmidt, R.K., Köllmer, T., Leinmüller, T., Böddeker, B., Schaefer, G.: Degradation of Transmission Range in VANETs caused by Interference, Praxis der Informationsverarbeitung und Kommunikation (PIK). Special Issue on Mobile Ad-hoc Networks 4 (2009)
2. Weinfield, A., Graham, S.: Methods to Reduce DSRC Channel Congestion and Improve V2V Communication Reliability. In: ITS World Congress, Busan, Korea (2010)
3. Weinfield, A.: Wireless Vehicular Safety Systems: Specialized Test Platform Supports Application Evaluation. In: ITS World Congress, Beijing, China (2007)
4. Jardosh, A.P., Ramachandran, K.N., Almeroth, K.C., Beldingroyer, E.M.: Understanding congestion in IEEE 802.11b wireless networks. In: Proceedings of the 2005 Internet Measurement Conference (2005)
5. Thorpe, C., Murphy, S., Murphy, L.: Analysis of Variation in IEEE802.11k Channel Load Measurements for Neighbouring WLAN Systems. In: ICT Mobile and Wireless Communications Summit (ICT-MobileSummit 2008), Stockholm, Sweden, June 10-12 (2008)
6. Ramachandran, K., Gruteser, M., Onishi, R., Hikita, T.: Experimental Analysis of Broadcast Reliability in Dense Vehicular Networks. In: VTC Fall (2007)
7. Kloiber, B., Strang, T., de Ponte-Müller, F., Garcia, C.R., Röckl, M.: An Approach for Performance Analysis of ETSI ITS-G5A MAC for Safety Applications. In: Proceedings of the The 10th International Conference on Intelligent Transport Systems Telecommunications (ITST), Kyoto, Japan (2010)
8. Schmidt, R.K., Brakemeier, A., Böddeker, B., Schäfer, G.: Architecture for Decentralized Mitigation of Local Congestion in VANETs. In: Proceedings of the The 10th International Conference on Intelligent Transport Systems Telecommunications (ITST), Kyoto, Japan (2010)
9. Institute of Electrical and Electronics Engineers, IEEE802.11p/D9.0 Wireless Access in Vehicular Environments, Draft 9.0 (November 2009)
10. Maurer, J., Fugen, T., Wiesbeck, W.: Physical Layer Simulations of IEEE802.11a for Vehicle-to-Vehicle Communications. In: Proceedings of the 62nd IEEE Vehicular Technology Conference, VTC 2005 (2005)
11. EDMO-Flugbetrieb GmbH, <http://www.edmo-airport.de/index.html>
12. SAFESPOT, Deliverable D3.3.3 - Local dynamic map specification, Technical Report (2008)

Real-World Measurements of Non-Line-Of-Sight Reception Quality for 5.9GHz IEEE 802.11p at Intersections

Thomas Mangel¹, Matthias Michl¹, Oliver Klemp¹, and Hannes Hartenstein²

¹ BMW Group Research and Development
Hanauer Straße 46, 80992 München, Germany

{thomas.mangel,matthias.michl,oliver.klemp}@bmw.de

² Karlsruher Institut für Technologie (KIT)
Geb. 20.21, Zirkel 2, 76131 Karlsruhe, Germany
hannes.hartenstein@kit.edu

Abstract. Vehicular Dedicated Short Range Communication (DSRC) promises to reduce accidents by enabling assistance systems such as cross-traffic assistance. This application requires movement information from vehicles that are in Non-Line-Of-Sight (NLOS) due to buildings at intersection corners. DSRC is foreseen to use IEEE 802.11p to deliver regular Cooperative Awareness Messages (CAM). Due to the high operational frequency of 5.9GHz, the ability to provide reliable NLOS reception is often put into question.

We performed an extensive field test specifically targeted to measure DSRC NLOS reception quality. As a novelty, tested intersections were methodically selected to reflect typical urban intersections (in Munich) and the test setup was specifically designed to provide comparable and generalizable results. This allowed us to determine the influence of factors like building positions. The collected data shows that NLOS reception is possible. Reception rates stay mostly well above 50% for distances of 50 meters to intersection center with blocked LOS.

Keywords: Car2Car, 802.11p, Non-Line-Of-Sight, NLOS, Testing.

1 Introduction

Vehicular safety communication has the potential to increase range and coverage of location and behavior awareness of vehicles, thus enabling highly developed pro-active safety systems such as e.g. cross-traffic assistance [1][2] at intersections. The idea is that each vehicle communicates information like position, speed and heading periodically (10 Hz are often assumed) to other vehicles in CAMs to enable the deduction of a highly accurate environment picture, used as basis for movement prediction. As communication technology, the WLAN based DSRC technology is developed, targeting ad-hoc communication based on 10 MHz wide frequency bands centered around 5.9 GHz in Europe. Medium and physical access is standardized as IEEE 802.11p [3].

For this radio technology, cross-traffic assistance at inner city intersections is one of the most challenging use cases. It needs to monitor two spatial dimensions and is very sensitive to heading estimation inaccuracies [2], thus implying the need for high information update rates. At the same time, the corners at inner city intersections will often be occupied by buildings. They can block the radio Line-Of-Sight (LOS). If the LOS is blocked, diffraction, reflection and refraction of radio waves can enable NLOS reception. But, the relatively high frequency of 5.9 GHz and a difficult radio fading environment might complicate the NLOS reception of packets.

An analysis [4] about building locations at intersections in the Munich administrative area verified the need for NLOS reception: If two cars are approaching an intersection with 50 km/h, only 30% of all intersection corners provide line of sight at 3 seconds to intersection center. Here, 3 seconds is often considered as a desired warning point before a potential collision [5,6].

While Road Side Units (RSU) or vehicles located in the intersection center might re-broadcast messages and reduce the need for NLOS reception, we still believe it is needed: Particularly urban street crossings with sparse traffic patterns are likely not to be equipped with dedicated RSU equipment. Such scenarios require robust signal transmission between the vehicles themselves which enter the crossing. These situations predominantly exhibit NLOS radio link conditions between the vehicles. This motivates to investigate NLOS reception.

Despite the existence of first investigations on 5.9GHz NLOS reception [7,8,9], it seems as if there is still a lack of in depth understanding about the expectable NLOS reception quality. For example, a verified NLOS channel model for simulations is still missing. Investigations performed so far lack a founded test site selection as well as a test setup that allows a generalization of results. In consequence, the general ability of proper NLOS reception is still often put in question by people dealing with DSRC.

To gain more insight, we performed an extensive field test, specifically targeted to measure the general ability of DSRC to provide NLOS reception and its quality in doing so. Hereby, we try to investigate the impact of specific parameters on the radio link such as building type or inter-building distance. We also intended to produce results that can serve as a basis for a future derivation of a channel model for packet level simulations in ns-2 [10] or similar simulators. To achieve these goals, we put special attention to a well-founded selection of representative test cases and to find a test setup that allows the derivation of predominant influence factors and to generalize results.

2 State of the Art

NLOS DSRC radio channel field measurements with channel sounders were performed in [7]. The results indicate that NLOS reception should generally be possible. Channel sounders measure the received energy impulse response over time. It potentially allows to identify different propagation paths and provides good inside in low level characteristics of a received signal.

First NLOS measurements with DSRC radios [8][9] show that a reception behind buildings is generally feasible with standard hardware.

In cellular communication research, algorithms were proposed to compute the NLOS path-loss for below-rooftop base stations [11][12][13]. As such base stations are located inside street canyons—for example at signal light posts—radio conditions should be similar to DSRC. Parameterized for DSRC, these models imply that NLOS reception should work fairly well. The equations also show that geometrical aspects—such as the inter-building distance—should influence reception quality. Such influence factors have not been investigated in measurement campaigns so far.

There exist first attempts to model DSRC NLOS reception for packet level simulations. In [9], one of the above mentioned cellular models was adopted by using DSRC parameters. In [14], path-loss characteristics of different of these models were evaluated with respect to vehicular communications. A unified model is proposed. However, a validation based on DSRC measurements is still missing.

The testing work performed so far especially lacks a profound scenario selection. In [7][8], different scenarios were specified based on an intuitive selection. Without a statistical decision basis, it is difficult to say if a representative intersection was tested or not.

There also exist no measurements with a defined test environment. Mostly, two moving cars are used. This leads to a difficult spacial result post processing and interpretation. Comparability between scenarios is complicated, as there is almost no possibility to recreate the same movement in different environments.

3 Test Setup

As can be seen from Section 2, there are in general two possibilities to measure a wireless channel. One way is a channel-sounder setup providing impulse response data and the other way to use off-the shelf radio hardware that can provide a signal power indication for received packets.

Using a radio, it is possible to evaluate information about packet reception success and signal power indication. The latter is only available for received packets. Channel-sounder have a higher sensitivity than radios and are therefore able to provide information for signals not receivable by a radio. They also provide a much more detailed view of the channel. However, it is difficult to combine the measured channel sounder output with a numerical model of signal processing at the receiver to conclude the real reception performance of a radio.

Another benefit of using a radio is the fact that the provided results are close to channel models as they are used in packet level simulators such as ns-2. These simulators compute the arriving signal power for one transmitted packet once for each possible receiver. Therefore, the simulation inherits the same per packet simplification as the reception power value from a radio.

Because the ability to deduct a channel model for packet level simulations was one of the main objectives and reception rate is one of the major performance metrics for applications, we selected to test with radios.

Another major goal was to get results that provide comparability between different intersections and even different intersection legs. If both, transmitter (Tx) and receiver (Rx) are mounted on driving vehicles, it is very difficult to control the distance between transmitter and receiver during testing and evaluation; certainly the most important parameter on reception power. Distance between transmitter and receiver also interferes in such a setup with the distance to intersection center, a value that partly defines the NLOS amount of the radio link. Furthermore, for comparability between intersections, exactly the same vehicle coordination would be needed.

We solved these problems by discretization—therefore fixing—of the transmit position. We selected two distances to the intersection center as transmit positions. For each position, a drive through on the crossing street is performed in both directions, producing continuous results in the receiver to intersection center dimension. A certain transmitter distance to center was tested in every side street of an intersection. This way, we were able to produce results that allow comparing the reception performance between different side streets and also different intersections.

The two selected distances to transmitter are 30 and 60 meters. They shall reflect two different situations in terms of cross traffic assistance: A warning is often considered to be triggered at 3 seconds before a collision might occur [56]. The last point of interception at 1 second. This corresponds to 42 and 14 meters at a speed of 50 km/h. Because the halt line is probably a more likely point of collision as the intersection center, and line of sight into a cross street is common at 15 meter, we increased the numbers to 2 and 4 seconds. With rounding applied, this leads to the selected 30 and 60 meters to intersection center.

Splitting the distance between transmitter and receiver up into the two dimensions “distance from transmitter to intersection center” and ”center to receiver” was inspired by the NLOS reception models for cellular communication with below-rooftop base stations. One of these models [1], computes path loss based on these two values. The split-up also reflects that a signal has to travel a longer distance in case of NLOS propagation than the direct distance between transmitter and receiver.

3.1 Used Hardware

Transmitter and receiver unit consisted in our setup each out of a radio, antenna, antenna-cable, a GPS receiver to determine position and time and a laptop that initiates transmissions respectively processes receptions and logs events.

We used a tripod as transmitter. It has a 35x35cm metal plate at 1.45m height as ground for the antenna. A tripod was favorable against a vehicle as it is better placeable at the fixed transmission distances while testing. Also, in some streets, a car would have blocked all traffic. Transmit positions were pre-determined via satellite images. Because of its drifting, GPS was not accurate enough.

The receiver vehicle was a BMW 5 Series GT with a roof height of 1.56m. Position was taken from the vehicle GPS via its CAN-Bus. This position is map matched, which means that a digital map and data like wheel spin and steering

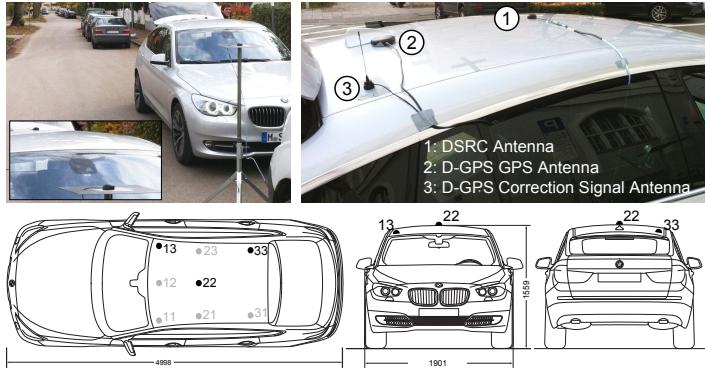


Fig. 1. Up-Left: The BMW 5 Series GT and the Tripod - Right: Antennas on the Roof. Lower Part: Antenna Positions. 13+22+33 Have Been Tested.

are also taken into account. The map matched position was much more accurate in urban canyons than a D-GPS. Especially, there is no drift if the car stands still. As GPS time was not available from the car, a D-GPS was still installed.

We used a LinkBird-MX V3 [15] communication box by NEC. The box contains two DCMA-82-N1 mini PCI cards with Atheros 802.11 radio chips. Each radio is equipped with two antenna jacks, providing switched antenna diversity. To be able to match results to antenna position, diversity was disabled and in consequence only one antenna was used.

To resemble a close to production system, we used small low-profile puck antennas from Nippon Antenna. They provide a gain of 0dB at horizon and +5dB at 15° according to their data sheet. The antenna cables inherit a loss of 2.4dB according to their data sheet. To find a near optimal antenna position, we tested three different positions in free-space conditions (see Section 4.1). The tripod, vehicle, antennas and tested antenna placements can be seen in Figure II.

3.2 Intersection Selection

A profound intersection selection is crucial to generate reasonable results. All of the NLOS models for below-rooftop base stations in the cellular domain take geometrical aspects into account. One of the major influence factors in the models is the inter-building distance (\equiv side street width) on the transmitter and receiver side street. The general assumption is: the smaller this distance, the less power in NLOS areas. This also implies that intersections with similar street width should show similar reception characteristics. We consider it worthwhile to verify this observation by selecting intersections with the same, as well as differing width.

Probably most important, selected intersections should reflect "typical" inner city intersections to enable a generalization of results. The scenario analysis in [4] shows that $\approx 90\%$ of all intersections in Munich have buildings at all corners. We focus on 4-leg intersections, as 3-leg intersection would require a result separation

between the straight street and the 3rd leg. Also, 4-leg intersections provide four LOS blocking buildings compared to only two at 3-leg intersections.

The most prominent cluster (50% of all intersections) in [4] showed a distance from center to building vertex of 17.8 meter. This translates to 25m inter-building distance in side streets. As less width means more complicated conditions for the radio link, and worst case conditions are from an application point of view important, we decided to select our "Main Case" as intersections with all side streets having an inter-building distance between 20 and 25 meters. For comparability reasons, we favored a street layout with $\approx 90^\circ$ angle between side streets. Testing multiple intersections with the same layout should show if the shape has most influence or if building facades and similar unique factors are more important.

We used the data basis from the analysis paper [4] to automatically find intersections with the desired shape and side street width in Munich. The resulting intersections were manually inspected via satellite images and then categorized in urban and suburban. As the classification is based on nearest building vertex, and front gardens (with bushes and trees) can further limit the field of view, we suspected suburban intersections to provide worse reception conditions at similar distances into the crossing street. To be able to check the influence, we selected multiple intersections from both categories: 3 suburban and 2 urban ones with the "Main Case" dimensions, identified by ID 1,2,3 and 10,11 in Table II.

To further test the influence of street width, we selected two more urban ones with larger distances (ID 20,21). One with slightly increased distances of 30 meter on each side street (ID 20). The other with 30 meter in one direction and really wide 55 meter in the other (ID 21). This intersection also has no real orthogonality at one leg, thus potentially showing the influence of a more acute angle. At this intersection, we tested with differing 60 and 100 meter transmit distance. At 30 meter, there was as good as line of sight into the side streets.

Table 1. Selected Intersections

ID	Streets	Center Lat/Lon	Inter Build- ing Distance	Description
1	Pommern-Konstanzer	48.184138,11.560239	23+21m	Suburban (Main Case)
2	Himmelschlüssel- Josef-Seifried	48.194885,11.531957	23+19m	Suburban (Main Case)
3	Tizian-Kratzer	48.161666,11.524957	\approx 21m	Suburban (Main Case)
10	Agnes-Teng	48.157592,11.568872	24+21m	Urban (Main Case)
11	Perlacher-Untersberger	48.110465,11.585836	27+21m	Urban (\approx Main Case)
20	Klug-Waisenhaus	48.163980,11.528899	\approx 30m	Urban Wide
21	Hanauer-Gneisenau	48.178500,11.529170	55+30m	Urban Very Wide, Non-Regular Shape
9	Gotebold-Driesch	48.179439,11.441817	18m	Suburban - Buildings at Only Two Corners
100	Free Space, Country Road	48.247295,11.347764		One Street, No Buildings, No Trees etc.

Last, we selected a potential worst case scenario: A tight suburban intersection with no buildings at two corners (ID 9). In result, there are major reflection areas missing for the DSRC signal. However, such a scenario is not common.

A free space test site was needed for antenna and plausibility testing. We selected a 1.8 km long country road with plowed soil right and left. Despite one small spot to one end, there were no bushes or trees within 200m at both sides.

In summary, the intersections can be classified in three major categories plus the worst case: Suburban "Main Case", Urban "Main Case" and Urban "Wide". All tests were performed between mid October and mid November 2010 under dry conditions. Trees still had most of their foliage. All intersections were tested under daytime traffic conditions.

3.3 Parameter Selection

As a result of the described setup, we get test parameters and dimensions as shown in Table 2. Tx-Distance has been selected as 30 and 60 meter as argued above. We also tested the LOS conditions in all intersections (Tx-Distance=0m), leading to three different values in total.

We are able to test different combinations of transmission power, rate and payload in one run by switching through combinations in each second. Tests in intersection 1 with 20/10dBm and 3/6Mbps showed that data rate has no influence on NLOS reception. Only a slightly higher signal to noise ratio is needed for reception with 6Mbps. Average reception power levels turned out the expectable 10 points lower with 10dBm transmit power compared to 20dBm. To get high sample numbers per distance, we constrained testing on a 20dBm/3Mbps configuration in all other intersections. The selected transmission power of 20dBm is a limitation of the radio. It has a maximum output of 21dBm. We used Channel 180 with 5.9GHz center frequency.

Table 2. Test Configuration

Tx-Power	20 dBm
Tx-DataRate	3.0 Mbps (BPSK Modulation)
Channel	10 Mhz, Number 180 (5.9 GHz Center Frequency)
Packet Size	200 Byte Payload + IP/MAC/Phy Headers
Tx-TransmissionRate	100 Hz
Tx Dist. to Intersection Center	0, 30, 60m (+ partly 100m)
Tx Lane	1 to 4 (Street Legs) + 0 (Middle of Intersection)
Drive Direction	Two Directions per Transmitter Position
Communication Box	NEC LinkBird v3 (Atheros Chipset)
Antenna	Nippon DSRC Puck, Horizon: 0dB Gain, 15°: +5dB
Rx Antenna Position	Roof Center (ID:22 in Figure 1)
Cable (Box to Antenna)	SUCOFLEX.104, 4m, Data Sheet: 2.4dB Loss
Transmitter	Tripod, 35x35cm Metal Plate at 1.45m Height
Receiver	BMW 5 Series GT, no Sunroof
Intersections	4 Suburban + 4 Urban + Free Space

Packet transmission rate was set to 100Hz and payload to 200Byte in all tests to keep complexity on a reasonable level. 200 Byte corresponds to the often assumed 50 Byte payload and 150 Byte security overhead for CAMs [16]. The resulting packet has a total size of $30(PHY/MAC) + 20(IP) + 8(UDP) + 200(Payload) = 258\text{Byte}$. The main configuration therefore uses only $(258\text{Byte} * 100\text{Hz}) / 3\text{Mbps} \equiv 6.6\%$ of the available resources.

The test setup leads to $4*2+1=9$ transmitter positions and to $4(\text{SideStreet})*2(\text{TxDistance})*2(\text{DriveDir.})+1(\text{Center})*2(\text{Streets})*2(\text{DriveDir.}) = 20$ runs for a typical tested intersection. In total, we tested 71 transmitter positions and performed roughly 170 test runs.

4 Measurement Results

A “run” consists of one traverse of an intersection on one street. It is identifiable by the association of TxLane-ID, Tx-Distance and a “From” and “To” Lane-ID indicating the driving direction.

We decided to plot all results—despite otherwise mentioned—in a geographic, rather than directional orientation. Therefore, the smaller Lane-ID of From and To-Lane is always plotted on the left. This becomes especially important, when runs are combined to provide averaged information: the results are averaged geographically. Therefore, two runs, 3→1 and 1→3 are combined to 1-3, showing the average reception for each side street.

All runs have been visualized on a map and as xy-plot. The map visualization is a Keyhole Markup Language (KML) file containing aggregated per second results of the run. Reception rate, as well as average received power is visualized as false-color coded dot, respective donut (Left part in Figure 2). One main

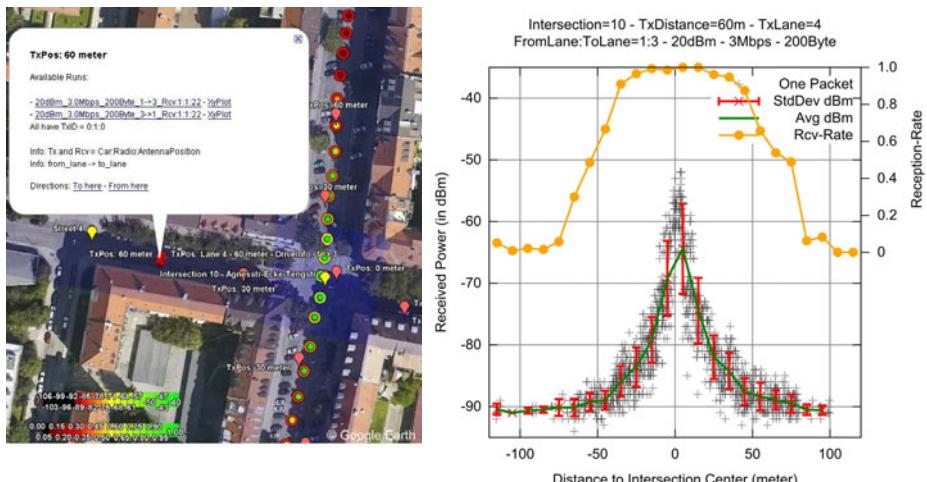


Fig. 2. KML based visualization of single runs. For each run, a run kml and a corresponding xy-plot are accessible via link from a single start kml.

KML file provides information on the transmitter positions, street IDs and area of interest. The xy-plot visualizes the power value of each received message, complemented by average and standard deviation and reception rate values. These are computed for distance-bins of 10 meter width. The complete per run evaluation is accessible under:

http://dsn.tm.kit.edu/download/bmw/nlos_test/

The example shows a run from intersection 10 with 60 meter tx-distance. Clipping at a threshold of \approx -92dBm becomes visible. This is 3dBm above the noise floor (-95dBm, retrieved from radio driver). Available data sheets (e.g. [7]) for Atheros 802.11 chipsets as used in the Linkbird also show a minimum reception threshold of \approx -92dBm with same signal modulation (802.11p@3Mbps \equiv 802.11a@6Mbps). The small standard deviation at higher ranges is a consequence of low reception rates: if a message could be received, its power is closely above the reception threshold. One should be aware that the average power curve does not correspond to path loss as soon as there is no reception rate of 100%. In general, the curve shows the average power of received messages.

4.1 Free Space Test and Antenna Position

We performed a free space test to verify a good antenna position and check proper operation of all components. Free space condition minimizes the influence of other effects like reflections. It also allows to determine if the test system performance is on par with theoretical predictions.

The three tested antenna positions were: absolute center of the car roof (22), front-right (13) and back-right (33). They are visualized in earlier Figure 1. We performed two drive-by's with each positioning, one in each direction. Both runs were averaged in the heading domain, averaging out the influence of left/right and environment differences on the two street stretches. The resulting approach/depart performance of the antennas is shown in Figure 3. While approaching the

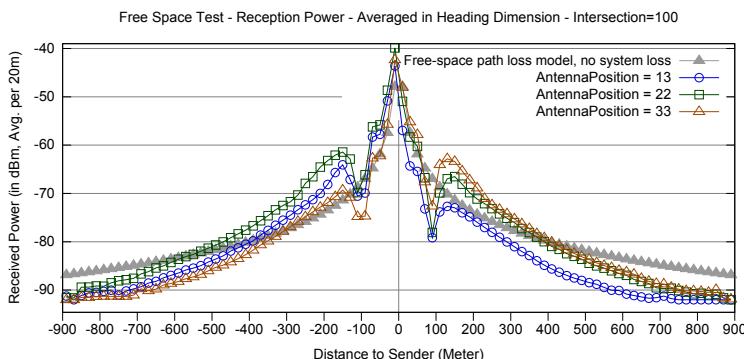


Fig. 3. Free Space Test with 3 Antenna Positions: 22=CenterCenter, 13=FrontRight, 33=BackRight. Driving Direction from -900 to +900. Averaged from Forth+Back Run.

transmitter, the center position performs best. At 300 meter distance, it has 3 dB advantage over the front antenna and 5 dB over the back antenna. While departing, the front position is 5 dB worse than center, while the rear is 1 dB better than the center position. Therefore, the center position provides the best overall performance. This is why we selected it in all NLOS tests.

A comparison to the well known Free-Space Path Loss (FSPL) model shows that our setup produces results close to theory. Adding approximately 1dB system loss (-2.4dB cable + $\approx 1\text{-}2\text{dB}$ antenna) at receiver and transmitter to the model, values would be very close to ours at high distances. Surprisingly, the real reception power at distances below 400 meter turns out even better than theoretically predicted. Of course, despite the prominent drop due to destructive two-ray ground reflection at 90 meter. Likely, ground reflection leads to the positive gain at distances around the dip.

4.2 Reception Quality under NLOS - Testing Results

We aggregated the results from single runs on two different levels to be able to compare the performance on a street, as well as on an intersection level. Averaging was done over the average-values per distance-bin. This way, fair weighting between runs is ensured, despite differing amounts of input samples per bin.

On street level, for each transmitter position, the back and forth runs were averaged. 30 and 60 meter tx-distance yield 4 results each. Plus two for the central transmitter position. On intersection level, the distance-bin values were averaged from all runs of a certain tx-distance. Therefore, three data-sets per intersection remain, one for 0, 30 and 60 meter tx-distance, respectively.

Street Level. Figure 4 shows the street level results for 60 meter at intersection 10. This urban intersection has fairly average performance as it can be seen later in Section 4.2. Considering the high transmitter distance of 60 meter, the

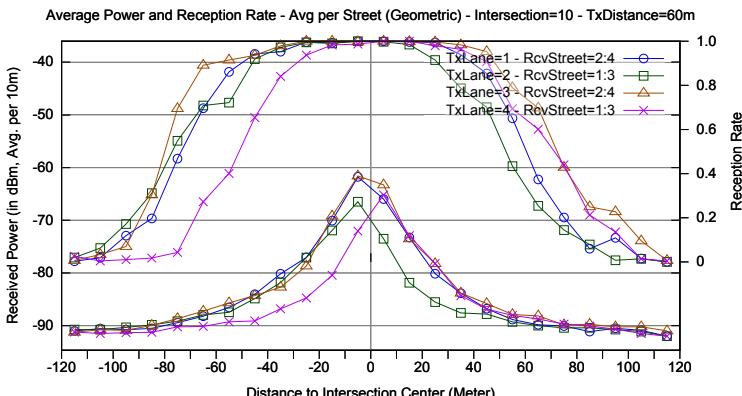


Fig. 4. Intersection 10 - Streets in an Intersection - Resulting Reception Rate and Power Over Distance for all 4 Transmitter Positions with 60m Distance to Center

reception rate over distance turns out fairly well. Reception rate drops under 50% not below 50 meter into the side street, mostly above.

The plot also shows that performance is pretty similar despite the transmitter positions being in 4 different side streets. 6 out of 8 transmitter to cross street leg results are very similar. Only when the transmitter was in street leg 4 and 2, performance in one cross street leg was worse. The other intersections show similar results. As we selected most intersections with a regular layout (same side street width), this confirms the assumption that side street width has most influence on results. If trees, bushes or building facades had major influence, the results would likely turn out more different despite the regular layout.

Intersection Level. The per transmission distance averaged values for the same intersection are plotted in Figure 5. The average power plot shows an expectable drop between the 30 and 60 meter tx-distance results due to 30 meter more distance in the transmitter street and slightly narrower view into the sight street. The gap between the 0 and 30 meter lines is much wider. As long as no missing data due to non-receptions introduces a slowing of the slope, the reduction is in the 18 to 20 dB range. The penalty due to missing LOS is of course only a part of this, as the 30 meter from transmitter to center also leads to a certain loss. For 30 meter tx-distance, reception rate stays really high

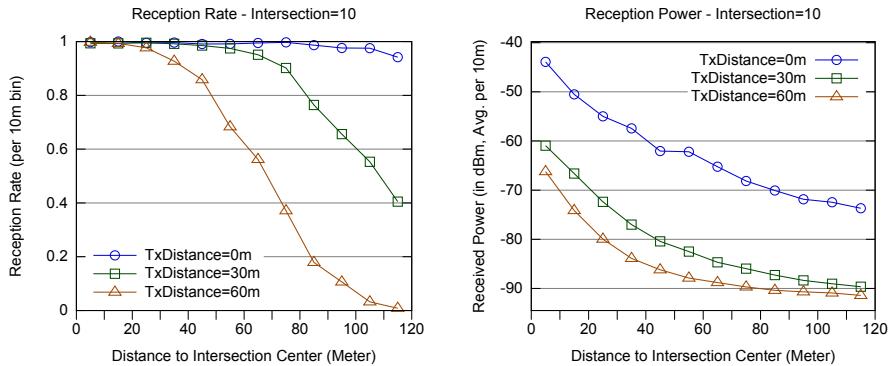


Fig. 5. Intersection 10 - Data Averaged for Same Transmit Distance to Center

(>90%) for a side street distance of up to 80 meter. For 60 meter tx-distance, the curve is shifted approximately 40 meter to the left. The center tx-position shows that with the used 20 dBm transmit power, a reception percentage of over 90% is feasible at 120 meter, despite certainly present interference from reflections.

Intersection Comparison. After showing the performance in one example intersection, this subsection compares the performance in different intersections with each other. Figure 6 shows the per intersection averaged results for all 8 intersections. The 30 meter tx-distance results are shown on the left, 60 meter on the right, reception rate is on the top, average received power on the bottom.

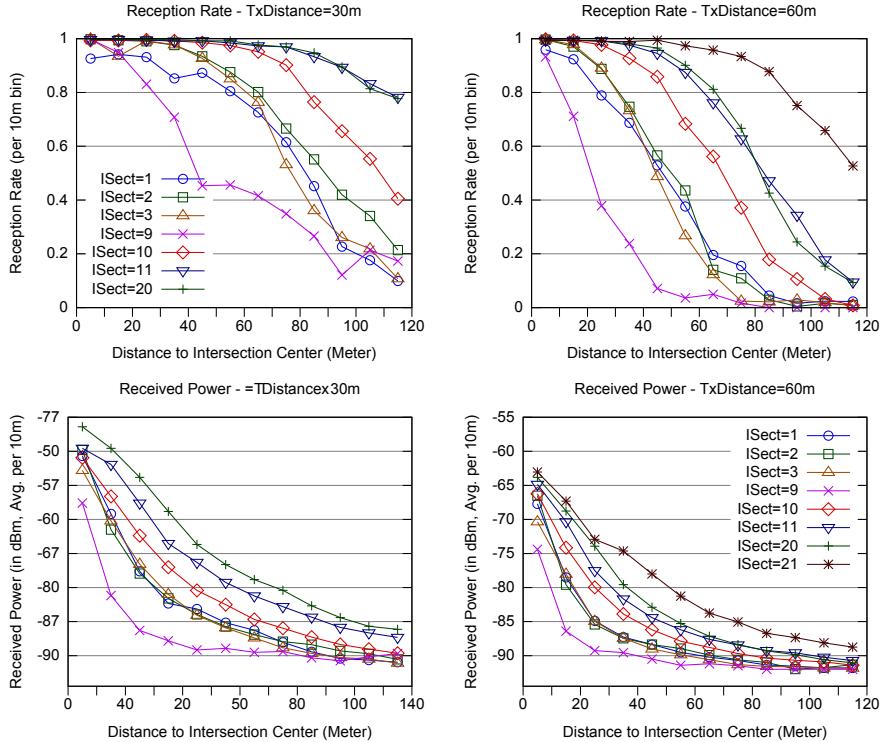


Fig. 6. Intersection Comparison - Reception Rate (Top) and Power (Bottom) - Per Intersection: Average for All Runs With Same Tx-Distance - Left: 30m, Right: 60m

First of all, note the comparatively small difference between the three suburban intersections, all having same building positions. Intersection 10, having the same dimensions as the three suburban ones, shows notably better reception rates and power values ($\approx +4\text{dB}$). This is most likely a result of less obstructions and cleaner building facades in the urban setting. Intersection 20, being also urban, but having 30 meter side-street width instead of 24+21m as for intersection 10, has another $\approx 6\text{dB}$ better values. Intersection 11 has values fitting between the ones of intersection 10 and 20. This matches its side-street width distances being in between 22 and 30 meter. For 60 meter transmitter distance to center, there is also a data row for the really wide intersection 21. It performs—as expected—even better.

These observations seem to proof the impact of building position on NLOS reception as presumed in cellular NLOS models with below-rooftop base stations. This is not surprising though: with increasing distance between the buildings of a street, the view angle into a side street is wider, resulting in more LOS and a more open angle for reflected beams, leading to a further traverse into a side street. The plots also show that suburban conditions are worse compared to urban ones.

A special case is intersection 9. The missing reflection areas due to non existence of buildings on two corners of the crossing provide a significant impact on signal power and reception rate (Note that we excluded the street legs with LOS from the averaging for comparability). But, even in this worst case scenario, there is still connectivity potential left: when two cars approach an intersection in the same moment and are 30 meter away from center, the reception rate is already at $\approx 50\%$.

Overall, reception rates turn out pretty well. Considering two vehicles approaching synchronously, reception should start at distances of ≈ 70 meter and more. 50% reception rate are reached at 50 meter to center or more. At this distance, a car needs 3.5s to reach an intersection center while driving at 50km/h.

5 Conclusion

For this paper, we performed an extensive field test with a profound scenario selection to investigate DSRC NLOS reception quality with off-the shelf radio equipment. The results show that DSRC NLOS reception is in general well feasible despite the high frequency of 5.9 GHz. If two vehicles synchronously approach an intersection from crossing streets, 50% reception rate is reached at 50 meter to intersection center. At 30 meter, rates are as high as 80%. These numbers hold for tight suburban intersections with worse than average conditions.

The collision avoidance use case needs information as early as 3 seconds to a potential collision. At 50 km/h, a vehicle travels 42 meter in this time. At this point, reception rates will be already at levels above 50%. Additionally, the investigation also shows that urban intersections provide even better NLOS reception than suburban intersections with front yards. And, if the inter-building distance rises, reception rates improve even more.

In general, we were able to verify that intersections with same inter-building distances perform very similar. Combined with the above described observation of better performance with higher inter-building distance, it seems to prove a correlation obtained from cellular below-rooftop base station models: the inter-building distance in a street has a major influence on the path-loss under NLOS conditions. Future work includes the comparison of obtained results to existing channel models.

In terms of our tested intersections, a worst case condition for DSRC NLOS reception was a side street with no buildings on the opposite side of the intersection. Missing reflection surfaces lead to considerable less signal power into a crossing street compared to intersections with all corners occupied.

As good as the numbers are, it has to be considered that a pretty high transmission power has been used. It leads to long transmission ranges in the street canyon the transmitter is in, possibly resulting in significant radio interference. Nevertheless, the most important situation where NLOS reception is needed is probably a suburban intersection with no traffic but two cars on collision course. In such a situation, competition levels will be unproblematic. In general, it is up

to future research to investigate the influence of high transmission power (due to NLOS reception demand). To do so, we intend to develop a validated DSRC NLOS path-loss model from our collected data that can be used in future DSRC simulations.

References

1. Nitz, G., Klanner, F.: Evaluation Of Advanced Driver Assistance Systems For Supportive Brake Application. In: Proceedings of FISITA 2008 World Automotive Congress, Munich (2008)
2. Klanner, F.: Entwicklung eines kommunikationsbasierten Querverkehrsassistenten im Fahrzeug. Fortschritt-Berichte VDI Reihe, Düsseldorf (2008)
3. IEEE Standards Association: Status of Project IEEE 802.11 Task Group p, http://grouper.ieee.org/groups/802/11/Reports/tgp_update.htm
4. Mangel, T., Schweizer, F., Kosch, T., Hartenstein, H.: Vehicular safety communication at intersections: Buildings, Non-Line-Of-Sight and representative scenarios. In: 2011 Eighth International Conference on Wireless On-Demand Network Systems and Services (WONS 2011), Bardonecchia, Italy, pp. 35–41 (January 2011)
5. INTERSAFE-2 - An FP7 Project Funded By The EUROPEAN COMMISSION: User Needs and Operational Requirements for a Cooperative Intersection Safety System - Deliverable D3.1 (2009)
6. TRL Limited on behalf of the European Commission: Automated Emergency Brake Systems: Technical requirements, costs and benefits (2008)
7. Karedal, J., Tufvesson, F., Abbas, T., Klemp, O., Paier, A., Bernado, L., Molisch, A.F.: Radio Channel Measurements at Street Intersections for Vehicle-to-Vehicle Safety Applications. In: 2010 IEEE 71st Vehicular Technology Conference (May 2010)
8. Miucic, R., Schaffnit, T.: Communication in Future Vehicle Cooperative Safety Systems: 5.9 GHz DSRC Non-Line-of-Sight Field Testing. SAE International Journal of Passenger Cars-Electronic and Electrical Systems 2(1), 56 (2009)
9. Giordano, E., Frank, R., Pau, G., Gerla, M.: CORNER: a Realistic Urban Propagation Model for VANET. In: Wireless On-demand Network Systems and Services, pp. 1–4 (2010)
10. NS-2: The Network Simulator NS-2, <http://www.isi.edu/nsnam/ns/>
11. El-Sallabi, H.M.: Fast path loss prediction by using virtual source technique for urban microcells. In: IEEE VTC 2000-Spring Tokyo, pp. 2183–2187 (2000)
12. Sun, Q., Tan, S., Teh, K.: Analytical formulae for path loss prediction in urban street grid microcellular environments. IEEE Transactions on Vehicular Technology 54, 1251–1258 (2005)
13. Denegri, L., Bixio, L., Lavagetto, F., Isra, A., Braccini, C.: An Analytical Model of Microcellular Propagation in Urban Canyons. In: 2007 IEEE 65th Vehicular Technology Conference - VTC 2007-Spring, pp. 402–406 (April 2007)
14. iTETRIS - An Integrated Wireless and Traffic Platform for Real-Time Road Traffic Management Solutions: D4.1 V2V Wireless Communications Modeling (2009)
15. NEC: NEC LinkBird-MX V3 Version 3 - Datasheet
16. Sichere Intelligente Mobilität Testfeld Deutschland (simTD): Deliverable D21.4 - Spezifikation der Kommunikationsprotokolle (2009)
17. dd-wrt.com: Mini-PCI WiFi Module : DCMA-82, <http://www.dd-wrt.com/shop/catalog/pdf/dcma82.pdf>

Interoperability Testing Suite for C2X Communication Components

Fabian de Ponte Müller¹, Juan María Reveriego Sierra², Bernhard Kloiber¹,
Matthias Röckl¹, and Thomas Strang¹

¹ Institute of Communications and Navigation,
German Aerospace Center, Munich,
Germany

² University of Málaga,
Málaga, Spain

Abstract. This paper presents a collection procedures to perform interoperability tests of C2X communication equipment. Following a cross layer approach, interoperability test purposes on radiocommunication equipment, networking layer algorithms, C2X applications and use cases are analyzed. Along with this analysis the main conclusions gathered from the execution of the proposed test cases on a concrete C2X prototype system will be presented.

Keywords: interoperability testing, C2X communication, geographical routing, physical layer, medium access, use case, test case.

1 Introduction

The European commission has distributed in the last years research funding for European field operational tests for Intelligent Transportation Systems (ITS) in order to accelerate the development of C2X prototypes across Europe. Most likely these prototypes include hardware and software components developed by different manufacturers. However, experiences have shown that during the integration phase on the field large effort has to be invested in solving interoperability issues between devices, specially in case of devices brought together by different partners. A solution to overcome this issue is conducting proper defined test cases to verify the correct interoperability between devices.

Many ITS applications rely on direct communication between vehicles (V2V) or vehicles and infrastructure (V2I). While numerous communication methods for ITS systems are defined [2], including 3G, Bluetooth, IrDA, 802.11a/b/g or 802.11p, it is the latter, which is likely to be in charge of direct vehicle to vehicle communication. 802.11p is rather a young standard with practically no final products available on the market. This work will concentrate on prototypes following IEEE 802.11p or its European equivalent ITS-G5.

Fig. 1 shows a usual representation of the protocol stack of an ITS system. Interoperability issues do not appear only on the lowest communication layers.

Misconfiguration or a faulty implementation of a standard can lead to interoperability problems on networking, facility or application layer. For this reason a layered approach has been followed throughout this work. First the lowest communication layers are tested, then moving upwards through the communication stack, the networking layer and the C2X applications are regarded. This way, interoperability failures from lower layers can be discarded when observing high level failures. In order to compose this suite of test cases, on every layer the corresponding standard or specification has been taken and analyzed for interoperability issues.

The test cases analyzed in the scope of this work have been carried out on an ITS implementation developed within the European project *PreDrive C2X* and which follows the common European architecture defined by COMeSafety [2]. The employed communication devices are prototypes from manufacturers such as NEC¹, DENSO² and Delphi³. These communication devices contain the lowest communication layers, namely the physical layer, the data link layer and the networking layer. The networking layer within the project is based on the GeoNet specification [5] for geographical addressing and routing, likely to be standardized for ITS at European level by ETSI. Two implementations of this standard, by NEC and Hitachi⁴, have been tested. The facility layer, which provides services for the running applications, and the applications itself were developed by different partners and are running on a open source OSGI implementation on an automotive PC. The applications that have been tested for interoperability are road works warning, car breakdown warning, green light optimal speed advisory, in-vehicle signage and electronic emergency brakelight.

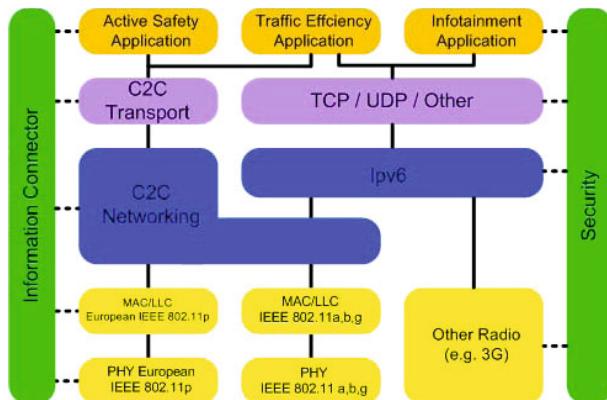


Fig. 1. Reference protocol stack for the common European architecture [2]

¹ NEC Europe Ltd.

² DENSO International America, Inc.

³ Delphi Delco Electronics Europe GmbH.

⁴ Hitachi Europe SAS.

2 Test Cases

The performance verification of a single equipment against its specification is known as *conformance testing*. Conformance testing is accomplished using a reference system at standardized interfaces, which are normally not accessible to the user or developer. In this sense we expect our systems under test (SUT) to have been proven for conformance by the system manufacturer in advance. *Interoperability testing*, on the other hand, proves the end-to-end functionality of the whole set of equipments in its conglomerate [3].

In interoperability testing there is always one item which is the subject of the test, the equipment under test (EUT) and a qualified equipment (QE) that has already been proven to interoperate. Due to the lack of a QE which all other EUT may be tested against, the test purposes presented here follow a test configuration where two or more EUT are tested against each other. When a EUT proves to pass a test it can be stated to be a QE to perform tests on further devices. Fig. 2

In the next sections different test cases on different communication layers will be identified. The test set-up and the test procedure will be commented and a set of verification parameters will be specified. Finally, the test verdicts for the given prototypes will be presented for each layer.

2.1 Physical Layer Testing

Physical layer testing is the first step in verifying the correct interoperability of C2X radiocommunication devices. Once direct communication between two devices is proven, further tests on networking or application layer can be performed without the concern of a failure at the lowest layers. The physical layer used in European systems for car-to-car and car-to-infrastructure communication is defined in ETSI ITS-G5 standard (ES 202-663 [9]), which is a slightly modified version of IEEE's 802.11p amendment for ITS systems. As the given devices under test work in the frequency band from 5855MHz to 5905MHz the requirements for G5A and G5B will be examined next.

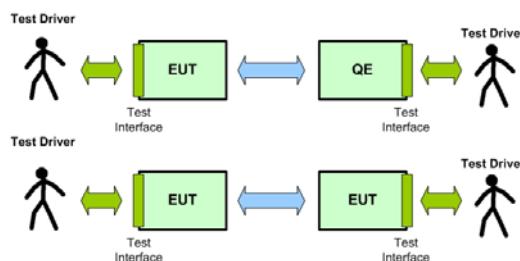


Fig. 2. Illustration of main interoperability testing concepts

- Frequency Allocation: ITS devices should be able to tune to any of the five 10MHz channels in the G5A and G5B band, namely channels G5CC (control channel) and G5SC1 to G5SC4 (service channel), which correspond to IEEE channels 172-180.
- Channel Bandwidth: According to the standard the supported channel bandwidth should be 10MHz .
- Dual receiver: The standard requires ITS devices to listen to the control channel while not transmitting. Practically, this results in a dual receiver concept, as radio devices are not able to switch without some delay.
- Data rate: The transmitted data rate is changed by using different modulation scheme/coding rate pairs. Default values for the data rate are 6Mbps on channels G5CC and G5SC1 and 12Mbps on the rest.

Test parameters related to RF output power, transmit power control, power spectral density or unwanted emissions fall out of the scope of interoperability testing and are part of previous conformance tests [8]. All the aforementioned communication units feature WLAN miniPCI cards based on Atheros chipsets, which prove to be flexible enough to be tuned with the help of a suited driver (madwifi, ath5k, etc.) to meet the European standard. Practically all ITS prototypes, however, can modify their transmission parameters prior utilization by means of some graphical configuration tool or a mere command line application, making the parameter setting transparent to the user/tester. For instance, the communication units used during this work have configuration files and configuration scripts where the radio interface can be set appropriately. According to this, different test cases have been identified. The first three test cases should be performed indoors, placing the devices at an appropriate distance from each other (i.e. far field). This way, the tests are independent on the output power, the receiver sensitivity or the channel characteristics.

- Transmission channel: Sequentially, a pair of communication devices should be tuned to a common channel and the ability to communicate should be verified. 0% PER (Packet Error Rate) is expected to PASS the test.
- Bit rate: This test case proves the ability of devices to communicate at different bit rates. The communication bit rate is given by the transmitter, which defines it in the header of the PHY frame. The header itself is coded always at 3Mbps . Although the default is 6Mbps (12Mbps on G5SC2) other possible bitrates are 3, 4.5, 9, 18, 24 and 27Mbps . Communication using these values should be verified requiring in each case a 0% PER to PASS the test.
- Dual receiver: ETSI ITS-G5 standard requires *de facto* that every ITS device features two radio interfaces, one of them tuned on channel G5CC. The test requires the setup of three devices, one radio transmitting on the G5CC, a

second radio transmitting on one of the G5SC and a third radio as EUT listening on both channels. Under this configuration simultaneous reception on a service and the control channel should be verified, requiring 0% PER to PASS the test.

- Transmission range: This test is important to be able to determine the achievable communication range with given radio devices, RF cables and antennas. Obviously the possible communication range depends on several environmental factors, which moreover are difficult to quantify. Obstacles causing shadowing, diffraction or attenuation, surfaces that cause multipath propagation, line of sight (LOS)/non line of sight (NLOS) propagation conditions, transmitter/receiver movement or in-band interferers and noise are examples of phenomena that have a negative impact on the transmission range. Unlike the tests mentioned so far, this one is carried out outdoors using two devices in different vehicles and mounting the antennas following the manufacturer's advisories (cabling, ground plane, etc.). The test should be carried out under controlled and reproducible conditions with regard to LOS, reflections, etc.. The devices are progressively separated from each other until packet loss increases over 0%. Having a GNSS receiver attached to the communication device, the distance between transmitter and receiver can be evaluated. Although the minimum transmission range is not standardized by ETSI, the experiences gathered during test trials in the last years yield a minimum range of 500m under LOS conditions and maximum transmit power. According to C2C-Communication Consortium [1], safety applications require a direct communication range of at least 300m.

In order to perform these tests, two communication devices from the same or different vendors are needed. On one of the communication devices an application for generating traffic is needed, while the other device should have an application to receive traffic, log it, calculate the packet error rate (PER) and verify its completeness. Test traffic contains an identifier, a timestamp, a sequence number and, if available, position information. As the test setup includes only one sending device, the isolation of the MAC layer algorithm is assured, i.e. the MAC layer will always sense an idle carrier. Table 1 summarizes the test results for the three radiocommunication devices under test.

One of the ITS devices did not have two radio interfaces installed, thus, being unable to receive on the control channel and on a service channel at the same time. The transmission range test revealed a poor communication range

Table 1. Test results for physical layer interoperability tests

Test Case	Radio A	Radio B	Radio C
Transmission channel	PASS	PASS	PASS
Bit rate	PASS	PASS	PASS
Dual receiver	PASS	PASS	FAIL
Transmission range	FAIL	FAIL	FAIL

of 100m. This test was performed outdoors under good line of sight conditions. The 5.9GHz antennas mounted on the roof of the vehicle were connected to the communication devices by 4m long RF cables. Further tests yielded that the quality of the RF cables was degraded, resulting in attenuations beyond 8dB.

2.2 MAC Layer Testing

The data link layer is in charge of direct communication between nodes by means of addressing and control mechanisms. The lower sublayer, the medium access control (MAC) layer, is in charge of assuring a coordinated access of nodes to the medium to avoid simultaneous transmission of packets that would lead to a collision. The ITS-G5 standard, as in 802.11p, uses the CSMA/CA algorithm to avoid collisions on the channel. The data link layer adds a 32Byte header to the given SDU (service data unit) which contains the source and destination MAC addresses, a BSSID address identifying the access point as well as further fields for protocol identification, quality of service (QoS), etc. The address fields as well as the mechanism to provide QoS differ from the IEEE 802.11 standard. This leads to the following set of interoperability test cases on MAC layer:

- BSSID: Due to the absence of an access point, the stations should operate outside the context of a BSS (Basic Service Set) and the BSSID address inside the MAC header should have all bits set to 1. Configuring the radio interface to work in *monitor mode* makes it possible to visualize the complete 32Byte MAC header and check that this value is correctly set.
- Destination MAC address: According to the standard, ITS stations should support unicast and broadcast communication. The distinction is done using the destination address field in the MAC header, which in broadcast mode should have all bits set to one.
- Source MAC address: ITS stations should continuously monitor whether a neighboring station is using the same MAC address and in this case generate a new address to resolve the conflict.
- QoS: As stations work outside the context of a BSS, QoS for ITS is provided on the basis of EDCA (Enhanced Distribution Channel Access) mechanism. Depending to which of four Access Categories (AC) the traffic belongs to, the interframe space length varies, increasing, this way, the priority to access the channel. To verify the correct behavior two devices send continuously (large packets with smallest repetition time) in different access categories while a third device acts as a receiver. At the receiver, packets with higher priority are expected to exhibit a smaller PER than the ones with lower priority to PASS the test.

The results of executing these test cases on the given C2X prototypes under test is summarized in table 2.

Table 2. Test results for MAC layer interoperability tests

Test Case	Radio A	Radio B	radio C
BSSID	PASS	PASS	PASS
Destination MAC address	PASS	PASS	PASS
Source MAC address	PASS	PASS	PASS
QoS	PASS	N/A	N/A

2.3 Network Layer Testing

In fig. 11 the networking layer can be identified above IEEE802.11p layer. Ad-hoc networks formed by C2X communication systems need a special routing algorithm, which is based on geographic addressing and should be able to forward data towards nodes within a destination area. GeoNet is a set of networking protocols developed within a European project [5] and that is likely of being the future networking layer for C2X communication. Four different geographical routing schemes are defined by GeoNet: geographic unicast, geographic anycast, geographic broadcast and topological broadcast. This work puts its focus on geographical and topological broadcast, as these schemes are the most relevant in current C2X communication.

Geographical broadcast (GeoBroadcast) delivers data over multiple hops until reaching all the nodes inside a given destination area. The forwarding algorithm behind the GeoBroadcast scheme is Greedy Perimeter Stateless Routing (GPSR) [4]. GPSR takes a decision using information about immediate neighbors and chooses as next forwarder the closest node to the destination area. This way geographical broadcast is suited for transmitting, for example, decentralized environmental notification (DEN) [12] messages that inform about events in a certain geographically delimited area. Figure 12 shows the GeoBroadcast process schematically.

On the other hand, topological broadcast (TopoBroadcast) forwards data to all nodes located within a specific number of hops. This scheme is suited for transmitting, for example, cooperative awareness messages (CAM)[11], a beacon with state information sent periodically by each C2X participant.

Two implementations by Hitachi and NEC of this algorithm have been used to test their interoperability. The following test cases have been identified:

- Tx and Rx of networking packets in a point-to-point scenario: The aim is to verify the ability of communication units to exchange networking packets directly. After verifying that devices can both send and receive network packets, it is possible to test further features in more complex scenarios. For this, two communication devices are placed in range of each other. One of the devices generates first a topological broadcast message with hop limit of one and then geographical broadcast message with a circular destination area that includes both devices. Both messages should be received at the receiving counterpart.

- Store and forward capability: The store and forward buffer is a component of the networking layer, which is used to store networking packets when no neighboring node is in range. When a new neighbor appears, the stored packets are forwarded to it. This test case checks if this functionality is performed correctly at each of the nodes using two communication devices. The sender generates a GeoBroadcast packet, while the receiver device has its networking layer switched off. When the networking layer of the second unit is switched on, it is possible to check if the sender device forwarded the stored packet correctly.
- GeoBroadcast packet forwarding: This test case checks the capability of a node to forward packets. Therefore, three communication units are needed to perform this test. One of them is the sender and the test will check if only the node closest to the destination area receives a forwarded packet.
- TopoBroadcast packet forwarding: Testing TopoBroadcast forwarding is performed with three devices, one in range of the next one, verifying that the central node decreases the hop counter when forwarding a packet and that the last receiver does not forward the packet as the counter reaches zero.
- Flooding inside the destination area 1: The next step is to check the broadcast process when forwarding packets inside the destination area. According to [5] nodes inside the destination area should forward a packet just once. Three communication units are needed to conduct this test case. When the three units are inside the destination area, it has to be checked that they just forward the packets once whereas when locating one unit outside the destination area, it has to be verified how the unit receives but does not forward the packets.
- Flooding inside the destination area 2: Further on, a node located outside the destination area should drop packets which reaches him from inside the area.

Table 3 summarizes the results of the four test cases performed on both available implementations A and B.

Table 3. Test results for networking layer interoperability tests

Test Case	Implementation A	Implementation B
Direct Communication	PASS	PASS
Store & Forward	PASS	FAIL
GB forwarding	PASS	PASS
TB forwarding	PASS	PASS
Flooding 1	PASS	PASS
Flooding 2	PASS	PASS

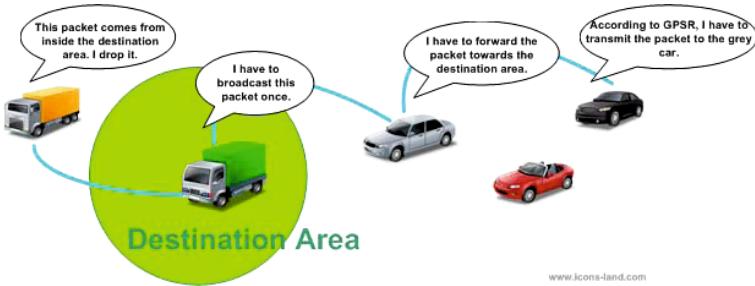


Fig. 3. GeoBroadcast Process

Direct node to node communication by means of topological broadcast messages is performed correctly with both implementations. When using GeoBroadcast communication, only one of the implementations correctly forwarded a previously stored message to a new neighbor. In principle both networking layers implement correctly the GPSR algorithm as described in [5]. Nevertheless, the third test case revealed a slight discrepancy with the specification. The generating node always sent the packet to all surrounding neighbors, instead of selecting the nearest node to the destination area or, in the absence of neighbors, storing the message in the store&forward buffer. This fact does not impede networking interoperability, but increases the load on the channel by creating several competing routes towards the destination area. Once the GeoBroadcast packet reaches the destination area, every node correctly forwards the packet once and discards packets coming from inside the area. The proposed networking layer test cases showed, that both implementations under test were fundamentally correct, but showed small flaws, that should be revised in further versions.

2.4 Application Layer Testing

Interoperability testing on application layer is somehow abstract. The large list of C2X use cases defined by the ETSI in its Basic Set of Applications [10] and their corresponding set of functional requirements, can be seen as a basis for conformance testing, where each of the requirements is checked for its rightness using a single unit and generating artificially the appropriate input. Further on, the correctness and completeness of the exchanged CAM, DENM and GLOSA messages according to their respective specifications, as well as timing requirements for CAM messages, are considered part of previous conformance testing.

Moving towards interoperability testing, the aim is verifying the right behavior of C2X use cases during their execution on the field involving several participants. Thus, the aim is to check if vehicles receive and show corresponding warnings when they are relevant for drivers. This layer is divided in two sublayers: the use-case sublayer and the facility sublayer. The facility sublayer offers services to use-cases such as positioning or relevance checking. Usually, use-cases receive messages from other vehicles, verify the relevance of the messages and warn the driver if necessary.

The C2X implementation developed in *PreDrive C2X* and used in the scope of this work, featured the following use cases:

- Road Works Warning (RWW)
- Emergency Electronic Brake Light (EEBL)
- Car Brakedown Warning - CBW
- Green Light Optimal Speed Advisory (GLOSA)
- In-vehicle Signage (IVS)

RWW and IVS are use cases, where infrastructure components send DEN messages periodically to the surrounding vehicles, whereas EEBL and CBW are use cases where the DEN messages are generated at a vehicle. While the former involve testing the receiving end, as the sender just repeats predefined messages, the latter require to include the sender and the receiver in the test case, as some kind of condition has to be fulfilled at the sender for starting generating DEN messages. In GLOSA, on the other hand, an infrastructure component sends traffic light status messages periodically to the approaching vehicles.

Although each use case is different in its purpose, there are similarities in their operation. Next, a list of test cases for these C2X applications according to [10] is presented:

- Test 1: DENM generation. The test checks if, according to the input sensors (inertial measurement data and brake status for EEBL or speed and emergency lights status for CBW), a DEN message is generated correctly inside the vehicle.
- Test 2: Relevance checking. This test verifies if an incoming DEN message is relevant for a driver. Position, speed and heading of the vehicle, distance to the event and radius are taken into consideration to evaluate the relevance.
- Test 3: Displaying of warning to the driver. This test checks whether a relevant warning with highest priority in the warning queue was actually forwarded to the driver's Human Machine Interface (HMI).
- Test 4: Revoking of warnings. This test verifies that an expired or not relevant warning is removed from the driver's HMI.
- Test 5: Overall delay. This test measures the delay since a DEN message is generated until the actual warning is displayed on the receiving end. Depending on the use case, the delay shall be as small as 100ms.

CODAR Viewer [6], a tool created by the DLR to monitor and visualize C2X test sites, was further developed to control field operational tests and extract test data out of the vehicles. This data is sent towards a server, where the

Table 4. Test results for C2X use-case interoperability tests

Test Case	App 1	App 2	App 3	App 4	App 5
Test 1	N/A	PASS	PASS	N/A	N/A
Test 2	FAIL	PASS	PASS	PASS	PASS
Test 3	PASS	PASS	PASS	PASS	FAIL
Test 4	PASS	PASS	PASS	PASS	FAIL
Test 5	PASS	PASS	PASS	PASS	PASS

running use-case is monitored and evaluated online by a tester. The proposed test cases were performed on the given C2X applications under test. Table 4 gives exemplary results for these test cases, which helped to detect failures during the implementation phase.

During the implementation phase, one of the applications showed problems with the correct displaying of warnings and with their revoking. Another application had too stringent conditions for accepting a warning as relevant with the consequence that no HMI alert was given to the driver. These issues were later solved correctly.

3 Conclusions and Future Work

This paper has presented a variety of interoperability test cases to be performed on different communication layers on European ITS systems. Starting on the physical layer direct ITS-G5 radio-to-radio communication has been tested, being frequency channel, bit rate or bandwidth the most common accessible parameters. One layer above, test cases for the MAC layer have been explained. Both, physical and MAC layer test cases have been applied on three different C2X communication devices from different manufacturers. On networking layer, the focus has been put on GeoNet specification, as the most likely set of algorithms to be standardized by the ETSI for ITS systems. A set of test cases for geographic and topological broadcast have been presented along with the results on their execution on two prototypical implementations. On application layer, five C2X use cases developed within the European project *PreDrive C2X* have been evaluated using the CODAR Viewer tool.

Additionally, interoperability testing on two further aspects of C2X systems is considered by the authors to be essential and to require exhaustive analysis. Both, security that deals with integrity, authenticity and trustworthiness in C2X communication to avoid malicious usage of the system, and privacy protection to preserve anonymity or to prevent recording of movement, will also require extensive interoperability testing due to the involvement of different entities (signer, verifier, certificate authority, etc.). Further on, another ITS architecture implementation, namely the one developed in the European project CVIS [7], should be analyzed for interoperability in future work. This architecture follows the COMeSafety definition and is based on ISO CALM. In spite of large similarities

on application, facility and lower communication layers with the implementation analyzed in this work, CALM is strongly focused on IPv6 as networking strategy.

Acknowledgment

The work for this paper was accomplished in the frame of the European project ITS TestBeds. The ITS prototypes used during this work were developed in the European project PreDrive C2X.

References

1. Baldessari, R., Boedekker, B., Deegener, M., Festag, A., Franz, W., Christopher Kellum, C., Kosch, T., Kovacs, A., Lenardi, M., Menig, C., Peichl, T., Roeckl, M., Seeberger, D., Strassberger, M., Stratil, H., Voegel, H.-J., Weyl, B., Zhang, W.: Car-2-car communication consortium - manifesto, 08 (2007)
2. Bossom, R., Brignolo, R., Ernst, T., Evensen, K., Frötscher, A., Höfs, W., Jääskeläinen, J., Jeftic, Z., Kompfner, P., Kosch, T., Kulp, I., Kung, A., Mokadem, A.-K., Schalk, A., Uhlemann, E., Wewetzer, C.: European ITS communication architecture — Overall framework — Proof of concept implementation. Technical report, EC Information Society Technologies Programme (February 2010)
3. ETSI Technical Committee Methods for Testing and Specification. Methods for testing and specification (mts); internet protocol testing (ipt); generic approach to interoperability testing. Technical report, ETSI, ETSI Guide ETSI ES 302 571 (September 2008)
4. Karp, B., Kung, H.T.: GPSR: greedy perimeter stateless routing for wireless networks, pp. 243–254 (2000)
5. Kovacs, A.: GeoNet Strep N216269 D2.2 Final GeoNet Specification (January 2010)
6. Kranz, M., Franz, A., Röckl, M., Andreas, L., Strang, T.: Codar viewer - a situation-aware driver assistance system. In: Advances in Pervasive Computing Pervasive 2008 Demonstration, vol. 236, pp. 126–129. Austrian Computer Society (2008)
7. Solberg, A., Schmid, A., Baggen, M., Alesiani, F., Stratil, H., Bossom, R., Forkert, S., Schlingelhof, M., Burkert, A., Haverkamp, S., Mathias, P., Jeftic, Z., Evensen, K., Olsen, E., Fischer, H.-J., Fazekas, I., Gaillet, J.-F.: D.CVIS.3.4 - Final Architecture and System Specifications (June 2010)
8. ETSI Technical Committee Intelligent Transport Systems. Intelligent transport systems (its); radiocommunications equipment operating in the 5855 mhz to 5 925 mhz frequency band; harmonized en converging the essential requirements of article 3.2 of the r&tte directive. Technical report, ETSI, ETSI Guide ETSI ES 202 237 (April 2007)
9. ETSI Technical Committee Intelligent Transport Systems. European Profile Standard for the Physical and Medium Access Control Layer of Intelligent Transport Systems Operating in the 5GHz Frequency Band. Technical report, ETSI, ETSI Standard ETSI ES 202 663 (November 2009)

10. ETSI Technical Committee Intelligent Transport Systems. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 1: Functional Requirements. Technical report, ETSI, Technical Specification ETSI TS 102637-1 (September 2010)
11. ETSI Technical Committee Intelligent Transport Systems. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service. Technical report, ETSI, Technical Specification ETSI TS 102637-2 (April 2010)
12. ETSI Technical Committee Intelligent Transport Systems. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 3: Specifications of Decentralized Environmental Notification Basic Service. Technical report, ETSI, Technical Specification ETSI TS 102637-3 (September 2010)

Towards Standardization of In-Car Sensors

Zubair Nabi¹, Atif Alvi¹, and Rashid Mehmood²

¹ Computer Science Department
Lahore University of Management Sciences, Pakistan
`{zubair.nabi,atif.alvi}@lums.edu.pk`
² School of Engineering
Swansea University, UK
`r.mehmood@swansea.ac.uk`

Abstract. In this paper we propose standardization of the firmware of in-car sensors to achieve software homogeneity across vendors. Such standardization enhances the reliability of the code by adhering to pre-set practices and definitions. It also increases the throughput of the programmers and opens up hardware platforms to third party developers. We make our case by advocating the use of TinyOS. TinyOS applications, coded in nesC introduce event-driven execution and component-centric design. Furthermore, programming in nesC reduces code size and potential bugs. We use the tyre pressure monitoring system of a car as a case study to illustrate our model. In our implementation, TinySec is employed to address the security and privacy problems that plague current systems. Throughout the paper we show that the use of a common software stack automatically leads to hardware standardization and security improvement.

Keywords: VANETs, In-Car Sensors, Wireless Sensor Networks, TinyOS, TinySec, TinyDB, Security.

1 Introduction

The average car today has 100 million lines of code and the cost of software and electronics accounts for 35 to 40 percent of the overall cost of the car [1]. This huge amount of code makes up the firmware of both sensors and the Electronic Control Units (ECU) used to control them. A typical intra-vehicular network consists of hundreds of sensors connected to ECUs. This network can be a hybrid of both wired (CAN-bus [2], FlexRay [3] etc.) and wireless standards (Bluetooth, ZigBee [4] etc.). These in-car sensors measure temperature, tyre pressure, fuel injection, engine knocking, air supply etc. to provide stability, safety and comfort to the occupants. To rein in the proliferation of in-car sensors, manufacturers are moving towards wireless connections. Wireless connectivity can save space, reduce the car weight and induce flexibility into the network.

Moreover, in-car sensors which are wirelessly connected, can be likened with Wireless Sensor Networks (WSN) [4]. Sensor nodes or *Motes* gather sensory information, perform basic processing and transmit data to other nodes in the

network. A typical mote consists of a CPU, radio, power source, and a variety of optional sensory elements. The CPU is typically an embedded microcontroller which performs all the processing. These microcontrollers run embedded operating systems – with TinyOS [5] as the most popular choice. TinyOS is an open source, event-based embedded operating system which supports concurrent programs with constrained memory requirements. The core OS has a signature of just 400 bytes which makes it an ideal choice for lower-power nodes.

In recent years, researchers have diverted their focus to standardization of various automotive software stacks. Of these, two deserve a mention: AUTOSAR [6] and EM2P [7]. AUTOSAR (Automotive Open System Architecture) is an ECU firmware standardization effort by manufacturers, vendors and tool developers which aims to define a common E/E (electrical/electronic) component based software architecture to facilitate scalability, integration, transferability and maintainability. AUTOSAR has safety as one of its prime goals and invites novel techniques to fulfil the stringent safety requirements of automotive applications. Likewise, EM2P (Emma Embedded Middleware Platform) is a middleware platform designed as part of the EMMA project. It enables the development of new applications for existing embedded sensors. EM2P hides the complexity of the underlying hardware and allows hardware vendors to open their interfaces to third party developers. It also links disparate sensors together by introducing a common wireless communication interface. Additionally, there has been increasing interest in the security deficiencies and mechanisms in vehicular networks. This includes exploiting vulnerabilities in wireless Tyre Pressure Monitoring sensors [8], and using the CAN-bus to infiltrate ECUs [9].

Therefore, in this paper we advocate the use of a common software and communication stack to achieve 1) Hardware homogeneity 2) Code reliability and 3) Security enrichment.

The rest of the paper is organized as follows. We give an overview of the TinyOS programming model in § 2. § 3 presents our Tyre Pressure Monitoring system case study. Related work is summarized in § 4. We conclude in § 5.

2 Background

TinyOS applications are written in nesC [10]—a programming language tailor-made for networked embedded systems. nesC enables event-driven concurrent execution and component-centric application design. TinyOS applications compiled in nesC run as single images on sensor nodes.

TinyOS applications are made up of one or more *components*. Components can either be *modules* or *configurations*. Modules (which can be Private or Public) implement one or more *interfaces*, where interfaces represent the functionality of the components. Configurations assemble components together. For example, a simple application say, `TemperatureSensor` would consist of `TemperatureSensor.nc` (interface file), `TemperatureSensor.h` (header file), and `TemperatureSensorC.nc` (public module) or `TemperatureSensorP.nc` (private module).

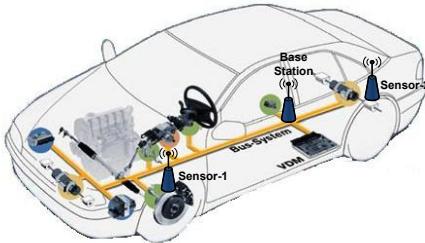


Fig. 1. Tyre pressure monitoring system which consists of 4 sensors—one for each tyre—and a base station on the CAN-bus. Original image from [13]

Also, sensor data from TinyOS can be acquired using a stream processing system: TinyDB [1]. TinyDB exposes an SQL-like interface to the user to define the data to be extracted and the rate of extraction. Extracted data is aggregated and routed to a PC using in-network processing algorithms. Data security can be enabled through TinySec [2], which is an in-software link layer security architecture with minimal impact on performance¹.

3 Case Study: Tyre Pressure Monitoring System

In this section we explain the basic model of a Tyre Pressure Monitoring System (TPMS). We then present the software implementation of a TPMS in TinyOS using nesC. We also describe the security checks that are incorporated into our implementation through TinySec.

To improve safety, fuel efficiency and the overall driving experience, TPMSs [4] based on wireless sensors will find their way into all vehicles in the United States and the European Union by 2012. A TPMS consists of sensor modules inside each tyre that relay information (temperature, pressure, etc.) in real time via an RF channel to an ECU. As shown in Figure 1, this Base Station or ECU is connected to the rest of the networks via the CAN-bus. The TPMS ECU analyzes pressure data from all tyres and raises a flag if any one of the tyres experiences any reading fluctuation.

Although, the deployment of TPMSs is on the rise, their software and security requirements are yet to be explored in detail. Rouf et al. [8] have evaluated the privacy and security mechanisms employed by commercial TPMSs. Their study has shown that TPMSs have little or no security checks in place. In fact, messages can be reverse-engineered to track vehicles which raises significant privacy concerns. Moreover, sensor data packets can be spoofed due to lack of basic authentication or validation. As a result, tyre pressure warning messages can be remotely injected to launch an attack.

¹ As shown in [2], TinySec has a negligible computational, bandwidth, size and energy overhead.

```

zubair@ubuntu: /opt/tinyos-2.1.1/apps/TPMS_Demo
File Edit View Terminal Help
zubair@ubuntu:/opt/tinyos-2.1.1/apps/TPMS Demo$ python test.py
Creating noise model for node 0
Creating noise model for node 1
Creating noise model for node 2
Creating noise model for node 3
Creating noise model for node 4
Application started at time: 1
Injecting normal pressure (30 psi) packet at time: 4941406251
Injecting normal pressure (30 psi) packet at time: 4941406254
Injecting normal pressure (30 psi) packet at time: 4941406257
Injecting normal pressure (30 psi) packet at time: 4941406260
Injecting normal pressure (30 psi) packet at time: 4941406263
Injecting low pressure (29 psi) packet at time: 4941406266
DEBUG (0): Received pressure reading of 30 psi.
DEBUG (0): Normal pressure.
DEBUG (0): Received pressure reading of 30 psi.
DEBUG (0): Normal pressure.
DEBUG (0): Received pressure reading of 30 psi.
DEBUG (0): Normal pressure.
DEBUG (0): Received pressure reading of 30 psi.
DEBUG (0): Normal pressure.
DEBUG (0): Received pressure reading of 30 psi.
DEBUG (0): Normal pressure.
DEBUG (0): Received pressure reading of 29 psi.
DEBUG (0): Turning on warning LEDs.
DEBUG (0): LEDS: Led0 on.
DEBUG (0): LEDS: Led1 on.
DEBUG (0): LEDS: Led2 on.

```

Fig. 2. Demo run of TOSSIM for the TPMS Application

3.1 Application Development

The use of TinyOS for this purpose enables us to address the challenges faced by proprietary software such as complex comprehensive data models and compatibility [15]. Heterogeneity of the various components and lack of standardization has led to a large number of proprietary tools and models. The relatively simple mode of programming can enable the manufacturer to reduce both time and cost pressure, as applications will have a shorter development cycle. Coding in nesC can easily be made to follow the best practices for automotive systems as described in the MISRA guidelines [16].

Appendix A shows the configuration file code for this application. TinyOS automatically takes care of the underlying operating system functions and also the network stack. It also enables automatic node and topology discovery.

We now describe the flow of the TPMS Application by using a simulation. TOSSIM [17]—a simulator to capture the network behaviour of TinyOS enabled sensor nodes—was used to simulate our model.

As shown in Figure 2, the application makes use of 5 motes. Mote 0 serves as the Base Station while Motes 1-4 serve as tyre pressure sensors (one for each tyre). We manually inject pressure readings to check the behaviour of the network. A pressure reading of 30 psi is considered as normal while any value below it raises a warning flag. To make the simulation as realistic as possible we insert noise into the channel by providing TOSSIM bit error rates for each link between nodes. TOSSIM uses this data to automatically generate Gaussian packet loss probability distributions for the corresponding link. From timer event 4941406251 to 4941406263 the pressure readings received by the Base Station (node 0) are normal (i.e. 30 psi). At timer event 4941406266 we intentionally

Table 1. TinySec Modes

Receiver Mode	Sender Mode
TINYSEC_RECEIVE_AUTHENTICATED	TINYSEC_AUTH_ONLY or TINYSEC_ENCRYPT_AND_AUTH
TINYSEC_RECEIVE_CRC	TINYSEC_DISABLED
TINYSEC_RECEIVE_ANY	Any

insert a low pressure reading of 29 psi. This packet when received at the Base Station raises a flag and as a result low pressure warning LEDs are toggled.

3.2 Security

TinySec security can be added to any application through its Makefile. TinySec exposes different operation modes to the application through two functions: `setTransmitMode()` and `setReceiveMode()`. Table 1 presents the compatibility of TinySec modes. Sender modes `TINYSEC_AUTH_ONLY` and `TINYSEC_ENCRYPT_AND_AUTH` work with receiver mode of `TINYSEC_RECEIVE_AUTHENTICATED`. Additionally, security can be disabled at the sender end through `TINYSEC_DISABLE` and can be received with a mode of `TINYSEC_RECEIVE_CRC`. Whereas, a receiver mode of `TINYSEC_RECEIVE_ANY` can be used with any sender mode.

Encryption and Authentication can also be implemented on top of existing software stacks but this would require substantial porting in some cases. TinySec on the other hand already exists as a module inside TinyOS and does not face any compatibility issues.

We now describe two scenarios where TinySec is effective in addressing security vulnerabilities in standard TPMSs.

Vehicle Tracking. According to [8], a packet and its identifier from a tyre sensor can be captured from a distance of 40m. Standard tyre sensor packets do not employ encryption. Hence, these packets can be used to identify and maliciously track any vehicle. This raises many privacy concerns.

TinySec uses cipher block chaining (CBC) [18] as the underlying encryption scheme with Skipjack as the default block cipher and an 8 byte Initialization Vector (IV). Skipjack is the most appropriate choice for software implementation on embedded microcontrollers with power constraints. By setting `setTransmitMode()` to `TINYSEC_ENCRYPT_AND_AUTH` we were able to encrypt all sensor packets and ensure privacy.

Remote Spoofing. Current TPMSs also do not employ any sort of authentication [8]. Therefore, anyone can inject malicious packets into the system. These packets can be used to turn on the low pressure warning light and even crash the TPMS ECU altogether.

TinySec uses cipher block chaining CBC-MAC [19] for authentication. CBC-MAC is fast and memory efficient. We enable authentication by setting `setTransmitMode()` to `TINYSEC_AUTH_ONLY` or `TINYSEC_ENCRYPT_AND_AUTH`. This allows our implementation of the TPMS to nullify the effect of remote spoofing.

3.3 Software Configuration

TPM systems need to be reconfigured during seasonal changes and also when the tyres themselves are changed. Traditionally, this reconfiguration is carried out by technicians, which requires a trip to the mechanic. To simplify this calibration, we make use of a rule layer on top of an ontology substrate. More precisely, we add a layer of SWRL [20] on top of the ECU configuration to enable users to calibrate on-board ECUs and sensors through the console. For instance, in case of TPMS the user can change the summer code to winter through the following SWRL rule:

$$\text{TPMSECU}(\text{s}) \wedge \text{temp}(\text{t}) \wedge \text{swrlb : lessThan}(\text{t}, 5) \rightarrow \text{hasCode}(\text{s}, \text{Winter}) \quad (1)$$

The above rule is in SWRL format [20]. The antecedent is composed of conjunctions of certain conditions whose fulfilment results in the consequent. In this case, if there is a TPM ECU (represented by the variable *s*) and some sensor readings at hand that measure existing ambient temperature (*t*), and if the current temperature is less than 5°C, then the TPM ECU has winter code loaded into it (that has the correct winter tyre pressures). The rule can easily be entered through a touch-screen GUI on the car console.

4 Related Work

In this section we summarize relevant related work and compare our work with it.

Rouf et al. [8] present an analysis of the security mechanism of TPMSs and suggest improvements to address the problems that they highlight. Their work has already been discussed in Section 3. Reference [4] makes use of ZigBee to implement a wireless sensor network in an automobile. Our work is different as we focus on the software stack and the security mechanism of in-car wireless sensor networks instead of the physical layer of communication. Reference [21] presents a study of the usage of various security solutions for intra-car wireless sensor networks. We carry their work one step forward by incorporating their suggestions into a real application. Kosher et al. [9] have shown that it is possible to infiltrate ECUs and disable a whole range of on-board safety-critical systems. This infiltration can be leveraged to disable brakes, stop the engine etc. Also, an attack can embed self-deleting malicious code into any ECU. The focus of their work is the security of the CAN-bus while the focal point of our work is the security of on-board wireless sensors. Our work is complementary to the AUTOSAR [6] and the EM2P [7] initiatives but while the former focuses on ECU configuration and the latter is a work in progress, we advocate the use of existing technology and practices for in-car sensor standardization.

5 Conclusion and Future Work

We have shown that the use of TinyOS leads to improved software reliability. It also simplifies and augments the security mechanisms. Using the tyre pressure

monitoring system as a case study, we demonstrated that application development in TinyOS is far superior to current models in terms of scalability, and reuse. Moreover, we have shown that TinySec can be used to fulfil the security needs of any in-car sensor system.

Our future work includes deploying our model on an actual vehicle and exploring its performance under various scenarios. Currently, we are in the process of designing and implementing a firewall for both intra and inter-vehicular communication.

References

1. Charette, R.N.: This car runs on code. *IEEE Spectrum* (2009)
2. Tindell, K.W., Hansson, H., Wellings, A.J.: Analysing real-time communications: controller area network (can), 259–263 (1994)
3. Berwanger, J., Ebner, C., Schedl, A., Belschner, R., Fluhrer, S., Lohrmann, P., Fuchs, E., Millinger, D., Sprachmann, M., Bogenberger, F., Hay, G., Krger, A., Rausch, M., Budde, W.O., Fuhrmann, P., Mores, R.: Flexray - the communication system for advanced automotive control systems. In: *SAE World Congress* (2001)
4. Tsai, H.M., Tonguz, O., Saraydar, C., Talty, T., Ames, M., Macdonald, A.: Zigbee-based intra-car wireless sensor networks: a case study. *IEEE Wireless Communications* 14(6), 67–77 (2007)
5. Levis, P., Madden, S., Polastre, J., Szewczyk, R., Whitehouse, K., Woo, A., Gay, D., Hill, J., Welsh, M., Brewer, E., Culler, D.: Tinyos: An operating system for sensor networks. In: Weber, W., Rabaey, J.M., Aarts, E. (eds.) *Ambient Intelligence*, pp. 115–148. Springer, Heidelberg (2005)
6. (AUTOSAR (Automotive Open System Architecture), <http://www.autosar.org>
7. (EMMA Project Embedded Middleware in Mobility Applications), <http://www.emmaproject.eu>
8. Rouf, I., Miller, R., Mustafa, H., Taylor, T., Oh, S., Xu, W., Gruteser, M., Trappe, W., Seskar, I.: Security and privacy vulnerabilities of in-car wireless networks: A tire pressure monitoring system case study. In: *Proceedings of the 19th USENIX Security Symposium* (2010)
9. Koscher, K., Czeskis, A., Roesner, F., Patel, S., Kohno, T., Checkoway, S., McCoy, D., Kantor, B., Anderson, D., Shacham, H., Savage, S.: Experimental security analysis of a modern automobile. In: *2010 IEEE Symposium on Security and Privacy*, pp. 447–462. IEEE, Los Alamitos (2010)
10. Gay, D., Levis, P., von Behren, R., Welsh, M., Brewer, E., Culler, D.: The nesc language: A holistic approach to networked embedded systems. In: *PLDI 2003: Proceedings of the ACM SIGPLAN 2003 Conference on Programming Language Design and Implementation*, pp. 1–11. ACM, New York (2003)
11. Madden, S.R., Franklin, M.J., Hellerstein, J.M., Hong, W.: Tinydb: an acquisitional query processing system for sensor networks. *ACM Trans. Database Syst.* 30(1), 122–173 (2005)
12. Karlof, C., Sastry, N., Wagner, D.: Tinysec: a link layer security architecture for wireless sensor networks. In: *SenSys 2004: Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems*, pp. 162–175. ACM, New York (2004)
13. (Infinite dimensional stochastic hybrid systems), <http://www.ece.ucsb.edu/~hespanha/projects/CCR-0311084-EHS03/>

14. Chatmon, C., van Le, T., Burmester, M.: Secure anonymous rfid authentication protocols. Technical Report TR-060112, Department of Computer Science, Florida State University, Tallahassee, Florida, USA (2006)
15. Broy, M.: Challenges in automotive software engineering. In: ICSE 2006: Proceedings of the 28th International Conference on Software Engineering, pp. 33–42. ACM, New York (2006)
16. MISRA-consortium: Guidelines for the use of the c language in vehicle based software (1998)
17. Levis, P., Lee, N., Welsh, M., Culler, D.: Tossim: accurate and scalable simulation of entire tinyos applications. In: SenSys 2003: Proceedings of the 1st International Conference on Embedded Networked Sensor Systems, pp. 126–137. ACM, New York (2003)
18. Bellare, M., Desai, A., Jokipii, E., Rogaway, P.: A concrete security treatment of symmetric encryption: Analysis of the des modes of operation. In: Proceedings of 38th Annual Symposium on Foundations of Computer Science, FOCS 1997 (1997)
19. Bellare, M., Kilian, J., Rogaway, P.: The security of the cipher block chaining message authentication code. J. Comput. Syst. Sci. 61(3), 362–399 (2000)
20. SWRL: A semantic web rule language combining OWL and RuleML, W3C member submission (2004), <http://www.w3.org/Submission/SWRL/>
21. Lee, H., Tsai, H.M., Tonguz, O.: On the security of intra-car wireless sensor networks, 1–5 (2009)

Appendix A: Tyre Pressure Monitoring Code

```

configuration TyrePressureAppC {
implementation
{
  components TyrePressureC, MainC, ActiveMessageC,
    new TimerMilliC(), new DemoSensorC() as PressureSensor,
    new AMSenderC(AM_PRESSURE) as PressureSenderC;

  TyrePressureC.Boot -> MainC;
  TyrePressureC.RadioControl -> ActiveMessageC;
  TyrePressureC.AMSend -> PressureSenderC;
  TyrePressureC.Timer -> TimerMilliC;
  TyrePressureC.Read -> PressureSensor;
}

```

Secure Automotive On-Board Protocols: A Case of Over-the-Air Firmware Updates

Muhammad Sabir Idrees¹, Hendrik Schweppe¹, Yves Roudier¹, Marko Wolf²,
Dirk Scheuermann³, and Olaf Henniger³

¹ EURECOM

{muhammad-sabir.idrees,hendrik.schweppe,yves.roudier}@eurecom.fr

² Escrypt GmbH

marko.wolf@escrypt.com

³ Fraunhofer SIT

{dirk.scheuermann,olaf.henniger}@sit.fraunhofer.de

Abstract. The software running on electronic devices is regularly updated, these days. A vehicle consists of many such devices, but is operated in a completely different manner than consumer devices. Update operations are safety critical in the automotive domain. Thus, they demand for a very well secured process. We propose an on-board security architecture which facilitates such update processes by combining hardware and software modules. In this paper, we present a protocol to show how this security architecture is employed in order to achieve secure firmware updates for automotive control units.

Keywords: Security protocols; security architectures; over the air firmware updates; software functionality.

1 Introduction

Current research activities in vehicular on-board IT architectures basically follow two key trends: unification of network communication and centralization of functionality. Recent on-board IT architectures comprise a very heterogeneous landscape of communication network technologies, e.g., CAN, LIN, FlexRay and MOST. Internet Protocol (IP) based communication is currently being researched as a technology for unifying the interconnection of electronic control units (ECUs) in future on-board communication systems [9]. In addition, there is a shift towards multipurpose ECUs and usage of flash memory technology in the microcontrollers. Besides these trends in the design of automotive on-board IT architectures, new external communication interfaces, fixed and wireless, are becoming an integral part of on-board architectures. One key factor for this development is the integration of future e-safety applications based on V2X communications (external communications of vehicles, e.g. with other vehicles – V2V, or with the infrastructure – V2I) [13][3] which have been identified as one promising measure for increasing the efficiency and quality of operational performance of all vehicles and corresponding intelligent transportation systems.

Firmware updates are crucial for the automotive domain, in which recalls are a very costly activity and thus should be avoided where possible. The practicability of remotely updating devices has been shown by Google for their Android telephones. With this, they have a powerful tool to react on discovered security flaws in very short time [21]. In the automotive domain, update intervals are calculated in quarters of a year and not quarters of a day right now. This paradigm is about to change and security mechanisms within the car provide the necessary building blocks. With the arising “always-connected” infrastructure, it will be possible to perform over-the-air (OTA) diagnosis and OTA firmware updates (see Fig. 1), for example. This will provide several advantages over hardwired access, such as saving time by faster firmware updates, which improves the efficiency of the system by installing firmware updates as soon as they are released by the car manufacturer. However, adding new in-vehicle services does not only facilitate novel applications, but also imposes stringent requirements on security, performance, reliability, and flexibility. As discussed in [8], in-vehicle components need not only be extremely reliable and defect free, but also resistant to the exploitation of vulnerabilities. Although on-board bus systems are not physically accessible (apart from via diagnostic interfaces), this provides only a limited degree of security for vehicles that are in wireless communication with other vehicles and devices (e.g., consumer devices connected to the vehicle).

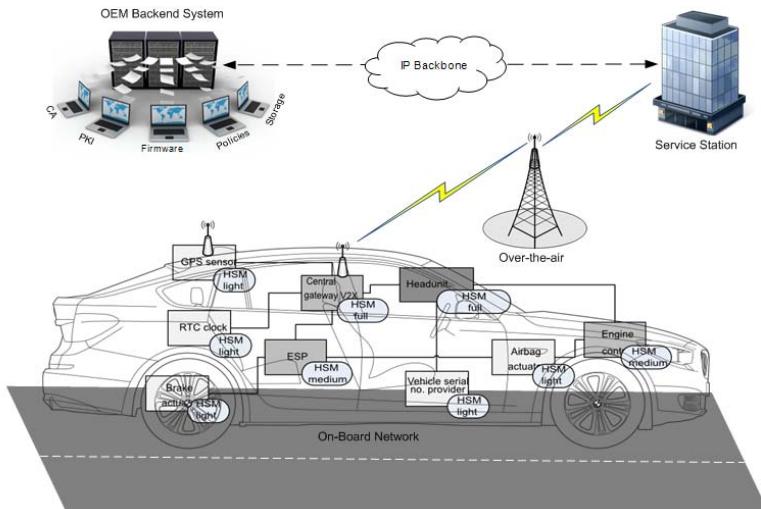


Fig. 1. Over-the-Air Firmware Updates

As seen in [8,5], attacks on the in-vehicle network have serious consequences for the driver. If an attacker can install malicious firmware, he can virtually control the functionality of the vehicle and perform arbitrary actions on the in-vehicle network [18]. Furthermore, since the ECU itself is an untrusted environment, there exist challenges in how to securely perform cryptographic operations (i.e., encryption/decryption, key storage). Thus, it does not make much

sense if the verifier software runs from the same flash as the software to be verified. In this paper we present a generic firmware update protocol, that can be used for both hardwired and remote wireless firmware flashing. The protocol has especially been designed with respect to the above mentioned functional and non-functional requirements. Our approach is based on hard security mechanisms. Hardware security measures are required in order to raise the security level of specific security services, e.g. for the storage of security credentials. We present a hardware security module design, which protects most critical parts of the architecture during firmware updates, such as secure key storage, secure operation of cryptographic algorithms, etc.

2 Secure On-Board Network Architecture

We believe that the combination of software and hardware based security solutions is necessary for meeting the requirements of on-board network security. However, depending on the risk level it should be analyzed for which use cases a security level using pure software security mechanisms is sufficient. Based on the security levels identified in [5], we focused on how to use hardware security services during firmware updates. Hardware security measures provided by hardware security modules (HSMs) are primarily used as root of trust for integrity measurement and responsible for performing all cryptography applications including symmetric/asymmetric encryption/decryption, symmetric integrity checking, digital signature creation/verification, and random number generation.

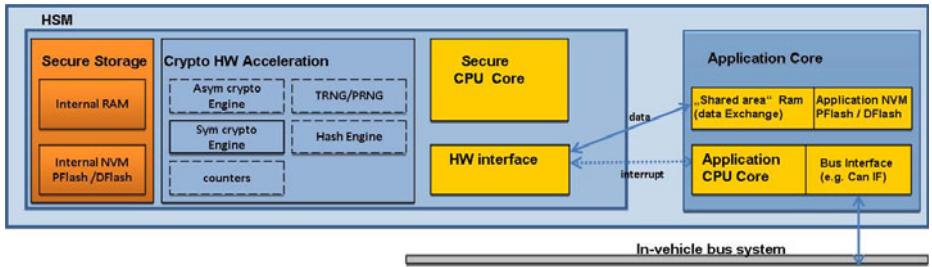


Fig. 2. HSM Architecture Overview

Considering the constraints of current vehicular on-board networks and the trend towards an increasing centralization of functions, a very flexible and scalable on-board HSM is required. The design of the HSM needs to consider the different available resources on sensors and actuators, ECUs and bus systems. We have decided to create three different variants of our HSM: The full, the medium and the light HSM. The simplest HSM is designed for sensors/actuators (see Fig. 1). At the ECU level, a more complex architecture can be applied (i.e., medium HSM), which for instance provides services for managing keys for the

on-board system and for protecting the ECU itself. In order to satisfy the performance requirements for signing and verifying messages for V2I communications (i.e., OTA firmware updates), a very efficient asymmetric cryptographic engine is required. Thus, the full HSM architecture is applied in this case, which provides the maximum level of functionality, security, and performance. Table 1 presents security features of the different HSM variants and comparison with other existing HSM approaches. These variants offer different levels of security functionality. The components of the HSM are divided into mandatory and optional components because, depending on the use case, different security requirements have to be fulfilled. The optional components are represented in Fig. 2 with dashed lines. Furthermore, it has been defined in compliance with the Secure Hardware Extension (SHE) specification proposed by the automotive HIS consortium [15]. All HSM modules (full to light) are able to understand and process SHE commands [4] accordingly.

Table 1. Components of different HSMs

Security Features	Hardware Security Module - HSM			SHE	TPM	Smart card
	Full	Medium	Light			
<i>Boot Integrity protection</i>	Auth & Secure	Auth & Secure	Auth & Secure	Secure	Auth	None
<i>HW crypto algorithms (incl. key generation)</i>	ECDSA, ECDH, AES/MAC, WHIRLPOOL / RSA, SHA1 HMAC	RSA, SHA1, AES/ MAC	AES/ MAC	AES/ MAC	RSA, SHA1 HMAC	ECC, RSA, AES, 3DES, MAC, SHA x..
<i>HW crypto acceleration</i>	ECC, AES, WHIRLPOOL	AES	AES	AES	None	None
<i>Internal CPU</i>	Programmable	Programmable	None	None	Preset	Programmable
<i>RNG</i>	TRNG	TRNG	PRNG w/ext. seed	PRNG w/ext. seed	TRNG	TRNG
<i>Counter</i>	16x64bit	16x64bit	None	None	4x32bit	None
<i>Internal NVM</i>	Yes	Yes	Optional	Yes	Indirect (Via SRK)	Yes
<i>Internal Clock</i>	Yes w/ext UTC Sync	Yes w/ext UTC Sync	Yes w/ext UTC Sync	No	No	No
<i>Parallel Access</i>	Multiple sessions	Multiple sessions	Multiple sessions	No	Multiple sessions	No
<i>Tamper Protection</i>	Indirect (passive, part of ASIC)	Indirect (passive, part of ASIC)	Indirect (passive, part of ASIC)	Indirect (passive, part of ASIC)	Yes (mfr.dep.)	Yes (active, up to EAL5)

There exist different approaches for integrating the HSM with the microcontroller: i). HSM in the same chip as the CPU with a state machine and ii). HSM in the same chip as the CPU but with a programmable secure core. In the solution we propose, a programmable CPU core is inside the same chip as the main microcontroller to perform cryptography operations (see Fig. 2). Note that when a software-based cryptography implementation is used, it can be easily modified (possibly not a highly efficient solution) but changing a state machine requires that hardware to be redesigned and is very expensive. It is necessary that the HSM be in the same chip as the application CPU and contains a microprocessor, to protect it from physical tampering.

3 Implementing Security Primitives

In this section we briefly review the hardware interface for the invocation of cryptographic hardware security blocks, higher-level security functionality and security management functionality (e.g., key import/export, signature, and message authentication code) that are required during OTA firmware update. More details about HSM functional calls/descriptions can be found in [6].

Signature: This function is used for demonstrating the authenticity and integrity of a message. A valid signature gives a recipient reason to believe that the message was created by a known sender, and that it was not altered in transit. For signature generation, a signature generation scheme $\text{sig}(m)_{S_k}$ takes as input a key k , and message m , outputs a signature $\hat{\sigma}$; we write $\text{sig}(m)_{S_k} = \{\hat{\sigma}\}_{S_k}$. Where k is the security parameter, outputs a pair of keys $(S_k; V_k)$. S_k is the signing key, which is kept secret, and V_k is the verification key which is made public. We also assume that a time stamp (UTC Time) is generated and then also covered by the signature calculation, and write $\vec{m} = (m + \text{Ts})$ to denote the message and a time stamp whose signature is $\hat{\sigma}$. For the signature verification, $\text{ver_sig}(\vec{m}, \hat{\sigma})_{V_k} \rightarrow \alpha$ function is defined, takes as input the signature $\hat{\sigma}$, the signature verification public key part V_k , and outputs the answer α which is either succeed (signature is valid) or fail (signature is invalid). As a precondition, the V_k must be loaded and enabled for verification.

Message Authentication Code – MAC: This function is used to protect both the data integrity and the authenticity of a message, by allowing verifiers (who also possess the secret key) to detect any changes to the message content. For generating a MAC as well as the message itself, the notation $\text{MAC}(m)_{M_k} = \{\hat{m}\}_{M_k}$ is used, so that it produces the message itself plus the cryptographic authentication code based on M_k and m . Here, M_k refers to a cryptographic key for MAC generation and m to the message to be authenticated. In the same way as for signatures, the use of the time stamp $\vec{m} = (m + \text{Ts})$ is covered by the MAC calculation. For the verification of a MAC, the notation $\text{ver_MAC}(\vec{m}, \hat{m})_{M_k}$ is used. Based on the M_k , it is verified whether \hat{m} corresponds to the message \vec{m} .

Key Creation: This security building block is used for the creation of a key on a hardware module, using HSM `Create_Random_Key` function. All properties of the key are determined and fixed during creation. This includes the cryptographic algorithm to be used, the use and further property use flags indicating what actions may be done with this key (i.e., sign and verify) as well as the authorization data needed for key usage. The `use_flag` parameter indicates the operations that may be performed with the key. In particular, the following flags are present:

- `sign|verify`: Key can be used to generate and/or verify digital signatures or H/MACs of any data.
- `encrypt|decrypt`: Key can be used to encrypt and/or decrypt any data.
- `secureboot`: Can be used to create/verify secure boot references.

- **keycreation:** Can be used for creation of new keys, e.g. via key derivation functions (symmetric) or DH key agreement (asymmetric).
- **securestorage:** Can be used to realize (locally bound) secure storage
- **utcsync:** Can be used for synchronizing internal tick counter to UTC.
- **transport:** Can be used to protect transports of keys (i.e., migration, swapping, move) between locations, according to individual transport flags (i.e., 0 = INT, 1 = MIG, 2 = OEM, 3 = EXT) [6].

Only the `use_flag` may explicitly be set by the creator whereas further property flags are set inherently. Once created, the key properties are unchangeable. As output, the function delivers a key handle for later usage of the key.

Key Export: With this function, keys stored on an HSM are transported to other HSMs or to other trusted parties. During transport, the key is encrypted ($\epsilon(k)_{T_k}$) with a special transport key (T_k) that may be symmetric or asymmetric. In addition, the authenticity of the key is protected by a key authenticity code which consists in a MAC or a signature appended to the encrypted key. The key authenticity code can be an explicit symmetric key enabled with `use_flag = verify` or an implicit symmetric key derived from a symmetric transport key or an implicit asymmetric key also enabled with `use_flag = verify`. The use of this key authenticity code is mandatory. As output, the function delivers the encrypted key together with its authentication code. As an important precondition, the specified transport key must be loaded and enabled to be used for transport. Furthermore, the transport flag of the key to be exported must be appropriately marked according to the type of module managing the transport key.

Key Import: This function is used for importing keys into another HSM or to other trusted parties. In this way, the function provides the counterpart to the previously described export function. The key may be imported either into the non-volatile memory or into the main memory (RAM) of the HSM. In the same manner as for Key_Export, the use of the key authenticity code is mandatory. As output, the function delivers a key handle to reference the key for later usage. As a precondition, the transport key must be loaded and enabled before. In addition, the authentication code verification key must be loaded if the key is protected by a signature.

Key Master – KM: We introduce a new functional entity, which we call the *KeyMaster*. As there exist multiple variants of the HSM, that support different cryptographic keys (symmetric/asymmetric), we had to take this into account for key distribution. The KM is a central element in the establishment of a session between entities. It holds public key (P_k) and pre-shared keys (Psk) of the individual ECUs, which are used as transport keys, to establish a secure session. This functional entity resides on a dedicated ECU or is integrated into another ECU. There may be more than one KM node in a vehicle for replication purposes.

Counter: As a further instrument to control the behavior of the HSM, the possible use of counters is introduced as an additional security building block. The concrete, central task of a counter is freshness enforcement to prevent different kinds of replay attacks. For handling these counters, the following HSM functions are provided: `Create_Counter`; `Read_Counter`, `Increment_Counter`, and `Delete_Counter`. Access authorization data needs to be provided as input data, and is later necessary to create, increment or delete the counter.

4 Secure Firmware Update Protocol

Ensuring secure firmware updates requires authenticity, integrity, freshness, and confidentiality, as identified in [5]. The specified protocol provides means to satisfy these security requirements. The next section will now show how HSMs are used and how they interact in order to ensure firmware updates are downloaded and installed securely into the vehicle.

4.1 OTA Update Requirements

Before sketching the protocol, we describe some additional constraints and requirements that have to be taken into account for secure firmware updates:

HSM in the Diagnostic Tool: As stated in [5], there are numerous scenarios, where an attacker targets the diagnostic tool (DT). For instance, the attacker might inject bogus authority keys into the ECU, through DT, which compromises the overall security of the vehicular on-board architecture. In particular, this means that the DT stores challenges and public strings for key recovery (i.e., ECU unlock key) and is therefore responsible for the security of the subsystem. Therefore, this information needs to be stored securely on the DT-side. An additional advantage of HSM is the resistance against physical tampering of the DT. Any damage to the HSM changes the behavior and therefore prevents the extraction of secret key material.

Bandwidth Limitations of In-vehicle Networking Technologies: Firmware update protocols comprise two parts: a V2I part, and an intra-vehicular part, the latter involving a large number of interconnected ECUs. In the on-board bus systems used, a specific restriction lies in the limited size of data packets. For the CAN bus, for example, this means that only eight bytes of payload may be transmitted at a time. For this purpose, secure common transport protocols (S-CTP) [7], extensions of the CTP defined in [2] are applied to diagnosis jobs, where typically larger data chunks need to be transmitted.

4.2 Protocol Description

Remote Diagnosis: In the OTA firmware update scenario, a service station using a DT connects remotely to a vehicle, using V2I communication, to assess the state of the vehicle (see Fig. 10). To know which version is installed, a diagnosis of the vehicle is required to have all necessary information

such as ECU type, firmware version, and date of last update. An employee of the station using the DT establishes a secure connection with the vehicle, at the ECU level, in order to determine the current state of the vehicle. To do so, DT creates a session key M_{sk} (exportable), by sending a HSM command `Create_Random_key` and specifies the set of allowed key properties such as, `use_flag= sign|verify, encrypt| decrypt`. It then calls `export` the M_k using $P_k\text{-ccu}$ (Public key of the central communication unit – CCU) as a transport key (T_k) and transmits it to the vehicle. Here, the CCU is the first receiving entity in the vehicle, responsible for receiving and distributing V2X messages to the in-vehicle network.

In the vehicle, the CCU, equipped with the Full HSM and acting as a KM node, receives the connection request. The authorization for the connection is verified in the CCU. The message \vec{m} is checked for freshness, integrity and the service station is authenticated. If the check succeeds, CCU-KM imports the key into the HSM. It then exports the received $Exported_M_k$ with the corresponding $P_k\text{-ecu}$ or Psk , depending on the ECU type, and distributes it to the target ECU in order to enable end-to-end communication. This message includes all information that is necessary to deliver this message to the correct ECU. On the receiving side, ECU verifies the integrity, authenticity and authorizations of CCU/DT based on the policy as to whether DT is allowed to deliver a message or not. If this is true, and the message is fresh, ECU imports the M_k in the HSM. Once M_k key have imported, an acknowledgment is sent back to DT (see Algorithm 1). After this acknowledgment frame, the DT sends, depending on the option chosen by the employee of the service station, requests to read out diagnosis information (State/Log information) from the ECU it wants to check.

Algorithm 1. Remote Diagnosis

Require: Signature verification Key V_k of DT, CCU, ECU are pre-loaded

Ensure: Establishing a fresh and authentic session between DT and ECU based on a symmetric session key M_k , where CCU-KM acts as a Key Master Node.

- DT \rightleftarrows HSM: $M_k\text{-handle} := create_random_key (use_flag = sign| verify, encrypt| decrypt, export, use_authorization_data)$
DAT: $Exported_M_k := key_export(T_k\text{-handle} = \langle P_k\text{-ccu-handle} \rangle, kh = \langle M_k\text{-handle} \rangle)$
 - 1. DT \rightarrow CCU-KM: $\{(Exported_M_k, Ts), \{\hat{\sigma}\}_{S_k\text{-dt}}\}$
 - CCU-KM \rightleftarrows HSM: $M_k\text{-handle} := key_import (T_k\text{-handle} = \langle S_k\text{-ccu-handle} \rangle, kh = \langle M_k\text{-handle} \rangle)$
 - CCU-KM: \rightleftarrows HSM: $Exported_M_k := key_export(T_k\text{-handle} = \langle P_k\text{-ecu-handle} | Psk\text{-handle} \rangle, kh = \langle M_k\text{-handle} \rangle)$
 - 2. CCU-KM \rightarrow ECU: $\{(Exported_M_k, Ts), \{\hat{\sigma}\}_{S_k\text{-ccu}}\}$
 - ECU \rightleftarrows HSM: $M_k\text{-handle} := key_import (T_k\text{-handle} = \langle S_k\text{-ecu-handle} | Psk\text{-handle} \rangle, kh = \langle M_k\text{-handle} \rangle)$
 - 3. CCU-KM \leftarrow ECU: $\{(ACK, Ts), \{\hat{\sigma}\}_{S_k\text{-ecu}}\}$
 - 4. DT \leftarrow CCU-KM: $\{(ACK, Ts), \{\hat{\sigma}\}_{S_k\text{-ccu}}\}$
-

Advance Notification: Due to legal reasons and to allow for flexible deployment, we consider that service station will send an advance notification of possible firmware updates, if the type is the expected one. This advance notification is intended to help customers plan for the effective deployment of

updates, and includes information about the number of new updates being released. These updates still need to be Approved for install before downloading. The customer receives this information on the vehicle human-machine interface (HMI) and can decide about possible deployment (i.e., `Install`, `Decline`, `Decide later`). Only updates that have the approval status `Install` will be downloaded to the vehicle. Disabling any ECU while vehicle is running may cause safety critical problems, depending on the function ECU is responsible for. We thus assume that additional checks will be performed by the on-board system, to ensure that the vehicle is stopped and has access to the infrastructure, before switching the ECU into the re-programming mode. Furthermore, we assume that the V2I communication is available throughout the OTA firmware update process.

ECU Reprogramming Mode: If the type is the expected one, the DT forces the ECU to switch from an application mode into a reprogramming mode by requesting a seed (Na). This seed is required to calculate an ECU specific key value to unlock the ECU for reprogramming. The ECU verifies desired security properties. If it is true, ECU sends a HSM command `SecM_Generate(seed)` to generate a seed. It then encrypts the seed $\epsilon(Na)_{M_k}$ for confidentiality enforcement, compute a MAC on $\vec{m} = (\epsilon(Na)_{M_k}, Ts)$ and transmits it to the DT. At the same time, the ECU sends a HSM command `SecM_ComputeKey(Na, SecM_key)` to compute the key on the HSM using Na . As output, the function delivers a $SecM_{key}$ key handle, we write $SecM_{key}=S_{mk}$, that is used to unlock the ECU.

Algorithm 2. ECU Reprogramming Mode

Require: DT and ECU have established a fresh and authentic connection based on a M_k . Vehicle is stopped and have access to infrastructure

Ensure: Authentic and confidential exchange of ECU unlock key.

1. $DT \rightarrow ECU: \{ (request_seed, Ts), \{\hat{m}\}_{M_k} \}$
– $ECU: \xrightarrow{\text{HSM}} \text{SecM_Generate}(seed)$
 2. $DT \leftarrow ECU: \{ (\epsilon(Na)_{M_k}, Ts), \{\hat{m}\}_{M_k} \}$
– $ECU: \xrightarrow{\text{HSM}} S_{mk} := \text{SecM_ComputeKey}(seed, SecM_key)$
– $DT: \xrightarrow{\text{HSM}} S_{mk} := \text{SecM_ComputeKey}(seed, SecM_key)$
– $DT \xrightarrow{\text{HSM}} Exported_S_{mk} := \text{key_export}(T_k\text{-handle} = < M_k\text{-handle} >, kh = < S_{mk}\text{-handle} >)$
 3. $DT \rightarrow ECU: \{ (Exported_S_{mk}, Ts), \{\hat{m}\}_{M_k} \}$
– $ECU: \xrightarrow{\text{HSM}} \text{SecM_CompareKey}(key, seed)$
 4. $DT \leftarrow ECU: \{ (ACK, Ts), \{\hat{m}\}_{M_k} \}$
-

The DT verifies $\{\hat{m}\}_{M_k}$, decrypts the received seed $(\epsilon^{-1}(Na)_{M_k})$, and computes the S_{mk} with the aid of the received seed (Na). Once the S_{mk} key value is computed, it is exported, using M_k as a transport key, and transmitted to the target ECU. The ECU verifies the $\{\hat{m}\}_{M_k}$ and compares the received S_{mk} key with the self-generated S_{mk} . If the two values are identical, the ECU is switched into unlock state (from application mode to the reprogramming mode) and sends an ACK message to the DT (see Algorithm 2). This message is sent after the ECU is switched into the unlock state to make sure the switch has been performed. The information whether a re-programming request has been received

or not shall be stored in non-volatile memory, e.g. EEPROM. Since switching from the application to the reprogramming mode shall be done via a hardware reset, all contents of volatile memory will be lost [15]. If the comparison failed, the flashloader [15] holds the ECU in locked state. ECU reprogramming is possible only in the unlocked state.

Firmware Encryption Key Exchange: In this phase we are considering two possible scenarios for exchanging firmware encryption keys: i) on-line solution and ii) off-line solution. In the on-line solution: the service station has access to an online infrastructure of the manufacturer, it can request the firmware and as well as the firmware encryption key – (*SSK*). The *SSK* is a stakeholder symmetric key pair [7], created externally, with *use_flag=decrypt*, key for stakeholder individual usage e.g., software update. Instead, in the case of off-line firmware is encrypted with the pre-installed *SSK*.

Considering current trends and advancements in the automotive industry, on-line solutions provide more reliability, flexibility and will eventually increase the security of the on-board network. Sharing the firmware encryption key only with specific ECUs makes an on-line solution more robust and generic compared with off-line approaches, where all vehicles share unique symmetric keys that are pre-installed in the vehicles. In addition, the existence of various security levels in the architecture [5], pleads for the specification of a validity period of the *SSK* (short term or long term keys), for an individual ECU. We suggest to use short term keys for firmware encryption. Short terms keys will expire after a short amount of time and thus, as there is no need for instant revocation if keys are compromised. This has the advantage that OEMs do not have to go through another key migration (installing new keys) process if keys are compromised. As such, the following section only details the on-line solution.

Algorithm 3. Firmware Encryption Key Exchange

Require: On-line access to OEM server and PKI infrastructure

Ensure: Authentic and confidential firmware encryption key exchange between OEM and ECU

1. $\text{DT} \rightarrow \text{OEM}: \{(request_firmware_encryption_key, Ts), \{\hat{o}\}_{Sk_{dt}}\}$
– OEM: *Exported_SSK:=key_export(T_k-handle=<P_k_ecu-handle|Psk-handle>, kh=<SSK-handle>)*
 2. $\text{DT} \leftarrow \text{OEM}: \{(Exported_SSK, Ts), \{\hat{o}\}_{Sk_{oem}}\}$
 3. $\text{DT} \rightarrow \text{ECU}: \{(Exported_SSK, Ts), \{\hat{m}\}_{M_k}\}$
– ECU: \rightleftarrows HSM: *SSK-handle :=key_import (T_k-handle =<S_k_ecu-handle|Psk-handle>, kh= <SSK-handle>)*
 4. $\text{DT} \leftarrow \text{ECU}: \{(ACK, Ts), \{\hat{m}\}_{M_k}\}$
-

After successfully reprogramming access at the ECU level, the DT sends a request (*request_firmware_encryption_key*) to the OEM server to get the firmware encryption key (see Algorithm 3). This request includes information about the ECU (i.e., ECU type, ECU identification number, firmware version, etc.). The OEM verifies the authenticity and integrity of the received message. If verified, OEM server retrieve the *Pk_ecu* from the Public Key Infrastructure (PKI), (possibly) maintained by an individual OEM, and exports *SSK* using

$P_k\text{-}ecu$ as a T_k . This is only feasible if the ECU is equipped with a full HSM. In the case of medium or light HSM-ECUs, the pre-shared key Psk will be used as a T_k . The OEM server exports the SSK and sends a signed message to the DT. As the SSK key blob is encrypted with the ECU key, It is not possible for the DT to retrieve the firmware encryption key. Next, the DT transmits the received firmware encryption key to the ECU. The ECU imports the SSK in the HSM using the `key_import` function. The `key_import` function provides the assurance to the ECU that the key is generated by the OEM, by verifying the authentication code send along with the encrypted key, and can only be decrypted by the specific ECU key. After importing the SSK in the ECU-HSM, the ECU sends an `acknowledgment` about the successful import of the SSK .

Firmware Download: Once the SSK is successfully imported into the ECU-HSM, the DT sends the received signed and encrypted firmware (\mathbb{F}_{rm}) along with its ECU Configuration Register (ECR) reference: $sig(\mathbb{F}_{rm}, \mathbb{E}_{cr}, Ts) \rightarrow \widehat{\sigma_{\mathbb{F}_{rm}}}^{\epsilon_{nc}} \rightarrow \epsilon(\widehat{\sigma_{\mathbb{F}_{rm}}})_{ssk}$, to the Random-Access Memory (RAM) of the ECU. Following the HSM `use_flag` approach, where multiple key-properties may be set, only the OEM server can sign and encrypt the firmware, whereas the receiving ECU can decrypt and verify the received firmware, using the same key material. The encrypted firmware is downloaded block by block (logical block). Each of those blocks is divided into segments, which are a set of bytes containing a start address and a length. The start address and the length of each segment is sent to the HSM during the segment initialization. For one block, a download request is sent from the DT to the HSM. The HSM initializes the decryption service and sends an answer to the DT. The download then starts segment by segment. After sending the last firmware segment, the DT sends a `transfer_exit` message to the ECU (see Algorithm 4).

Algorithm 4. Firmware Download

Require: Signed and encrypted firmware from OEM

Ensure: Authentic, fresh and Confidential firmware downloaded in the ECU

1. $DT \rightarrow ECU: \left\{ \left(\epsilon(\widehat{\sigma_{\mathbb{F}_{rm}}})_{SSK}, Ts \right), \{\hat{m}\}_{M_k} \right\}$
 - $ECU \rightleftharpoons HSM: \text{SecM_InitDecryption}(\epsilon(\widehat{\sigma_{\mathbb{F}_{rm}}})_{ssk})$
 - $ECU \rightleftharpoons HSM: \text{SecM_Decryption}(\epsilon^{-1}(\widehat{\sigma_{\mathbb{F}_{rm}}})_{ssk})$
 2. $DT \rightarrow ECU: \left\{ (request_transfer_exit, Ts), \{\hat{m}\}_{M_k} \right\}$
 - $ECU \rightleftharpoons HSM: \text{SecM_DeinitDecryption}()$
 3. $DT \leftarrow ECU: \left\{ (ACK, Ts), \{\hat{m}\}_{M_k} \right\}$
-

Firmware Installation and Verification: For an installation of the firmware, we consider the standard firmware installation procedure defined in [15], where each logical block is erased and reprogrammed. However, before the flash driver can be used to re-program an ECU, its compatibility with the underlying hardware, the calling software environment and with prior versions of the firmware has to be checked. This compatibility check is performed by means of a version information stored in the HSM monotonic counters. The HSM `Read_Counter` function is used to read out the value of a counter. The counter is referenced by

a *counter-identifier* previously increased after every authentic and successful installation of the firmware. These monotonic counters are defined to perform such a checking of its current version against the new firmware version in order to prevent the downgrading attacks meant to install older firmware.

For the verification, we defined a two step verification process: In the first step, before re-programming, the ECU verifies the signature of the firmware data. This is verified by using the pre-installed Manufacturer Verification Key *MVK*. It proves that the software was indeed released from the OEM. In the second step: we construct a tiny trusted computing base (TCB) during the installation phase. We compute an ECR trusted chain at each step of the firmware installation. The ECR reference is needed to ascertain the integrity/authenticity of the firmware data. An *Extend_ECR* function is defined to build the ECR trusted chain. This function is used for updating the ECR with a new hash value. The new value is provided as input and chained with the existing value stored in the ECR, using a hash update function. As output, the function delivers the updated ECR value.

After a successful installation of the new firmware data, software consistency check is performed. The check for software dependencies shall be done by means of a callback routine provided by the ECU supplier. This check is done after reprogramming and before setting the new ECR reference. Next, the *Compare_ECR* function is called. This comparison can only be performed after all writing procedures for the logical block have been finished. This function allows the direct comparison of the current ECR with a reference ECR value received with the firmware. It is also possible that the ECR reference may be contained inside the firmware itself. In this case the flashloader shall call a routine provided by the ECU supplier to obtain the ECR reference. If the check succeed, the *HSM Preset_ECR* function is called. This function is used to manage references to ECR values by ECR indices in the context of a secure boot. After successfully setting the ECR value, the HSM (*Increment_Counter*) function is called to increment the monotonic counter with the new value. At the last step, the actual hardware reset is executed, the flashloader deletes (i.e. over-writes) the routines for erasing and/or programming the flash memory from the ECU's RAM [15], thereby making sure those routines are not present on the ECU in application mode. After the reset, the application is started.

Error Handling: Each function of the HSM returns a status after its successful or unsuccessful execution. Some functions may deliver further function specific error codes. The value of the status shows the positive execution of the function or the reason for the failure. In case of a failure, the flash process must stop with an error code and the ECU enters the locked state.

5 Related Work

The past decade has seen a tremendous growth in the vehicular communication domain, yet no comprehensive security architecture solution has been defined that covers all aspects of on-board communication (data protection, secure communication, secure and tamper proof execution platform for applications). On

the other hand, several projects, namely GST [10], C2C-CC [3], IEEE Wave [22] and SeVeCOM [20] have been concerned with inter-vehicular communication and have come up with security architectures for protecting V2X communications. These proposals essentially aim at communication specific security requirements in a host-based security architecture style, as attackers are assumed to be within a network where no security perimeter can be defined (ad-hoc communication). These proposals consider the car mostly as a single entity, communicating with other cars using secure protocols.

Mahmud et al. [14] present a security architecture and discuss secure firmware upload, which depends however on a number of prerequisites and assumptions (i.e., sending multiple copies to ensure firmware updates) in order to make secure firmware update. However, sending multiple copies is not realistic and imposes several constraints on the infrastructure. This proposal does not consider automotive on-board networks, where domains are traditionally separated, due to functional and non-functional requirements. Kim et al. [12] present remote progressive updates for flash-based networked embedded systems. In their solution a link-time technique is proposed which reduces the energy consumption during installation. However, no security concern is addressed in this proposal.

Nilsson et al. discuss in [16][17] provide a lightweight protocol and verification for secure firmware updates over the air (SFOTA). In the SFOTA protocol, different properties are ensured during firmware update protocol (i.e., data integrity, data confidentiality, and data freshness). However, this approach also relies on strong imposed assumptions in order to ensure the secure software upload: the authentication of the vehicle is not considered, keys are assumed to be stored securely and the authors use a single encryption key for all the ECUs in a car. Furthermore, no specific execution platform requirements are put forward by this proposal. In [18], key management issues are discussed in relation with software updates. A rekeying protocol is defined in order to distribute keys with only specific nodes in the group. It also uses a multicast approach to update the software on a group of node. Furthermore, as mentioned above, this approach also does not consider execution platform requirements. It does not discuss about computation attacks, where the attacker can learn and modify the firmware, during the installation phase or simply prevent to update the counter, for later replay attacks.

Hagai [19] presents an approach that takes hardware into account by providing a secured runtime environment with a so-called Trust Zone on an ARM processor. In contrast the solutions of [10][11] are software based. The so called *tools* and *enablers*, which are low-level and application-level security functions in [1] also cover a number of on-board automotive use-cases, while leaving the essential link to the external communication domain uncovered. The approach most closely related to our work is that of the Hersteller-Initiative Software – HIS [15]. The flashing process defined by the HIS provides a good basis for the OEMs, but the recommended protocol does not provide all the necessary security functionalities (i.e., freshness). Furthermore, this process only addresses

hardwired firmware updates and does not provide any information about which key is used for firmware encryption, in a heterogeneous landscape of communication network technologies.

6 Conclusion

We have presented a firmware update protocol for a new security architecture to be deployed within the vehicle. We showed how a root of trust in hardware can sensibly be combined with software modules. These modules and primitives have been applied to show how firmware updates can be done securely and over-the-air, while respecting existing standards and infrastructure. In contrast to existing approaches, the protocols presented in this paper describe a complete process, which involves the service provider, the vehicle infrastructure as well as the manufacturer and the workshop. By using secure in-vehicle communication and a trusted platform model, we show how to establish a secure end-to-end link between the manufacturer, the workshop and the vehicle. Despite the fact that a trusted platform model entails certain constraints, such as the obligation to bind cryptographic keys to a given boot configuration, we showed how the protocols we presented deal with the update of the platform reference registers during the boot phase of an ECU.

Acknowledgments

This work has been carried out in the EVITA (E-safety Vehicle Intrusion proTected Applications) project, funded by the European Commission within the Seventh Framework Programme for research and technological development.

References

1. Bar-El, H.: Intra-vehicle information security framework. In: Proceedings of the 7th escar Conference, Düsseldorf, Germany (2009)
2. Busse, M., Pleil, M.: Data exchange concepts for gateways. Technical Report Deliverable D1.2-10, EASIS Project (2006)
3. C2C-CC. Car2Car Communication Consortium,
<http://www.car-to-car.org/>
4. Escherich, R., Ledendecker, I., Schmal, C., Kuhls, B., Grothe, C., Scharberth, F.: SHE – Secure Hardware Extension – Functional Specification Version 1.1
5. Ruddel, A., et al.: Security Requirements for Automotive On-Board Networks based on Dark-side Scenarios. Technical Report Deliverable D2.3, EVITA Project (2009)
6. Weyl, B., et al.: Secure On-board Architecture Specification. Technical Report Deliverable D3.2, EVITA Project (2010)
7. Schweppe, H., et al.: Secure On-Board Protocols Specification. Technical Report Deliverable D3.3, EVITA Project (2010)
8. Koscher, K., et al.: Experimental Security Analysis of a Modern Automobile. In: Proc. of the 31st IEEE Symposium on Security and Privacy (May 2010)
9. Rahmani, M., et al.: A novel network architecture for in-vehicle audio and video streams. In: IFIP – BcN (2007)

10. GST. Global systems for telematics, EU FP6 project, <http://www.gst-forum.org/>
11. Hergenhan, A., Heiser, G.: Operating Systems Technology for Converged ECUs. Embedded Security in Cars (2008)
12. Kim, J., Chou, P.H.: Remote progressive firmware update for flash-based networked embedded systems. In: ISLPED 2009, pp. 407–412 (2009)
13. Kosch, T.: Local Danger Warning based on Vehicle Ad-hoc Networks: Prototype and Simulation. In: WIT 2004, pp. 3–7 (2004)
14. Mahmud, S.M., Shanker, S., Hossain, I.: Secure software upload in an intelligent vehicle via wireless communication links. In: Proc. IEEE Intelligent Vehicles Symposium, pp. 588–593 (2005)
15. Miehling, T., Vondracek, P., Huber, M., Chodura, H., Bauersachs, G.: HIS flashloader specification version 1.1. Technical report, HIS Consortium (2006)
16. Nilsson, D.K., Larson, U.E.: Secure Firmware Updates Over the Air in Intelligent Vehicles. In: Proc. ICC Workshops (2008)
17. Nilsson, D.K., Sun, L., Nakajima, T.: A Framework for Self-Verification of Firmware Updates Over the Air in Vehicle ECUs. In: GLOBECOM (2008)
18. Nilsson, D.K., et al.: Key management and secure software updates in wireless process control environments. In: WiSec 2008 (2008)
19. Towards a secure automotive platform. White paper, secunet (2009)
20. SeVeCOM. Secure Vehicle Communication, <http://www.sevecom.org/>
21. Shabtai, A., Fledel, Y., Kanonov, U., Elovici, Y., Dolev, S.: Google Android: A State-of-the-Art Review of Security Mechanisms (2009)
22. IEEE WAVE. Wireless Access in Vehicular Environments, IEEE standard 1609.2

Author Index

- Aguado, Marina 69
Akbar, Mohammad Sajjad 95
Alonso, Luis 11
Alvi, Atif 216
Astorga, Jasone 69
Berbineau, Marion 83
Bouvry, Pascal 131
Castelain, Damien 23
Cocheril, Yann 83
Collado, Carlos 11
Corlay, Patrick 83
Coudoux, François-Xavier 83
Damas, Luís 143
Danoy, Grégoire 131
de Ponte Müller, Fabian 203
Diwanji, Vivek 58
Ehnert, Philipp 155
Facchi, Christian 106
Fatani, Imade Fahd Eddine 83
Ferreira, Michel 143
Flores, Enrique 11
Gomes, Pedro 143
González, Jesús 11
González-Arbesú, José M. 11
Hartenstein, Hannes 189
Heirich, Oliver 45
Henniger, Olaf 224
Herrscher, Daniel 165
Hierro, José 11
Higuera, Jorge 11
Huarte, Maider 69
Idrees, Muhammad Sabir 224
Islam, Shahidul 119
Kartsakli, Elli 11
Klemp, Oliver 189
Kloiber, Bernhard 176, 203
Kucharzyk, Uwe 1
Laya, Andres 11
Lehner, Andreas 45
Lill, Dirk 119
Lim, Hyung-Taek 165
Mangel, Thomas 189
Martínez, Raquel 11
Matias, Jon 69
Mehmood, Rashid 216
Michl, Matthias 189
Nabi, Zubair 216
Navarro, Isabel 11
Noureddine, Hadi 23
Nsiala, Crépin 83
Olaverri-Monreal, Cristina 143
Pigné, Yoann 131
Pyndiah, Ramesh 23
Qayyum, Amir 95
Rao, Subrahmanya Venkata Radha Krishna 58
Rasheed, Asim 95
Reveriego Sierra, Juan María 203
Rico García, Cristina 45
Röckl, Matthias 203
Röglinger, Sebastian 106
Roudier, Yves 224
Schappacher, Manuel 119
Scheuermann, Dirk 224
Schlichter, Johann 155
Schmidt, Robert K. 176
Schüttler, Florian 176
Schweiger, Benno 155
Schweppe, Hendrik 224
Sikora, Axel 119
Strang, Thomas 45, 176, 203
Valenzuela, José Luis 11
Villeforceix, Bernadette 34
Vlad, Adrian 11
Weckemann, Kay 165
Wolf, Marko 224
Zwingelstein-Colin, Marie 83