Rate-Adaptation Based Congestion Control for Vehicle Safety Communications

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

von der Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Leonie Tessa Tielert

aus München

Tag der mündlichen Prüfung:	25.06.2014
Erster Gutachter:	Prof. Dr. rer. nat. Hannes Hartenstein Karlsruher Institut für Technologie (KIT)
Zweiter Gutachter:	Prof. Dr. Marco Gruteser Rutgers University, New Jersey

Zusammenfassung

Im Rahmen der auf dem W-LAN-Standard IEEE 802.11p basierenden Fahrzeug-zu-Fahrzeug-Kommunikation tauschen letztere periodisch kurze Statusmeldungen untereinander aus, die ihre aktuelle Position, Geschwindigkeit und Fahrtrichtung enthalten. Basierend auf diesen empfangenen Nachrichten kann ein Fahrzeug entscheiden, ob eine potentiell gefährliche Situation wie ein unsicherer Spurwechsel vorliegt und den Fahrer entsprechend warnen. Falls der Austausch dieser Nachrichten nicht geregelt wird, kann bei höherer Fahrzeugdichte der Kommunikationskanal überlastet werden, was den Empfang der Nachrichten und schließlich die Funktionalität der entsprechenden Fahrerassistenzsysteme beeinträchtigen kann. Die vorliegende Arbeit entwickelt systematisch eine Entwurfsmethodik für die Staukontrolle auf dem Fahrzeug-zu-Fahrzeug-Kommunikationskanal und präsentiert schließlich ein resultierendes Protokoll namens PULSAR. Im Gegensatz zu den meisten existierenden Ansätzen, die sich typischerweise ausschließlich auf die Regelung der Kanalauslastung konzentrierten, stellt die vorliegende Arbeit einen Ansatz vor, wie die Regelung innerhalb durch die Sicherheitsanwendungen vorgegebener Grenzen geschehen kann. Indem statt der absoluten die *relative* Senderate angepasst wird, zielt der vorgestellte Ansatz darauf ab, den teilnehmenden Fahrzeugen einen vergleichbaren Sicherheitsnutzen statt - wie typischerweise sonst in Kommunikationsnetzen üblich - den gleichen Anteil an der verfügbaren Kanalkapazität zur Verfügung zu stellen. Des weiteren gingen bisherige Ansätze typischerweise weder darauf ein, was der gewünschte Zielwert der Regelung wäre, noch darauf, warum der gewählte Ansatz zu diesem Ergebnis konvergieren würde. Die vorliegende Arbeit präsentiert bezüglich der ersten Frage eine ausführliche Analyse und bezüglich der zweiten Frage einen Regelungsansatz, der die Senderate und die Sendeleistung eines Fahrzeuges gleichzeitig anpasst. Im Gegensatz dazu konzentrierten sich bisherige Ansätze typischerweise nur auf die Anpassung eines Freiheitsgrades. Während der Vorstellung des Protokolls PULSAR geht die vorliegende Arbeit detailliert auf Entwurfsentscheidungen bezüglich der Erfassung und Messung der Kanalauslastung, der Anpassung der Senderate und dem Austausch von Kontrollinformationen ein, die das Ergebnis der Regelung in Hinblick auf Fairness, Konvergenz und Stabilität beeinflussen. Ein Vergleich mit anderen dem Stand der Technik entsprechenden Ansätzen zeigt, dass „Details“ im Protokollentwurf einen signifikanten Einfluss auf das räumliche und zeitliche Konvergenzverhalten des Protokolls haben können.

Abstract

In Vehicle Safety Communications (VSC) based on the wireless LAN standard IEEE 802.11p, vehicles are envisioned to periodically exchange short messages indicating their current position, speed and heading in order to establish a mutual awareness of each other's presence. Based on this knowledge, a vehicle can warn its driver of potentially dangerous situations, e.g., an unsafe lane change maneuver. However, if the vehicle density is high and the transmission of messages is not regulated, the communication channel can become congested, impairing reception performance and eventually the performance of driver assistance safety applications. In this thesis, we systematically develop a design methodology for congestion control in VSC and present a resulting protocol named PULSAR. In contrast to the majority of the related work which typically focused on controlling channel load only, we thereby integrate a concept which allows the adaptation to operate within the limits defined by safety applications. By adapting a vehicle's *relative* transmission rate instead of the absolute value, our approach allows to establish a fairness based on safety benefit among vehicles rather than a fairness based on the share of channel resources as it is typically applied in communication networks. In addition, while the related work typically did not address what the optimal outcome for congestion control in VSC would look like and why their approach would converge to this result, we provide a detailed analysis for the first question and a control strategy jointly adapting message generation rate and transmit power for the second question. In contrast, existing approaches typically focused on the adaptation of one control dimension only. In the description of PULSAR, we discuss several intricate protocol design decisions regarding channel load assessment, transmission rate adaptation and information sharing which influence the fairness, convergence and stability of the adaptation. A comparison with state-of-the-art approaches shows that "details" in protocol design matter regarding the spatio-temporal outcome of the adaptation.

Acknowledgments

The foundation of the work presented in this thesis was laid during my stay with Mercedes-Benz Research & Development North America, Inc. (MBRDNA) in 2010. I would like to express my sincere gratitude to my former colleagues from MBRDNA, in particular Daniel Jiang, Qi Chen and Luca Delgrossi, without whom this work would not have been possible. I'm grateful to have had the opportunity to work with them and I would like to thank them for a successful and highly enjoyable collaboration over many years.

At the same time, I would like to express my sincere gratitude to my advisor Prof. Dr. Hannes Hartenstein for his continuous support, encouragement and guidance over the years. His example and his leadership helped me to grow both as a scientist and as a person. I would further like to thank Prof. Dr. Marco Gruteser for his valuable comments during my visit in New Jersey and for taking the time to review my thesis. Our discussions have helped me to see my dissertation topic from a broader perspective.

In the recent years, the DSN research group has become like a second home to me. I would like to thank all of my present and former colleagues for their support and comradeship as well as for many insightful discussions which helped me to improve the quality of my work. I would also like to thank all the students who worked with me over the course of the years, from whom I was able to learn as much as they hopefully were able to learn from me.

This work has further been influenced substantially by my colleagues from the ETSI Specialist Task Force 447, Katrin Sjöberg, Dieter Smely, Jérôme Härri, Jan de Jongh and Andreas Festag. I would like to thank them for many fruitful discussions which have broadened my horizon.

Last but not least, I would like to thank my friends and family for their love and support during the course of my PhD studies, especially my boyfriend and partner Erik Haaf.

Contents

Zusammenfassung iii

Abstract v

Acknowledgments vii

List of Tables xiii

List of Figures xv

List of Abbreviations xix

List of Symbols xxiii

1 Introduction 1

2 Vehicle Safety Communications 11
2.1 Technology and Standards . . . 12
2.1.1 Frequency Band . . . 12
2.1.2 IEEE 802.11p . . . 13
2.1.3 Protocol Stacks in Europe and the United States . . . 18
2.1.4 Global Positioning System (GPS) . . . 24
2.2 Safety Applications and Their Requirements . . . 25
2.2.1 Focus in the United States . . . 25
2.2.2 Focus in Europe . . . 27
2.2.3 Cooperative Awareness . . . 30
2.3 Scalability and Channel Congestion . . . 32
2.3.1 Degrees of Freedom . . . 34
2.3.2 Transmission Range, Carrier Sense Range and Interference Range . . . 34
2.3.3 Metrics for Channel Load . . . 40
2.4 Summary of Characteristics and Challenges . . . 43

3 Congestion Control in Wired, Wireless and Vehicular Networks 45
3.1 Basic Concepts . . . 46

3.1.1 Congestion Avoidance and Control . . . 46
3.1.2 Fairness and Efficiency . . . 48
3.1.3 Binary Adaptation Schemes . . . 49
3.1.4 Taxonomy According to Yang and Reddy . . . 51
3.2 Transmission Control Protocol (TCP) . . . 53
3.2.1 Congestion Detection . . . 53
3.2.2 Rate Adaptation . . . 54
3.2.3 TCP over Wireless Links . . . 56
3.3 Wireless Networks . . . 57
3.3.1 Cellular Networks . . . 57
3.3.2 IEEE 802.11 Wireless LAN . . . 57
3.3.3 Wireless Sensor Networks . . . 58
3.3.4 Mobile Ad-Hoc Networks (MANETs) . . . 59
3.4 Vehicular Networks . . . 59
3.4.1 Packet Size Control . . . 60
3.4.2 Data Rate Control . . . 62
3.4.3 Contention Window Control . . . 64
3.4.4 Carrier Sense Threshold Control . . . 68
3.4.5 Transmit Power Control (TPC) . . . 73
3.4.6 Transmission Rate Control (TRC) . . . 76
3.4.7 Joint Control of Multiple Dimensions . . . 79
3.4.8 Summary and Conclusions . . . 80

4 Analysis of Optimization Problem and Solution Space 83
4.1 Optimization Criterion . . . 84
4.1.1 Spatial Aspect: Target Distance . . . 84
4.1.2 Temporal Aspect: Inter-Reception Time . . . 86
4.2 Optimized Parameter Configurations . . . 88
4.2.1 Methodology . . . 88
4.2.2 Optimization for One Group of Vehicles . . . 89
4.2.3 Optimization for a Certain Channel Load . . . 91
4.2.4 Optimization for Two Intermixed Groups of Vehicles . . . 93
4.3 Reception Performance Without Congestion Control . . . 96

5 Design of Congestion Control for VSC 103
5.1 Control Strategies . . . 104
5.1.1 Transmit Power Control (TPC) . . . 104
5.1.2 Transmission Rate Control (TRC) . . . 105
5.1.3 Joint Control of Transmit Power and Transmission Rate . . . 106
5.1.4 Discussion . . . 109
5.2 System Model and Convergence Target . . . 111
5.2.1 Rate Control Model According to Kenney et al. . . . 112
5.2.2 Relationship of Communication Density and CBR . . . 113
5.3 Aspects of Fairness . . . 119
5.3.1 Local Fairness . . . 119

5.3.2 Global Fairness . . . 121
5.3.3 Participation Fairness . . . 122
5.4 Adaptation of the Relative Transmission Rate . . . 123
5.4.1 General Concept . . . 124
5.4.2 Concrete Example Based on Inter-Reception Distance . . . 126
5.4.3 Implementation using AIMD . . . 130
5.5 Summary of Design Principles . . . 131

6 Protocol Design and Description **135**
6.1 General Structure . . . 135
6.2 Local Channel Load Assessment . . . 136
6.2.1 CBR Filtering Techniques . . . 137
6.2.2 Impact of the Nyquist-Shannon Sampling Theorem . . . 140
6.3 Transmission Rate Adaptation . . . 141
6.3.1 Message Rescheduling . . . 142
6.3.2 Synchronization . . . 144
6.3.3 Target Rate Mechanism . . . 149
6.3.4 Adaptation of the Relative Transmission Rate . . . 151
6.4 Exchange and Aggregation of Channel Information . . . 152
6.4.1 Motivation . . . 153
6.4.2 One-Hop Piggybacking . . . 154
6.4.3 Piggybacking Efficiency . . . 158
6.4.4 Two-Hop Piggybacking . . . 160
6.5 Summary . . . 164

7 Evaluation and Discussion **167**
7.1 Responsiveness and Smoothness . . . 167
7.1.1 AIMD Parameter Choice . . . 168
7.1.2 Other Influencing Factors . . . 170
7.1.3 Resulting Trade-Offs . . . 174
7.2 Fairness and Stability . . . 175
7.2.1 Long-Term Convergence in a Static Scenario . . . 175
7.2.2 Performance in a Dynamic Highway Scenario . . . 184
7.2.3 Dynamic Change of Minimum and Maximum Rate . . . 191
7.3 Impact on an Exemplary Safety Application . . . 191
7.3.1 Considered Safety Application . . . 192
7.3.2 Scenario and Setup . . . 194
7.3.3 Convergence and Stability of PULSAR . . . 195
7.3.4 Impact on the Observed Inter-Reception Distance . . . 200
7.4 Comparison With Other Approaches . . . 204
7.4.1 Khorakhun et al. . . . 205
7.4.2 SOURC . . . 206
7.4.3 LIMERIC . . . 211
7.5 Discussion . . . 216
7.5.1 Piggybacking in Urban Scenarios . . . 216

7.5.2 Enhanced Distributed Channel Access (EDCA) 216

8 Conclusions and Outlook **219**

A Simulation Environment and Configuration **223**
A.1 Propagation model 223
A.2 Network Simulator 224

Bibliography **227**

List of Tables

2.1 Available data rates in a 10 MHz channel using IEEE 802.11p and their required SINR values for reception . . . 14
2.2 Default EDCA parameters of IEEE 802.11p . . . 19
2.3 High-priority safety applications and their communication requirements according to the VSC project . . . 26
2.4 V2V safety applications evaluated by the VSC-A project . . . 28
2.5 Communication-relevant minimum performance requirements of V2V safety applications as defined by the VSC-A project . . . 29
2.6 CAM-based use cases of the ETSI Basic Set of Applications (BSA) . . . 30
2.7 Awareness metrics related to Packet Inter-Reception Time (IRT) . . . 33
2.8 Degrees of freedom influencing the scalability of VSC and their allowed value ranges . . . 35
2.9 Exemplary Communication Density (CD) and Beaconing Load (BL) values . . . 42
3.1 Congestion and awareness control . . . 61
3.2 Classification of selected related approaches . . . 81
4.1 Simulation parameters and their evaluated range . . . 90
5.1 Examples of fairness concepts in the related work . . . 120
6.1 PULSAR's header format . . . 165
6.2 Realization of design principles in PULSAR . . . 165
6.3 PULSAR as a congestion avoidance scheme according to Jain et al. . . . 165
7.1 Factors influencing PULSAR's responsiveness and smoothness . . . 176
7.2 k^{th} percentiles P_k of Inter-Reception Distance (IRD) observed in the traffic jam scenario . . . 205
A.1 Default simulation parameters . . . 225

List of Figures

1.1 Crash statistics of Germany (1953-2011) . . . 2
1.2 Overview of the structure and contents of this thesis . . . 10
2.1 Spectrum and channel allocation for C-ITS in Europe and the United States . . . 13
2.2 Carrier Sense Multiple Access (CSMA) example . . . 16
2.3 Hidden terminal and exposed terminal in unicast communication . . . 18
2.4 Overview of protocol stacks and corresponding standards for VSC in Europe and the United States . . . 20
2.5 Structure and size range of a Cooperative Awareness Message (CAM) and Basic Safety Message (BSM) (estimation) . . . 23
2.6 Target classification (simplified) and corresponding safety applications of the VSC-A project . . . 27
2.7 ETSI safety applications and corresponding time horizons . . . 32
2.8 Trade-off between communication range and the maximum number of packets/s for different data rates . . . 37
2.9 Relationship of data rate, interference and communication range . . . 38
2.10 Interference range according to the SIR and SINR models . . . 39
2.11 Exemplary deterministic carrier sense ranges for different carrier sense thresholds . . . 40
2.12 Schematic illustration of Channel Busy Ratio (CBR) calculation with channel probing . . . 44
3.1 Congestion avoidance and control . . . 47
3.2 Example scenario to illustrate different fairness concepts . . . 50
3.3 Vector representation of linear control algorithms according to Chiu et al. . . . 52
3.4 Smoothness and responsiveness according to Chiu et al. . . . 53
3.5 Taxonomy of congestion control algorithms according to Yang et al. 54
3.6 Example of congestion window adaptation in TCP Tahoe . . . 56
3.7 Packet collision probability for fixed contention window size . . . 66
3.8 Relationship between sender-receiver distance, carrier sense range, interference range, hidden node area and exposed node area . . . 69
3.9 Schematic illustration of the impact of the carrier sense range on hidden terminals in a highway crossing scenario . . . 72

4.1 Schematic illustration of a German freeway situation with resulting target distances . . . 86
4.2 Examples of the observed average IRT values for a certain vehicle density and target distance . . . 91
4.3 Parameter combinations optimizing average IRT . . . 92
4.4 Sensitivity of the parameter combinations optimizing average IRT . . . 93
4.5 Parameter combinations optimizing the 99th percentile of IRT . . . 94
4.6 Packet Delivery Ratio (PDR) corresponding to the parameter configurations optimizing average IRT . . . 95
4.7 Channel Busy Ratio (CBR) corresponding to the parameter configurations optimizing average IRT . . . 96
4.8 Parameter combinations optimizing average IRT at different CBR levels 97
4.9 Average IRT results for two groups of vehicles optimizing for different target distances . . . 98
4.10 Parameter combinations resulting in Pareto optimal points . . . 99
4.11 Ratio of Pareto optimal setups using the transmit power identified as optimal in the homogeneous case . . . 100
4.12 Average IRT for different fixed transmission parameter settings compared to minimum values . . . 101
4.13 Difference to minimum average IRT for different fixed transmission parameter settings . . . 102
5.1 Difference in average IRT between the minimum value for a fixed transmission rate and the global minimum value for each distance . . . 105
5.2 Difference in average IRT between the minimum value for a fixed transmit power and the global minimum value for each distance . . . 107
5.3 Difference in average IRT between the suggested joint control strategy and the global minimum value for each distance . . . 109
5.4 Observed CBR and goodput in a simulation where 100 nodes share the same location . . . 113
5.5 Channel performance metrics observed in the optimization study with uniform node configuration . . . 115
5.6 Relationship between Communication Density (CD) and different performance metrics . . . 116
5.7 Relationship between Channel Busy Ratio (CBR) and different performance metrics . . . 118
5.8 Schematic illustration of local and global fairness from a rate control perspective (with uniform maximum and minimum rate) . . . 121
5.9 Theoretical carrier sense range and participation range for a transmit power of 20 dBm in our setup . . . 123
5.10 Illustration of the concept of relative transmission rate adaptation . . . 125
5.11 Inter-Reception Distance (IRD) functions based on the 99.9th percentile of IRT . . . 129
5.12 Schematic illustration of the relationship between maximum Inter-Reception Distance (IRD) and relative transmission rate . . . 130

5.13 Relationship of transmission rate, transmission interval and observed CBR in a scenario where 100 nodes share the same location 131
5.14 Examples of relative AIMD adaptations in the vector representation of Jain et al. 132
5.15 Convergence of relative AIMD over time (continued example) . . . 133
6.1 Overview and building blocks of the PULSAR protocol 137
6.2 Impact of different filtering techniques on the observed CBR 138
6.3 Example of an AIMD adaptation based on CBR filtering with and without the fulfillment of the sampling theorem 141
6.4 Impact of the rescheduling mechanism when decreasing transmission rate 143
6.5 Impact of the rescheduling mechanism when increasing transmission rate 144
6.6 Impact of the rescheduling mechanism on the observed CBR 145
6.7 Impact of the rescheduling mechanism on AIMD-based rate adaptation 146
6.8 Example of an asynchronous AIMD adaptation with three nodes . . 147
6.9 Asynchronous AIMD-based rate adaptation in a scenario excluding radio-wave propagation effects 148
6.10 Impact of the target rate mechanism on the convergence of AIMD . 150
6.11 Exemplary adaptation of the relative transmission rate using AIMD 153
6.12 AIMD-based adaptation on a linear road (local information, without target rate) 155
6.13 AIMD-based adaptation on a linear road (local information, with target rate) 155
6.14 AIMD-based adaptation on a linear road (one-hop information, 1-norm, with target rate) 157
6.15 AIMD-based adaptation on a linear road (one-hop information, maximum norm, with target rate) 157
6.16 Convergence of AIMD with target rate using one-hop piggybacking in the crossing scenario 159
6.17 Piggybacking efficiency (probability of receiving one out of n messages) 160
6.18 Illustration of two-hop piggybacking and CBR aggregation 162
6.19 Convergence of AIMD with target rate using two-hop piggybacking in the crossing scenario 163
7.1 Convergence of AIMD for two different parameter choices 169
7.2 Impact of the number of nodes on responsiveness and smoothness . 171
7.3 Impact of the adaptation interval on responsiveness and smoothness 172
7.4 Impact of the target rate mechanisms on responsiveness and smoothness 173
7.5 Impact of two-hop piggybacking on responsiveness and smoothness 174
7.6 Comparison of the convergence of PULSAR and water filling in the static crossing scenario 177
7.7 PULSAR's rate adaptation over time in the crossing scenario 179
7.8 Stabilizing effect of target rate mechanism 181

7.9 Impact of the adaptation interval on PULSAR's convergence in the crossing scenario . . . 182
7.10 Probability of receiving a piggybacked bit within a certain time limit 183
7.11 Impact of the participation distance on PULSAR's convergence in the crossing scenario . . . 185
7.12 Impact of AIMD parameters on PULSAR's convergence in the crossing scenario . . . 186
7.13 Snapshots of vehicle positions in scenario with two approaching groups of vehicles . . . 187
7.14 Snapshots of PULSAR's transmission rate and observed channel load in a scenario with two passing groups of vehicles . . . 189
7.15 Impact of the target rate mechanism on PULSAR's convergence in a scenario where two groups of vehicles pass each other . . . 190
7.16 Adaptation by an exemplary vehicle in a scenario where two groups of vehicles pass each other . . . 192
7.17 Stress test of dynamic change in minimum and maximum rate . . . 193
7.18 Vehicle speed distribution at simulation time 50 s in the traffic jam scenario . . . 195
7.19 Representative snapshot of PULSAR's performance in the traffic jam scenario . . . 198
7.20 Transmission rate and rate boundaries of an exemplary vehicle applying relative rate adaptation in the traffic jam scenario . . . 200
7.21 Relative transmission rate and resulting CBR at simulation time 50 s in the traffic jam scenario . . . 201
7.22 Observed CBR over the course of a simulation using dynamic rate boundaries and transmit power values in the traffic jam scenario . . . 202
7.23 Distribution of observed IRDs in the traffic jam scenario . . . 203
7.24 Convergence of the approach by Khorakhun et al. in the linear road scenario . . . 206
7.25 Convergence of the approach by Khorakhun et al. in the static crossing scenario . . . 207
7.26 Convergence of SOURC in the static crossing scenario . . . 209
7.27 Adaptation by an exemplary node running SOURC in the scenario where two groups of vehicles pass each other . . . 210
7.28 Convergence of SOURC in scenario where 200 nodes share the same location . . . 211
7.29 Convergence of LIMERIC in a scenario where 200 nodes share the same location . . . 213
7.30 Convergence of LIMERIC in the static crossing scenario . . . 214
7.31 Adaptation by an exemplary node running LIMERIC in the scenario where two groups of vehicles pass each other . . . 215
7.32 Schematic illustration of two-hop piggybacking in a crossing scenario 217
A.1 RSS per distance according to different propagation models . . . 224

List of Abbreviations

ABS	Antilock Braking System
AC	Access Category
ACK	Acknowledgment
AIAD	Additive Increase Additive Decrease
AIFS	Arbitration Interframe Space
AIFSN	AIFS Number
AIMD	Additive Increase Multiplicative Decrease
AP	Access Point
APP	Application layer
ARF	Auto Rate Fallback
BL	Beaconing Load
BS	Base Station
BSA	Basic Set of Applications
BSM	Basic Safety Message
BSS	Basic Service Set
BSW	Blind Spot Warning
BTP	Basic Transport Protocol
C-ITS	Cooperative Intelligent Transport Systems
C2C-CC	Car-to-Car Communications Consortium
CA	Cooperative Awareness
CAM	Cooperative Awareness Message
CAMP	Crash Avoidance Metrics Partnership
CAT	Channel Access Time
CBR	Channel Busy Ratio
CBT	Channel Busy Time
CCA	Clear Channel Assessment
CD	Communication Density
CDF	Cumulative Density Function
CEN	European Committee for Standardization
CLW	Control Loss Warning
CSMA	Carrier Sense Multiple Access
CSTh	Carrier Sense Threshold
CTS	Clear To Send

DCC	Decentralized Congestion Control
DCF	Distributed Coordination Function
DEN	Decentralized Environment Notification
DENM	Decentralized Environment Notification Message
DGPS	Differential GPS
DIFS	DCF Interframe Spacing
DNPW	Do Not Pass Warning
DSC	DCC Sensitivity Control
DSRC	Dedicated Short Range Communications
ECDSA	Elliptic Curve Digital Signature Algorithm
ECN	Explicit Congestion Notification
EDCA	Enhanced Distributed Channel Access
EEBL	Emergency Electronic Brake Lights
EMDV	Emergency Message Dissemination for Vehicular environments
ETSI	European Telecommunications Standards Institute
EWMA	Exponentially Weighted Moving Average
FCC	U.S. Federal Communication Commission
FCW	Forward Collision Warning
FEC	Forward Error Correction
FOT	Field Operational Test
GLOSA	Green Light Optimized Speed Advisory
GNSS	Global Navigation Satellite Systems
GPS	Global Positioning System
HV	Host Vehicle
I2V	Infrastructure to Vehicle
ICRW	Intersection Collision Risk Warning
IEEE	Institute for Electrical and Electronics Engineers
IMA	Intersection Movement Assist
IP	Internet Protocol
IRD	Inter-Reception Distance
IRT	Packet Inter-Reception Time
ISO	International Organization for Standardization
ITS	Intelligent Transportation System
LAN	Local Area Network
LCRW	Longitudinal Collision Risk Warning
LCW	Lane Change Warning
LDM	Local Dynamic Map

LIMERIC	LInear MEssage Rate Integrated Control
LLC	Logical Link Control
LOS	Line of Sight
LTE	Long Term Evolution
MAC	Medium Access Control
MANET	Mobile Ad-Hoc Network
MIAD	Multiplicative Increase Additive Decrease
MIMD	Multiplicative Increase Multiplicative Decrease
MN	Mobile Node
MSS	Maximum Segment Size
NLOS	Non Line of sight
OCB	Outside the Context of a BSS
OEM	Original Equipment Manufacturer
OSI	Open Systems Interconnection
PDR	Packet Delivery Ratio
PER	Packet Error Rate
PHY	Physical layer
PMF	Probability Mass Function
PULSAR	Periodically Updated Load Sensitive Adaptive Rate control
QoS	Quality of Service
RHS	Road Hazard Signaling
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RTO	Retransmission Timeout
RTS	Ready To Send
RTT	Round Trip Time
RV	Remote Vehicle
SACK	Selective Acknowledgment
SAE	Society for Automotive Engineers
SIFS	Short Interframe Spacing
SINR	Signal to Interference and Noise Ratio
SIR	Signal to Interference Ratio
SMA	Simple Moving Average
SNAP	SubNetwork Access Protocol
SOURC	Self-Organized Update Rate Control
SRTT	Smoothed RTT

TC	Technical Committee
TCP	Transmission Control Protocol
TDC	Transmit Datarate Control
TPC	Transmit Power Control
TRC	Transmission Rate Control
TTC	Time To Collision
TXOP	Transmission Opportunity
UDP	User Datagram Protocol
USDOT	United States Department of Transportation
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
V2X	Vehicle to X
VANET	Vehicular Ad-Hoc Network
VSC	Vehicle Safety Communications
VSC-A	Vehicle Safety Communications - Applications
WAVE	Wireless Access in Vehicular Environments
WSA	WAVE Service Announcement
WSM	WAVE Short Message
WSMP	WAVE Short Message Protocol
WSN	Wireless Sensor Network

List of Symbols

BL	Beaconing Load
CD	Communication density
L	Reference loss
M	Message size
U	Channel Busy Ratio (CBR) (channel utilization)
γ	Path loss exponent
ρ	Vehicle density
a	Acceleration
v	Speed
CW_{min}	Minimum contention window
N_o	Noise floor
P_{RxTh}	RSS required to receive a packet in an interference-free environment
P_{Rx}	Received Signal Strength (RSS)
P_{Tx}	Transmit power
P_{csN}	Difference between carrier sense threshold and noise floor
P_{cs}	Carrier sense threshold
S_o	Signal to Interference Ratio (SIR) required for reception
S_N	Signal to Interference and Noise Ratio (SINR) required for reception
U_t	Target CBR (channel utilization)
α_I	Additive increase parameter (PULSAR)
$\alpha_{I,rel}$	Relative additive increase parameter (PULSAR)
β_D	Multiplicative decrease parameter (PULSAR)
$\beta_{D,rel}$	Relative multiplicative decrease parameter (PULSAR)
d_{Rx}	Inter-reception distance
d_{Tx}	Transmission range
d_{brk}	Braking distance
d_{cs}	Carrier sense range
d_{dec}	Deceleration distance
d_{hop}	Hop distance (PULSAR)
d_{int}	Interference range
d_{min}	Minimum sender-interferer distance which does not diminish communication range
d_{neigh}	Neighbor distance (SOURC)
d_{part}	Participation distance (PULSAR)
d_{rel}	Relative distance between two vehicles
d_{sr}	Sender-receiver separation distance

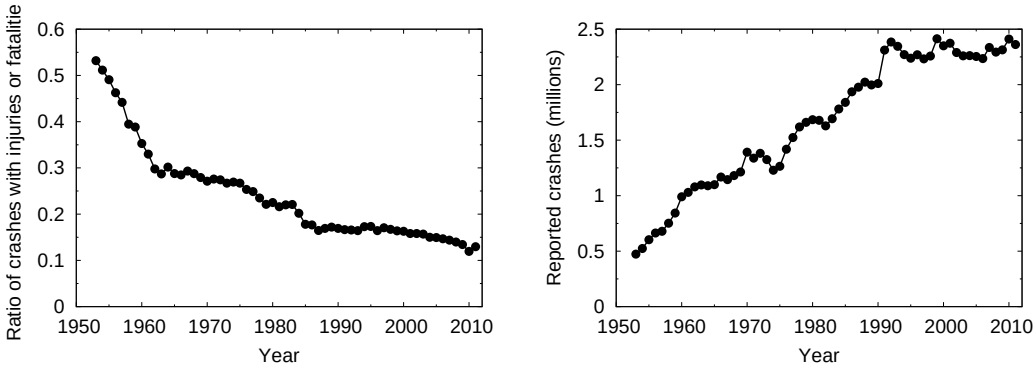
d_t	Target distance
p_{Rx}	Probability of packet reception
r_{max}	Maximum transmission rate
r_{min}	Minimum transmission rate
r_{rel}	Relative transmission rate
r_t	Target rate (in PULSAR)
r	Transmission (message generation) rate
s_e	Smoothness (oscillation size)
t_e	Responsiveness (convergence time)
t_{Rx}	Packet Inter-Reception Time (IRT)
t_{Tx}	Transmission interval
t_{adp}	Adaptation interval (PULSAR)
t_{max}	Maximum age of the CBR values considered in the aggregation (PULSAR)
t_{mon}	Channel monitoring time (PULSAR)
t_{rct}	Reaction time of the driver
t_{sys}	System latency
v_{max}	Maximum speed
v_{rel}	Relative speed

1 Introduction

Within the last decades, research efforts in academia and industry have resulted in a constant improvement of the passive and active safety of vehicles. While passive safety systems such as seat belts and airbags are designed to reduce the harm done to passengers in a crash, active safety systems like the Antilock Braking System (ABS) aim at preventing crashes in the first place and at reducing the collision impact when a crash is unavoidable [1]. In Germany, two trends can be observed in crash statistics, as illustrated by Figure 1.1 on the following page. On the one hand, the ratio of reported crashes with injuries or fatalities to the total number of crashes decreased from 53% to 13% between 1953 and 2011 (Figure 1.1a). On the other hand, the total number of reported crashes increased by 400% during the same time period, stagnating since the 1990s (Figure 1.1b). In other words, while crashes tend to become less severe, their total number is still high. From a research point of view, the development of advanced driver assistance systems to improve active safety thus holds great potential in order to further reduce the number of crashes and thus of traffic casualties.

Driver assistance systems based on radar or cameras can already be found on the market, offering for example emergency braking and lane keeping assistance[1]. Typically, such systems can be found in premium class vehicles, being a competitive factor among Original Equipment Manufacturers (OEMs) [1]. While this may change over the years as the technology becomes less expensive, a major drawback of such systems in general is that they are based on the local view of a vehicle only. In other words, it is difficult to detect hidden hazards such as a broken down car behind a curve of the road or a car approaching an intersection from another direction. Thus, the

[1] E.g. *http://www.bosch-automotivetechnology.com*



(a) Ratio of crashes with injuries or fatalities

(b) Total number of reported crashes

Figure 1.1: Crash statistics of Germany (1953-2011)[2]

next generation of driver assistance systems is envisioned to introduce *cooperation* between vehicles by means of communication technology.

In October 2012, twelve European vehicle manufacturers signed a *Memorandum of Understanding* to "bring Cooperative Intelligent Transport Systems and Services (C-ITS) onto European roads", starting in 2015 [2]. The new communication system is envisioned to facilitate an information exchange from vehicle to vehicle (V2V) and/or from vehicle to infrastructure (V2I), summarized as vehicle-to-X (V2X) communication. In February 2014, the United States Department of Transportation (USDOT) announced that "it will begin taking steps to enable vehicle-to-vehicle (V2V) communication technology for light vehicles". A key enabler in the realization of C-ITS is the widespread availability of Global Navigation Satellite Systems (GNSS) such as the Global Positioning System (GPS). While the idea of connecting cars fascinated people already in the 1930s and concrete ideas for V2X communication applications existed in the 1980s, a "game changer" occurred in 1999, when the U.S. Federal Communication Commission (FCC) allocated 75 MHz of bandwidth in the 5.9 GHz band for C-ITS [3]. Europe followed its course, reserving 30 MHz of spectrum for safety communications and 20 MHz for non-safety usage, also in the 5.9 GHz band [4][3].

V2V communication facilitates a variety of new safety applications such as Emergency Electronic Brake Lights (EEBL), which warns the driver if a vehicle in front of him is braking hard, even one to which he has no direct line of sight, and Intersection Movement Assist (IMA), which warns if there is a high probability of colliding with another vehicle coming from the side [6]. Further, the technology has the potential to substitute (or complement) cameras and radars in existing driver assistance systems

[2] Source: Statistisches Bundesamt (Federal Statistical Office); up to 1991: Bundesrepublik Deutschland (West Germany)

[3] Japan has allocated spectrum for V2V safety applications as well. However, it is out of the scope of this thesis to determine if the introduced concepts are compatible with the foreseen frequency band of 700 MHz and the deviating protocol stack [5].

such as Blind Spot Warning (BSW), offering opportunities to reduce manufacturing costs and make these systems available sooner to the general public [7].

In addition to benefits for traffic safety, V2X communication is envisioned to improve traffic efficiency and driving convenience. For example, the foreseen application Green Light Optimized Speed Advisory (GLOSA) offers speed adaptation advice to the driver approaching a signalized intersection based on phase shift information received from the traffic light [8]. By reducing stopping and accelerating maneuvers, GLOSA is expected to not only make the driving experience more pleasurable but to facilitate a significant reduction of the fuel consumption and emissions of conventional vehicles as well as of the energy consumption of electric vehicles [9][10]. Other non-safety applications relate for example to the discovery of free parking spaces and advanced navigation based on the current traffic situation [8][11].

From a technical point of view, two main alternatives are currently under discussion for the realization of C-ITS. One candidate is IEEE 802.11p, a wireless Local Area Network (LAN) variant specifically adapted to the challenges of vehicular communications in the 5.9 GHz band [12]. Unlike conventional wireless LANs which are typically coordinated centrally by an access point, vehicles are envisioned communicate directly with each other, forming a Vehicular Ad-Hoc Network (VANET)[3]. Next to its low-cost availability, one of the main advantages of this technology is that it enables information exchange with low latency, i.e., less than 100 ms, which is particularly important for safety communications [13].

The second technological alternative is cellular networks, in particular Long Term Evolution (LTE). In contrast to its 2G and 3G predecessors, 4G LTE may technically be able to meet the stringent latency requirements of safety applications [1][14]. However, it is not clear if cellular networks could accommodate the significant amount of data generated by vehicles in addition to the load generated by human users [1][14]. Another open issue is how a potentially mandatory safety system could be financed if it requires a monthly fee to be paid. While cellular networks can be expected to play a role in the future of C-ITS, especially for non-safety applications like GLOSA [1], the focus of this thesis is on IEEE 802.11p based Vehicle Safety Communications (VSC) and the challenges of its realization.

Problem statement

The predominant message type [1][15] for VSC is a periodically transmitted single-hop broadcast message, referred to as a Cooperative Awareness Message (CAM) [16] in Europe and as a Basic Safety Message (BSM) [17] in the United States. A CAM/BSM[4] contains the sending vehicle's current position, speed, heading and other

[4] CAMs and BSMs are also referred to as *beacons* in the context of VSC. This term should not be confused with beacons used in IEEE 802.11 to advertise the presence of an access point.

status information, determined based a GNSS device as well as local sensors. By default, CAMs/BSMs are generated at a rate of ten messages per second, i.e., 10 Hz [8][13]. Based on the reception of CAMs/BSMs, vehicles can establish a *cooperative awareness*, i.e., they can track the position and heading of their neighbors and infer the likelihood of a crash [18][19]. Cooperative awareness is the basis for many safety applications, e.g., IMA. With increasing vehicle density, however, the number of generated beacon messages can exceed the capacity of the channel, resulting in packet collisions and poor reception performance [19][20]. Eventually, the functionality of safety applications can be impaired [21]. Thus, a congestion control protocol is required to adjust the load injected into the channel.

In computer science, congestion control has played a major role since the 1980s when the early Internet collapsed due to increasing network load [22]. In the aftermath, the Transmission Control Protocol (TCP) was augmented with congestion control and has since constantly been analyzed and improved. However, classical end-to-end oriented congestion control approaches like TCP are not directly applicable to VSC due to its special characteristics. The broadcast nature of CAMs/BSMs requires a distributed selection of transmission parameters in real time, without message acknowledgments and in a constantly changing environment and vehicle density. In addition, congestion control needs to adapt transmission parameters in a way which does not conflict with safety applications' requirements in terms of message reception. While a number of congestion control approaches for VSC have been suggested, they typically either do not address their impact on safety applications' performance, e.g., [20], or they do not address a fairness concept, e.g., [23]. In addition, no consensus can be inferred from the related work as to how and why the available degrees of freedom should be adapted.

To close this gap, this thesis addresses the following research questions:

1. Which *transmission parameter combinations* should congestion control converge to in order not to conflict with safety applications' requirements?
2. Which *control strategy* is able to find or approximate the optimal transmission parameters for a certain channel load limit?
3. How can this control strategy be fleshed out with concrete *protocol mechanisms* in order to achieve fairness and stability?
4. How can congestion control be designed to provide a similar *safety benefit* to vehicles rather than a similar share of channel resources?

Challenges

As mentioned above, the special characteristics of VSC result in a number of challenges for the design of congestion control, in particular:

Lack of a clearly defined optimization criterion. In contrast to other communication networks, the objective of VSC is to maximize cooperative awareness rather than network throughput. However, there is no universal definition of "awareness", since safety applications' requirements depend not only on the application itself, but also on the driving situation. In addition, different applications may have conflicting requirements. Due to this complexity, a number of different optimization criteria for congestion control can be found in the related work. In addition, many approaches focused purely on channel load and did not address their impact on awareness.

Many degrees of freedom with intricate interplays. Channel load in VSC can be influenced by adapting a number of parameters, in particular data rate, transmit power, carrier sense threshold, contention window size, message size and message generation rate, cf. Section 2.3.1 on page 34. Existing congestion control approaches for VSC adapt one or more of the aforementioned parameters, the majority focusing on transmit power or message generation rate. However, even when limiting the choice to the latter two parameters, there is no consensus in the related work as to how and why or why not to adapt each degree of freedom.

Achieving stability and fairness in a distributed and rapidly changing environment. The high mobility of vehicles results in many different challenges for congestion control. First, a centralized assignment of transmission parameters is impracticable and a decentralized algorithm is necessary. Second, the number of vehicles sharing the channel at a certain location is constantly changing, requiring a fast converging adaptation. Third, fairness plays a major role, since vehicles not able to communicate their presence sufficiently may become a safety risk to others. Finally, mobility affects the characteristics of the wireless channel due to changes in the environment, e.g., a truck obstructing line of sight, and thus makes the outcome of the chosen communication parameters difficult to predict.

Main contributions

Following the logic of the underlying research questions and challenges, the main contributions of this dissertation are:

Definition of an optimization criterion. Following the reasoning that in theory, infinitesimally small time intervals between status updates from neighboring vehicles would be ideal for the functionality of safety applications [24], we define the minimization of Packet Inter-Reception Time (IRT) i.e., the time interval between two successive message[5] receptions for a sender-receiver pair [7], at a situation-dependent *target distance* as the optimization criterion. We provide a detailed analysis of the

[5] While the terms *message* and *packet* refer to different layers in the ISO/OSI protocol stack, they are often used interchangeably in the context of VSC, since CAMs/BSMs are expected to be less than 800 Bytes [17][16] and thus a segmentation is not necessary.

distribution of IRT in a controlled environment and discuss the advantages and disadvantages of using different statistical properties of IRT for optimization. By taking on a receiver-centric view in the optimization process, the approach introduced in this thesis differs from the majority of related congestion control approaches which typically focused only on channel input, i.e., the sending side.

Identification of optimal parameter combinations and their characteristics. Using the defined optimization criterion, we evaluate a large parameter space of transmit power and transmission rate[6] in a scenario with a uniform configuration of nodes by means of simulations. The main finding is that the transmission power which minimizes average (and percentile of) IRT at a particular distance is very similar for different vehicle densities. Since the constant change in vehicle density is a major challenge for congestion control, this result helps considerably to simplify protocol design. We further show that the same characteristics can be observed if a channel load limit is applied. Finally, we demonstrate that the obtained optimal transmission power values can also be used to approximate the Pareto border for two groups of vehicles optimizing for different target distances at the same time.

Analysis of the potential of common congestion control strategies to approximate the identified optimal configurations. Many existing congestion control approaches for vehicular networks are based on Transmit Power Control (TPC) or Transmission Rate Control (TRC), without a discernible dominance of either one. In this thesis, we analyze the potential of each control strategy to approximate the identified optimal parameter combinations. Quantifying the resulting difference, we show that both strategies leave room for improvement. Between the two, TRC produces the more predictable result with respect to the outcome at the desired target distance.

Introduction of a joint power/rate congestion control strategy. Based on the characteristics of the identified optimal parameter configurations and the insights gained in the analysis of TPC and TRC, we suggest a joint control strategy which uses TRC as the primary congestion control mechanism. Additionally, TPC is applied to reduce unnecessary interference if allowed by the driving context and as a fallback mechanism. We show that the suggested strategy efficiently approximates the identified optimal configurations.

Identification and analysis of the requirements for congestion control in VSC. We distill the state of the art into five design principles for congestion control. Next to decentralization and deference to safety applications, we distinguish three aspects of fairness, i.e., *local fairness*, *global fairness* and *participation fairness*. While related approaches typically addressed only a subset of these principles, we discuss in detail why each of them is required.

Introduction of a concept to integrate safety applications' requirements into congestion control. While previous approaches typically focused on the control of

[6] In this thesis, the term *transmission rate* refers to the generation rate of CAMs/BSMs on application layer and not to the applied *data rate* on physical layer.

channel load *or* the adaptation of transmission parameters to meet the reception requirements of safety applications, we suggest to combine both objectives by adapting transmission rate within boundaries set individually for each vehicle based on its driving context. We provide a concrete example how a state of the art awareness approach [24] can be simplified and transformed to create an exemplary input function for the adaptation of a vehicle's *relative* transmission rate. In the evaluation, we show that based on the resulting Inter-Reception Distance (IRD) metric, the relative transmission rate adaptation can make a significant difference in the reception performance of a (simplified) safety application compared to a conventional congestion control approach which works without the guidance of safety applications.

Description of a concrete congestion control protocol. We introduce and evaluate Periodically Updated Load Sensitive Adaptive Rate control (PULSAR), a concrete implementation of the derived congestion control strategy and design principles. The protocol consists of three modular components, i.e., local channel load assessment, information sharing and transmission rate adaptation. Of these building blocks, especially the information sharing component plays a crucial rule with respect to fairness. By sharing congestion information over two hops, it enables vehicles to adapt their transmission rate based on the congestion state within their carrier sense range. In the evaluation, we compare PULSAR to other state-of-the-art approaches and show that it converges fair and efficiently in challenging scenarios specifically designed to stress the different protocol mechanisms.

Parts of the aforementioned contributions have been previously published in

- T. Tielert, D. Jiang, Q. Chen, L. Delgrossi, and H. Hartenstein, "Design Methodology and Evaluation of Rate Adaptation Based Congestion Control for Vehicle Safety Communications," in IEEE Vehicular Networking Conference (VNC), Best Paper Award, 2011 [25]

- T. Tielert, D. Jiang, H. Hartenstein, and L. Delgrossi, "Joint Power / Rate Congestion Control Optimizing Packet Reception in Vehicle Safety Communications," in The Tenth ACM International Workshop on VehiculAr Inter-NETworking, Systems, and Applications (VANET), 2013 [26]

- Q. Chen, D. Jiang, T. Tielert and L. Delgrossi, "Mathematical Modeling of Channel Load in Vehicle Safety Communications", in 4th International Symposium on Wireless Vehicular Communications (WIVEC), 2011 [27]

- A. Festag, P. Papadimitratos, and T. Tielert, "Design and Performance of Secure Geocast for Vehicular Communication," in IEEE Transactions on Vehicular Technology, Vol. 59, No. 5, pp. 2456-2471, June 2010 [15]

- T. Tielert, M. Killat, H. Hartenstein, R. Luz and S. Hausberger and T. Benz, "The Impact of Traffic-Light-to-Vehicle Communication on Fuel Consumption and Emissions", in Internet of Things Conference, 2010 [9]

- T. Tielert, D. Rieger, H. Hartenstein, R. Luz and S. Hausberger "Can V2X Communication Help Electric Vehicles Save Energy?", in 12th International Conference on ITS Telecommunications (ITST), Best Paper Award, 2012 [10]
- J. Härri, M. Killat, T. Tielert, J. Mittag and H. Hartenstein, "DEMO: Simulation-as-a-Service for ITS Applications", in 3rd IEEE International Symposium on Wireless Vehicular Communications (WiVeC), 2010 [28]

Outline

The thesis is structured as follows. An overview is provided in Figure 1.2 on page 10.

Chapter 2 on page 11 provides background information on VSC, addressing the foreseen safety applications, technical aspects and challenges as well as channel capacity and the need for congestion control. Further, we discuss the available degrees of freedom and their interrelationships.

Chapter 3 on page 45 provides an overview of the state of the art. We first address basic congestion control concepts from milestone papers. Then, we provide a short overview of the functionality of TCP as it is the most widely known congestion control protocol and shares some similar aspects with the suggested approach. Next, we discuss congestion control in other wireless networks. Finally, we survey existing congestion control approaches for vehicular communications, structured by control dimension and optimization objective. For each degree of freedom, we explain why or why not we consider them in our study. We reduce the number of considered degrees of freedom to two, i.e., to transmission rate and transmit power.

In **Chapter 4 on page 83**, we analyze the underlying optimization problem and solution space. First, we define the optimization criterion for our study as described above. Based on the defined criterion, we identify the optimal transmission parameter combinations for a uniform parameter configuration among vehicles as well as for two (groups of) vehicles optimizing simultaneously for different distances. To demonstrate the need for congestion control, we conclude the chapter by quantifying the difference between the identified optimal outcome and the result for some exemplary fixed parameter combinations.

Chapter 5 on page 103 introduces a design methodology for congestion control based on the findings in the previous chapter. First, we analyze the potential of TPC and TRC to approximate the identified optimal points. Then, we introduce the joint control strategy as explained above. We discuss a suitable system model, convergence target as well as different aspects of fairness. In the next step, we introduce the concept of adapting the relative transmission rate.

In **Chapter 6 on page 135**, we derive the logic behind the suggested protocol PULSAR. We describe each of PULSAR's three building blocks separately, introducing each

module's mechanisms and discussing different design decisions and potential pitfalls. At the end of the chapter, we summarize all mechanisms of PULSAR and lay out how they address the underlying design principles.

In **Chapter 7 on page 167**, we evaluate the performance of PULSAR. First, we study the impact of different factors on PULSAR's convergence to stability and fairness. Then, we demonstrate the feasibility of a dynamic adaptation of transmission rate and transmit power based on a vehicle's driving context. Based on the metric of IRD, we show why we expect the relative rate adaptation approach to have benefits for the performance of an exemplary safety application. In the next step, we compare PULSAR's performance to other state of the art congestion control approaches. We conclude the chapter by discussing potential limitations of our work.

Chapter 8 on page 219 concludes the thesis and addresses potential directions for further research.

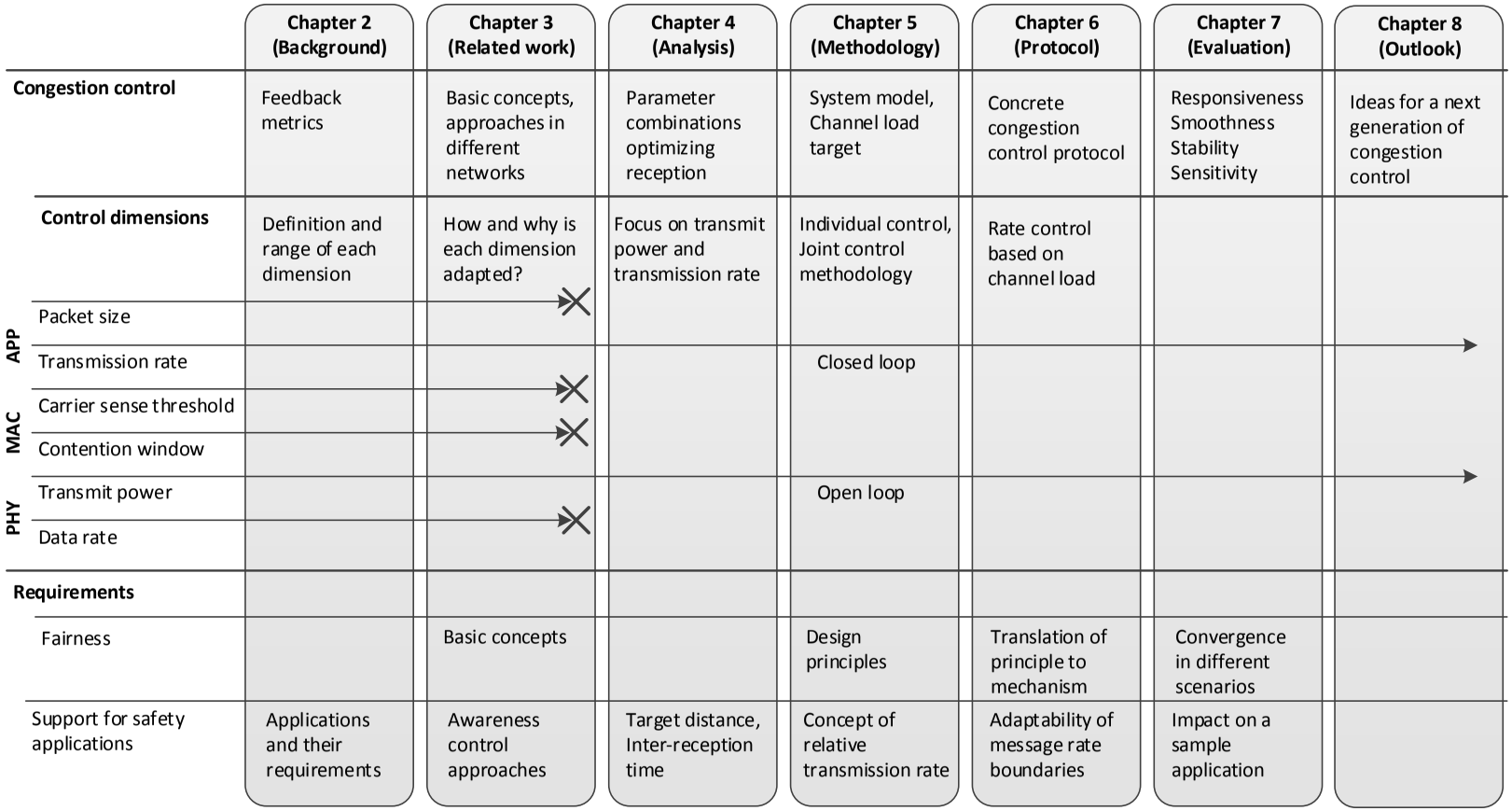


Figure 1.2: Overview of the structure and contents of this thesis

2

Vehicle Safety Communications

The term Vehicle Safety Communications (VSC) was coined as a project name in 2002, when the United States Department of Transportation (USDOT) and seven car manufacturers in the Crash Avoidance Metrics Partnership (CAMP) joined forces to "estimate the potential benefits of communication-based vehicle safety applications and define their communications requirements" [13]. Since then, a number of research initiatives worldwide have been dedicated to the topic. Recent activities include for example sim^{TD}[1], a German Field Operational Test (FOT) with 120 vehicles and 450 drivers, DRIVE-C2X[2], a project funded by the European Union with FOTs in seven locations across Europe, and the Safety Pilot Model Deployment[3] in the United States, a long-term FOT with approximately 3000 vehicles. In addition, an ongoing project of the USDOT and CAMP investigates the interoperability and scalability of Vehicle to Vehicle (V2V) safety communications [29]. Further information on research activities can be found in [3], [30], [31] and [32].

In this thesis, we use the term VSC to summarize safety-oriented V2V communication based on single-hop broadcast by means of IEEE 802.11p. This chapter provides background information on VSC as far as it is helpful to understand the research problem and the contributions of this thesis. For a comprehensive introduction to Vehicle to X (V2X) communication in general and VSC in particular, we kindly refer the reader to [1], [3] and [33]. Additional information regarding the standardization process in Europe and the United States can be found in [30], [29] and [34].

[1] Sichere Mobilität Testfeld Deutschland, *http://www.simtd.de*

[2] *http://www.drive-c2x.eu*

[3] *http://www.its.dot.gov/safety_pilot/*

2.1 Technology and Standards

In this section, we address the technological basis for VSC. We follow a bottom-up approach, starting with the allocated frequency band. Then, we discuss relevant aspects of the IEEE 802.11p standard and describe the foreseen protocol stack in Europe and the United States. We end this section by touching on Global Positioning System (GPS) and its implications for channel congestion.

2.1.1 Frequency Band

In the United States, 75 MHz of bandwidth between 5.850 and 5.925 GHz have been allocated for Dedicated Short Range Communications (DSRC)[4] in the context of Cooperative Intelligent Transport Systems (C-ITS) [1]. As illustrated by Figure 2.1 on the facing page, the DSRC spectrum is divided into a guard band of 5 MHz and seven channels of 10 MHz each[5]. Channels are assigned even numbers between 172 and 184. The center channel 178 is denoted as the *control channel*, while the other six are *service channels*. Since vehicles are not expected to be equipped with seven radios to listen to all channels simultaneously, the control channel is defined as a common "meeting point" which all vehicles are required to listen to on a regular basis [29]. The control channel is reserved for safety messaging and service announcements only. Of the six service channels, numbers 172 and 184 are reserved for VSC, while the others are open for safety or non-safety usage. Basic Safety Messages (BSMs) are expected to be sent on channel 172.

In Europe, a similar frequency band has been allocated. As visualized by the figure, the spectrum has been divided into different logical parts. Channels 176 to 180 are referred to as *ITS-G5A* and are reserved for safety usage, channel 180 being the control channel [35]. Cooperative Awareness Messages (CAMs) are expected to be transmitted on the control channel [37]. Of the remaining channels, numbers 172 and 174 are referred to as *ITS-G5B* and are dedicated to non-safety usage. Additionally, channels 182 and 184 may be allocated in the future under the name *ITS-G5D* [4][35]. Note that *ITS-G5C* is located between 5,470 GHz and 5,725 GHz and is foreseen for Infrastructure to Vehicle (I2V) communications [35].

The frequency band of 5.9 GHz chosen by both Europe and the United States comes with a number of technical challenges. Compared to lower frequencies, a signal at 5.9 GHz is subject to higher attenuation and absorption, resulting in a lower trans-

[4] The term DSRC can be ambiguous. While in the United States, it is often used as a synonym for IEEE 802.11p based V2X communications, it refers to wireless toll collect systems outside the 5.9 GHz band in Europe and Japan [1]. In this thesis, we adopt the U.S. American notion unless stated differently.

[5] The channel width of 10 MHz was chosen because it was found to be favorable with respect to effects like Doppler spreads [29], cf. also [36]. However, it is still an open issue whether 20 MHz channels would be a better option with respect to scalability [29].

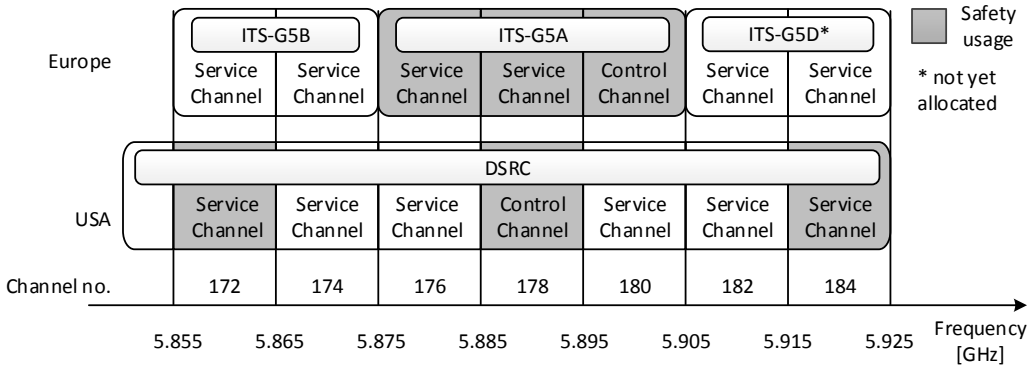


Figure 2.1: Spectrum and channel allocation for Cooperative Intelligent Transport Systems (C-ITS) in Europe and the United States [1][35]

mission range at the same output power [38]. In addition, high frequency signals are reflected more easily by surfaces and pass less easily through obstacles [39]. In the 5.9 GHz band, even minor obstacles like hedges and small trucks can lead to a significant reduction in the probability of reception [40]. Especially the latter factor, the obstruction through obstacles, caused Japan to move V2V safety communications to the 700 MHz band in spite of an already existing Vehicle to Infrastructure (V2I) infrastructure in the 5.8 GHz band [1]. Unlike signals at 5.9 GHz, signals at 700 MHz can pass through buildings and other vehicles [41].

2.1.2 IEEE 802.11p

Following the allocation of the 5.9 GHz frequency band in the United States, a decision was made to develop a new standard based on the IEEE 802.11a wireless Local Area Network (LAN) protocol [3]. The resulting standard IEEE 802.11p[6], consisting of a Physical layer (PHY) and a Medium Access Control (MAC) layer [12], forms the basis of the U.S. American Wireless Access in Vehicular Environments (WAVE) protocol stack, cf. Section 2.1.3 on page 18. In a nutshell, IEEE 802.11p is an adapted version of IEEE 802.11a with certain parameter changes to make the communication more robust against the challenges of a vehicular environment, e.g., severe multipath propagation and high relative speeds. In addition, it introduces a new communication mode called Outside the Context of a BSS (OCB). This concept allows stations to communicate directly with each other, without establishing or joining a Basic Service Set (BSS) first, and thus has advantages in terms of latency.

In a 10 MHz channel, IEEE 802.11p offers eight data rates as listed in Table 2.1 on the following page. Each data rate results from a combination of a *modulation scheme* and a *coding rate*. The modulation scheme determines how the physical waveform is

[6]Since 2012, the former amendment IEEE 802.11p [12] has been integrated into the IEEE 802.11 standard [42]. In this thesis, we still use the term IEEE 802.11p to emphasize the vehicular context.

Modulation	Coding rate	Data rate [Mbit/s]	SINR [dB]
BPSK	1/2	3	5
BPSK	3/4	4.5	6
QPSK	1/2	6	8
QPSK	3/4	9	11
16-QAM	1/2	12	15
16-QAM	3/4	18	20
64-QAM	2/3	24	25
64-QAM	3/4	27	N/A

Table 2.1: Available data rates in a 10 MHz channel using IEEE 802.11p [42] and their required SINR values for reception [43]

modified to convey information, e.g., by changing its phase and/or amplitude. The coding rate describes the ratio of the conveyed bits actually carrying data and the total number of bits, of which the rest is used for error correction. The last column of the table lists the required Signal to Interference and Noise Ratio (SINR) for the reception at each data rate, according to measurement results by Jiang et al. in [43]. We will get back to this topic in Section 2.3.1 on page 34.

For scalability and congestion control, two parameters of the MAC layer of IEEE 802.11p are of particular interest, i.e., *contention window* and *carrier sense threshold*, since they influence the temporal and spatial reuse of the channel. In the following, we thus provide an overview of the functionality of the MAC layer and the role of the two mentioned parameters. For more details on IEEE 802.11p, please refer to [3], [29] or the standard itself [42].

Carrier Sense Multiple Access (CSMA) and Contention Window

The medium access control of IEEE 802.11p is based on the Distributed Coordination Function (DCF) of IEEE 802.11 augmented by the Enhanced Distributed Channel Access mechanism of IEEE 802.11e. In the following, we first describe the basic functionality of DCF before getting back to EDCA on page 17.

As the name says, the DCF operates without central coordination. It employs Carrier Sense Multiple Access (CSMA), a random channel access mechanism based on the principle "listen before talk". Before transmitting a frame, each station monitors the channel for a certain time period, denoted as DCF Interframe Spacing (DIFS). If the channel was idle during this time, the station is allowed to access the medium. Otherwise, it has to perform a *backoff* procedure, i.e., the station waits for a random number of idle time slots before transmitting. The countdown starts with a random value between zero and the size of the contention window and is decremented while the medium is idle. If the medium becomes busy, the countdown is halted and resumed

after an idle period of DIFS has passed. An example for the coordinated channel access by three stations A, B and C is shown in Figure 2.2a on the following page. Station B finds the channel idle and starts a transmission after it has waited for the duration of DIFS. Stations A and C detect the ongoing transmission and delay their respective transmission until the medium is idle again. Station A draws a smaller backoff number than station C, i.e., three versus eight, and accesses the medium first. The example also shows the interruption of the countdown at station C during A's transmission.

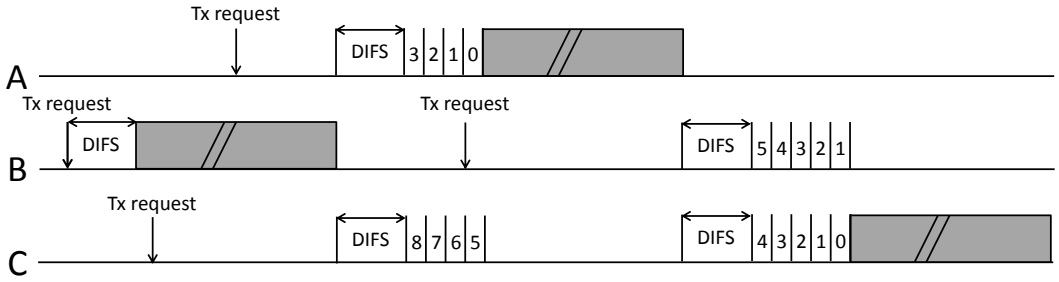
Obviously, the contention window size has an impact on CSMA's performance. If it is chosen too low compared to the number of contending stations, frame collisions due to simultaneous countdowns increase, while if it is too high, the channel is not used efficiently and stations have to wait longer to access the medium. The latter aspect is especially important for safety communications with its stringent latency requirements. It should be noted here that the *binary exponential backoff* procedure of IEEE 802.11, i.e., the doubling of the contention window size upon the detection of a frame collision, is not applicable to broadcast communications due to the absence of frame acknowledgments. The contention window thus stays at its minimum value CW_{min}, making its choice an important parameter for scalability.

Carrier Sensing, Hidden Terminals and Exposed Terminals

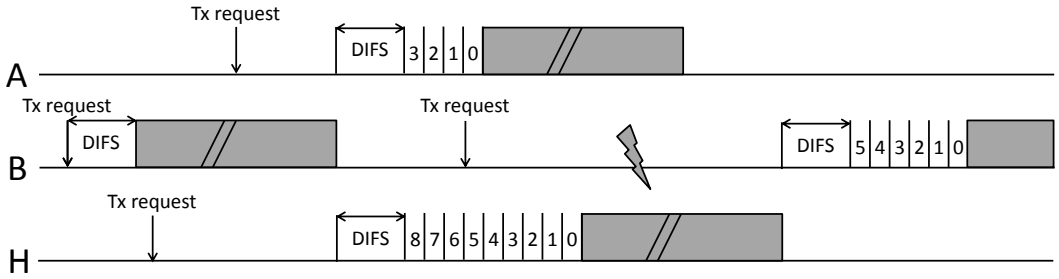
In the CSMA process, the state of the medium is determined by a binary function called Clear Channel Assessment (CCA). It consists of two mechanisms, *physical carrier sense* and *virtual carrier sense*. The physical carrier sense indicates channel busy if the total energy received at the antenna exceeds a certain value called *carrier sense threshold* P_{cs}. The virtual carrier sense is triggered if the preamble and header of a frame have been successfully received. Then, the medium is considered busy for the frame duration specified in the PHY header[7]. It should be noted here that a successful reception of preamble and header does not necessarily imply a successful reception of the entire frame. The CCA function returns busy if one or both of physical and virtual carrier sense indicate a busy channel [42].

The carrier sense threshold is a parameter that determines how "aggressively" a station accesses the medium. If it is set low, a station refrains from accessing the channel for received signals with low power which likely originate from more distanced nodes. In the reverse conclusion, a high carrier sense threshold may cause the station to ignore transmissions from closer neighbors. If a station transmits while another is accessing the channel, both signals may overlap at a receiving station, making it impossible to decode any one of them. The occurrence of such a *packet collision* is often caused by the *hidden terminal problem*. A hidden terminal is one which is not able to detect an ongoing transmission based on the carrier sensing mechanism. A second phenomenon related to the CSMA mechanism is the *exposed terminal*

[7] To help with this process, the signaling part of the PHY header is always modulated with the lowest available data rate of 3 Mbit/s, cf. Table 2.1 on the preceding page



(a) Coordinated



(b) Uncoordinated

Figure 2.2: Carrier Sense Multiple Access (CSMA) example

problem. An exposed terminal is one which unnecessarily refrains from transmitting even though it would not interfere with the ongoing transmission.

Figure 2.3 on page 18 illustrates a unicast scenario in which node A transmits a packet to node B. The example is based on a homogeneous transmit power and carrier sense range for all nodes. The solid line and dashed line circles indicate a node's *transmission range*, i.e., the range within which a packet can be successfully received, and its *carrier sense range*, i.e., the range up to which the received signal strength of the vehicle's transmission triggers a receiving radio into channel busy state, respectively. In the example, node H is a hidden terminal, since it cannot sense the ongoing transmission. On the other hand, node E is an exposed terminal, since it refrains from transmitting to node C, even though it would not disrupt the communication between A and B.

Note, however, that in a broadcast scenario, node E would rightly refrain from transmitting, since node C would be an intended receiver of A's transmission. In [44], Stanica et al. thus argue that the vehicular broadcast environment is exposed-node free. However, a node which refrains from transmitting too often due to a low carrier sense threshold may experience a very long channel access delay. In [45], Schmidt et al. thus argue that in the vehicular broadcast context, an exposed terminal is one which has to drop packets in its internal message queue.

Figure 2.2b illustrates the impact of the hidden terminal problem on the sample transmission studied above. Instead of nodes A, B and C, we now consider nodes A, B and H in Figure 2.3 on the following page. Since A and H are uncoordinated,

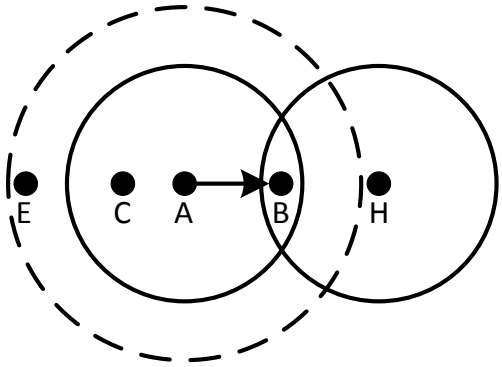


Figure 2.3: Hidden terminal (H) and exposed terminal (E) in unicast communication

both count down their backoff slots and start transmitting. In consequence, a packet collision occurs at B. After the medium is idle again, a new round begins and B accesses the channel. More information about the coordination and incoordination behavior of CSMA can be found in [46], [47] and [48].

To counter the hidden terminal problem, IEEE 802.11 employs a channel reservation based mechanism. Before transmitting a frame, a station first sends a short Ready To Send (RTS) message containing the frame duration and the intended receiver. The receiving station returns a prioritized Clear To Send (CTS) message repeating the frame duration. Neighboring stations receiving the CTS message infer that they may be hidden terminals and refrain from transmitting during the specified duration [49]. However, like frame acknowledgments, the RTS/CTS mechanism is not applicable to broadcast communications for scalability reasons.

Priority Access via Enhanced Distributed Channel Access (EDCA)

In the RTS/CTS mechanism, priority access for CTS messages is achieved by assigning them a shorter waiting interval than DIFS, denoted as Short Interframe Spacing (SIFS). A similar functionality is included in IEEE 802.11p to prioritize between different (types of) messages. The Enhanced Distributed Channel Access (EDCA) mechanism, adopted from IEEE 802.11e, is an enhancement of the DCF which allows for different levels of Quality of Service (QoS). The main difference[8] between EDCA and DCF is that DIFS is replaced by Arbitration Interframe Space (AIFS), an interval which is calculated as the sum of SIFS and a multiplicative of the default slot time[9]. The multiplier, denoted as AIFS Number (AIFSN), determines the priority of the frame to be sent, since frames with a shorter interframe spacing get to access the channel before all others. EDCA includes four Access Categories (ACs) which are referred to by their designation in IEEE 802.11e, i.e., background (AC_BK), best effort (AC_BE),

[8] Another difference is the introduction of a contention-free transmission period denoted as Transmission Opportunity (TXOP). However, in IEEE 802.11p TXOP is set to zero [50]. Thus, it is not considered in this thesis.

[9] The slot time in IEEE 802.11p is 13 μs for a 10 MHz channel [42].

video (AC_VI) and voice (AC_VO), in ascending order of priority. Each AC has its own queue and acts as an independent DCF station. Next to different AIFSNs, ACs also differ in the assigned contention window sizes. An overview of the default values specified in the IEEE 802.11p standard is provided in Table 2.2 on page 19. Note that a further difference between DCF and EDCA is the way the backoff counter is decremented, cf. Section 7.5.2 on page 216.

The mapping of V2X message types to access categories has not been standardized yet. Typically, it is assumed that each message type is associated with a specific AC. A different approach is described in [51], where the authors suggest to randomly assign ACs to messages in order to improve the overall reception probability.

2.1.3 Protocol Stacks in Europe and the United States

Both Europe and the United States have developed protocol stacks for C-ITS. The responsible standardization bodies in Europe are the European Telecommunications Standards Institute (ETSI) and the European Committee for Standardization (CEN). In the United States, standardization is in the hands of the Institute for Electrical and Electronics Engineers (IEEE) and the Society for Automotive Engineers (SAE). Additional C-ITS standards are being developed by the International Organization for Standardization (ISO).

Figure 2.4 on page 20 provides an overview of relevant[10] protocols and standards with respect to the layers of the ISO/OSI reference model[11]. In the United States, the lower four layers of the protocol stack are typically referred to as Wireless Access in Vehicular Environments (WAVE) [1], while the upper three layers are summarized into one [29]. In Europe, the protocol stack is being defined by the Technical Committee (TC) for Intelligent Transportation System (ITS) of ETSI. It has been compacted into three layers, of which the entities *Access, Network & Transport* and *Facilities* summarize the lower two, middle two and upper three layers of the ISO/OSI stack, respectively [52]. The ETSI stack additionally specifies an *application layer*, currently defining three safety applications, cf. Section 2.2 on page 25. Finally, both protocol stacks feature cross-layer entities for security and management. In the following, we go through the different layers from bottom to top. More information can be found in [29] and [53] for the United States and in [34] for Europe.

[10] Both protocol stacks include protocols for non-safety communications as well. These TCP/IP related branches of the protocol stack are omitted here for presentation clarity. The interested reader is kindly referred to [29] and [52] for more information.

[11] The Open Systems Interconnection (OSI) model developed by ISO standardizes different layers in a communication architecture. For details, please refer to networking textbooks, e.g., [49].

AC	Description	CW_{min}	CW_{max}	AIFSN	AIFS [μs][a]
AC_BK	Background	15	1023	9	149
AC_BE	Best effort	15	1023	6	110
AC_VI	Video	7	15	3	71
AC_VO	Voice	3	7	2	58

Table 2.2: Default EDCA parameters of IEEE 802.11p according to [42]

[a] Based on a slot time of 13 μs in a 10 MHz channel.

Physical and Data Link Layers

In both Europe and the United States, the IEEE 802.11p PHY and MAC layers form the basis of the protocol stack. Further, both protocol stacks employ the IEEE 802.2 Logical Link Control (LLC) sublayer which offers a connectionless, unacknowledged data transfer service [54]. Multiplexing is achieved by means of the SubNetwork Access Protocol (SNAP) [55]. The main difference between the lower layers of the two stacks is the WAVE multi-channel operation standard IEEE 1609.4 [56] which is not foreseen in Europe. The standard allows a radio to switch between different channels and introduces the concepts of control and service channels. The basic idea is to divide time into synchronized alternating 50 ms control channel periods and 50 ms service channel periods. Vehicles are required to return to the control channel during the control channel intervals and are allowed to switch to a service channel during the rest of the time.

However, the drawback of the time division is that only 46% of the total time are available for safety messaging (50 ms control channel interval minus 4 ms guard interval) [29]. Since the resulting channel capacity is not expected to be sufficient for safety applications, IEEE 1609.4 has been made optional in the United States and additionally, a consensus is forming to use channel 172 for safety messaging without channel switching [29]. A similar debate to use two radios instead of one in order to allow continuous listening to the control channel is currently going on in Europe [57].

Network and Transport Layers

Two types of network and transport layers are foreseen in both protocol stacks, one tailored toward safety applications and one toward non-safety applications. For the latter, the well-known Internet suite is available, consisting of Internet Protocol (IP) version 6, Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). For safety applications, custom solutions have been developed.

In Europe, a series of network layer standards has been summarized under the term *GeoNetworking*. The basic idea is that typically, safety information is relevant for all vehicles within a certain geographical area. Due to the high mobility of vehicles, it thus makes sense to address a message not to a specific vehicle but to all vehicles

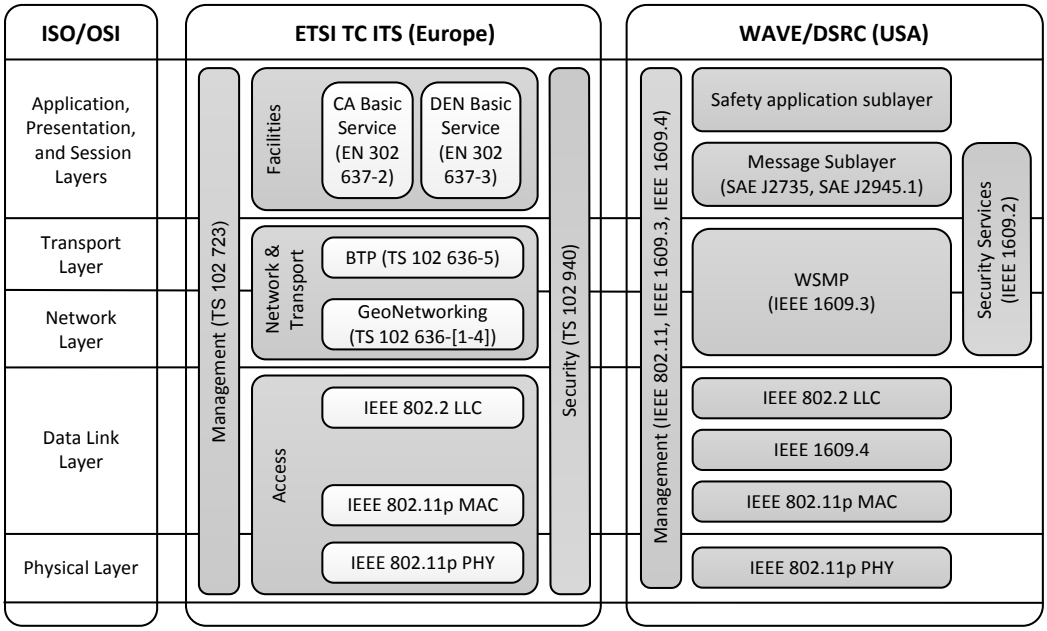


Figure 2.4: Overview of protocol stacks and corresponding standards for VSC in Europe and the United States

within the respective target area. If the distance between source and destination is larger than the communication range, the message is relayed over multiple hops. Corresponding *Geocast* protocols are designed to efficiently disseminate messages over large distances in a short period of time. For example, one of the protocol candidates for the GeoNetworking standard [58] named Emergency Message Dissemination for Vehicular environments (EMDV) [20] was shown in simulations to disseminate a message within a 2 km radius in less than 200 ms [15].

On top of GeoNetworking, the ETSI protocol stack employs the Basic Transport Protocol (BTP), a lightweight, connectionless end-to-end transport protocol. The main purpose of BTP is to provide a port-based multiplexing functionality for facilities layer instances [59], cf. the description of the upper layers below.

In contrast to the European approach, the U.S. American protocol stack is targeted toward one-hop[12] safety messages [53]. Consequently, the routing functionality has been omitted and a lightweight transport layer protocol named WAVE Short Message Protocol (WSMP) has been developed [60]. The main advantage of WSMP compared to IP is that it significantly reduces packet overhead [29]. Similar to BTP, WSMP offers a multiplexing functionality for upper layers. The packets sent using WSMP are referred to as WAVE Short Messages (WSMs). In addition, WAVE Service Announcements (WSAs) are specified to allow the transmission of information regarding locally offered services.

[12] Note that this does not necessarily mean that DSRC applications will not be using multi-hop communications. Such a functionality could still be provided by upper layers.

Upper Layers

The facilities layer in the ETSI protocol stack has the objective to fulfill common functional and operational requirements of V2X applications [61], structured into communication, information and application support. For safety communications, especially the Local Dynamic Map (LDM), Cooperative Awareness (CA) Basic service and the Decentralized Environment Notification (DEN) Basic Service are of interest. The LDM is essentially a neighbor table containing the local data gathered about other vehicles' positions and statuses. The CA and DEN basic services are responsible for generating CAMs and Decentralized Environment Notification Messages (DENMs), respectively.

CAMs are single-hop messages with mandatory and optional parts. The mandatory High Frequency (HF) container includes frequently changing vehicle information like position, speed and heading. The Low Frequency (LF) container, mandatory if the last low frequency CAM was sent 500 ms ago or more, contains long-term characteristics of the vehicle like its physical size as well as additional information like path history [62]. CAMs are generated by the CA basic service on a periodic basis, with a minimum and maximum interval of 100 ms and 1000 ms, respectively. According to the current state of the standard[13], the CAM interval is determined by a set of trigger rules, depending on how much the heading, position and speed of the transmitting vehicle has changed since the last generated CAM. In addition, the application layer is optionally allowed to lower the maximum CAM interval. However, any of these rules can be overridden by Decentralized Congestion Control (DCC) by setting a lower bound to the CAM interval according to channel congestion [62].

In contrast to CAMs, DENMs are triggered directly by the application layer upon the detection of a dangerous situation, e.g., a hard braking event of the transmitting vehicle. DENMs can be single hop or multi hop, depending on the targeted geographical area. They are repeated periodically at their destination until the event expires. While the initial DENM is exempt from DCC [63][64], the repetition rate of DENMs could be controlled by DCC [64].

The counterpart to the European CAM in the United States is designated Basic Safety Message (BSM). Like the CAM, the BSM has a mandatory *Part I* with frequently changing vehicle information and an optional *Part II* with additional information like path history [29]. BSMs are sent as WSMs and can be triggered periodically or event based by safety applications. The BSM is one of fifteen message types defined in the DSRC Message Set Dictionary [17]. In addition, a new standard SAE J2945.1 specifying the minimum performance requirements of safety applications is currently under way. The document is going to address the requirements regarding BSM generation rate and transmit power, accuracy of sensor and position information as well as channel congestion control [29].

[13] As this thesis is being written, the CAM generation rules are under revision and may change in the future. The description given here is based on the Draft standard ETSI EN 302 637-2 v.1.3.0 [62].

Management and Security

Both protocol stacks include cross-layer entities dedicated to management and security. Management entities are responsible for the maintenance of cross-layer parameters as well as for the transmission and reception of service advertisement messages [29][52].

The security entity can be seen as a specific part of the management [52]. In the United States and with adaptations in Europe, security is based on IEEE 1609.2 which defines mechanisms to authenticate and encrypt messages, in particular WSMs and WSAs [65]. For authentication, the standard employs asymmetric cryptography in the form of the Elliptic Curve Digital Signature Algorithm (ECDSA). Compared to other algorithms like RSA (Rivest, Shamir and Adleman), ECDSA creates significantly smaller signatures and thus less packet overhead [1]. The price to pay, however, is an increased processing time which can become an issue if a vehicle has to check hundreds of incoming messages per second, cf. [15].

To verify a signature, a receiver requires a certificate which contains the sender's public key, signed by the issuing certificate authority. For VSC, the authentication of safety messages is vital but comes at the price of a significant overhead in packet size. ECDSA signatures and certificates are typically in the order of 55 to 65 Bytes and 120 to 180 Bytes, respectively [15][29][66][67][68]. The overhead of certificates can be reduced by 50 to 60 Bytes via *implicit certificates* [29]. To further reduce overhead, it is also under discussion to attach a *certificate digest*, i.e., a hash of the certificate in the order of a few bytes, to the majority of packets and to send explicit certificates only with every n^{th} message [29].

Security in vehicular communications is a research topic of its own and can only be touched on in this thesis. In particular, safety applications require a short-term trackability of vehicles, while a long-term trackability is undesired for privacy reasons [69]. For more information, please refer to, e.g., [1], [66], [67], [69] or [70].

Size and Structure of a CAM/BSM

For congestion control, the size of the messages to be transmitted is of high importance, cf. Section 2.3 on page 32. However, due to the modularity and flexibility of the structure of CAMs and BSMs, it is difficult to specify a definite message size. Figure 2.5 on the facing page illustrates an estimation of the minimum and maximum lengths of the different components of both message types in relation to the expected security overhead. While the protocol headers of the lower layers are fixed, the size of the payload and the frequency at which the payload elements are sent out are currently still under investigation [29][62]. Especially the estimations of the CAM payload size are tentative. The current draft version of the standard [62] specifies the CAM in the format ASN.1, but does not provide concrete numbers in

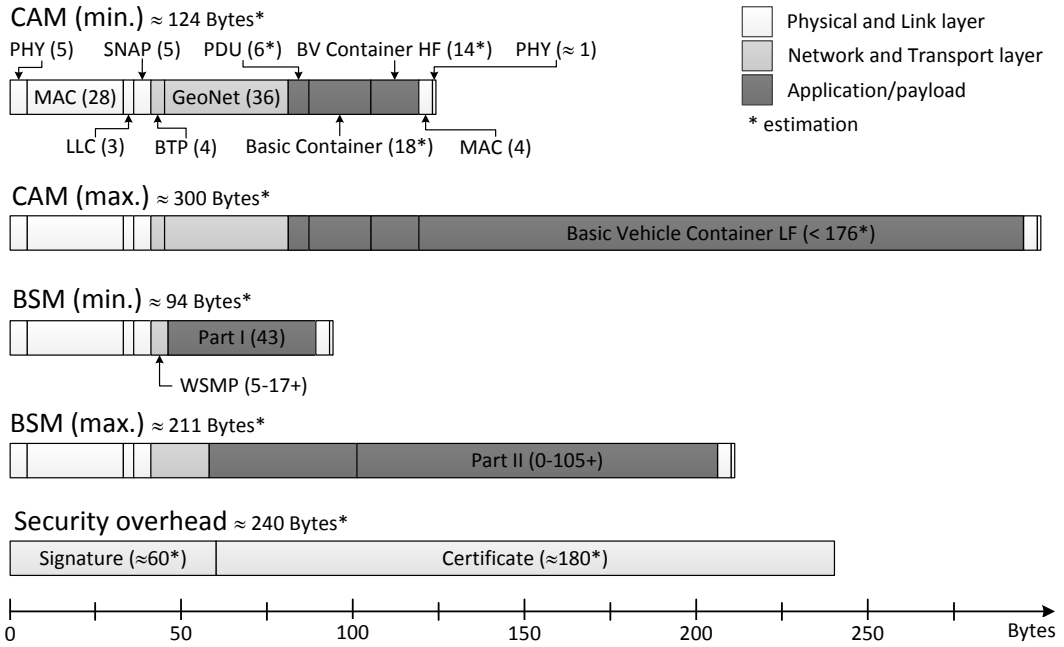


Figure 2.5: Structure and size range of a Cooperative Awareness Message (CAM) and Basic Safety Message (BSM) (estimation based on [29], [34], [53] and the respective standards)

terms of Bytes. The values shown are preliminary results from ETSI published in [34] and may change in the future.

Based on the results shown in Figure 2.5, a reasonable average size for a beacon, i.e., a CAM or BSM, appears to be somewhere near 160 Bytes. Adding the estimated 240 Bytes of security overhead, a total packet size of 400 Bytes is chosen in this thesis unless specified otherwise. This assumption is in line with other works, e.g., [34][48][71].

Decentralized Congestion Control (DCC)

As mentioned above, channel congestion control in the United States is currently under development and is expected to operate in the upper layers of the protocol stack [1][29]. In Europe, the foreseen Decentralized Congestion Control (DCC) has a cross-layer architecture consisting of components in the access, network, facilities and management layer. As of writing this thesis, only the access layer part has been published [72]. It specifies a general framework with four options, i.e., Transmit Power Control (TPC), Transmission Rate Control (TRC), Transmit Datarate Control (TDC) and DCC Sensitivity Control (DSC). Thereby, TRC refers to how often packets are sent, TDC refers to the modulation scheme and DSC refers to the carrier sense threshold. The latter is adapted to resolve *local channel congestion*, i.e., a condition in which packets are dropped locally because the channel could not be accessed.

According to [72], the state of the channel is determined by means of *channel probing*, i.e., a sampling of the current busy/idle state of the channel and the calculation of the ratio of time the channel was perceived as busy, cf. Section 2.3.3 on page 40. Depending on the determined channel load, DCC is in one of three states, i.e., *relaxed*, *active* or *restrictive*, of which the *active* state may consist of different sub-states. Each state and sub-state is associated with a fixed set of parameters to be applied to the different degrees of freedom. State transitions are triggered by means of fixed channel load thresholds.

Different studies have shown in simulations that the current version of the DCC access standard may result in an oscillatory behavior, e.g., [73] and [74]. It is thus possible that the algorithm is going to be updated in the future.

2.1.4 Global Positioning System (GPS)

To facilitate cooperative awareness, vehicles require a localization technology with lane-level accuracy, i.e., 1 m [64][75] to 1.5 m [1]. Typically, the Global Positioning System (GPS) is used for this purpose. GPS position updates can be obtained at a rate of 10 Hz with current receivers [18]. To determine its position, a vehicle requires location and timing information from at least four GPS satellites. Based on the measured signal propagation time, it calculates the distances between itself and the satellites and triangulates its current position. The fourth satellite is required to compensate for the reduced timing precision of common devices [1]. The current GPS standard specifies a *worst case* accuracy of 7.8 m in 95% of the cases under normal usage [76]. However, GPS errors can typically be expected to be smaller, at least in open field conditions. For example, a measurement study from 2011 indicates a 95% cut-off error of 2.2 m[14]. Note that in an urban environment, GPS may be significantly less accurate. To enhance GPS accuracy, different technologies are available, e.g., Differential GPS (DGPS). In addition, a vehicle's on-board sensors can be applied to extrapolate its position if satellite connection is lost, e.g., due to a tunnel [1].

For cooperative awareness, the *relative positioning* of vehicles is generally more important than their absolute position [1]. An unwanted side effect of the deployment of GPS receivers from different manufacturers is that their individual positioning errors may add up if positioning information is shared [1]. To solve this issue, it is under discussion to attach raw GPS measurement data to the transmitted messages, e.g., in Part II of the BSM [1]. However, this practice would significantly increase packet size and thus add to channel congestion.

[14] *http://www.gps.gov/systems/gps/performance/accuracy*

2.2 Safety Applications and Their Requirements

On top of the protocol stacks described in the previous section, a variety of safety and non-safety applications have been suggested. In the following, we focus on safety applications and what is currently known about their requirements. Since there are differences between the United States and Europe regarding the selection and implementation of safety applications, we look at both regions separately. Then, we discuss the concept of *cooperative awareness* and review frequently used metrics for its quantification. For more information on efficiency and infotainment V2X applications, we kindly refer the reader to, e.g., [77].

2.2.1 Focus in the United States

In the United States, a series of research projects has been dedicated to the definition, implementation and testing of DSRC safety applications. In the following, we discuss three projects conducted by the USDOT and CAMP, a research organization of leading U.S. American vehicle manufacturers.

Vehicle Safety Communications (VSC) Project (2002-2004)

More than 75 application scenarios were identified and analyzed in the aforementioned VSC project, resulting in a description of 34 selected safety applications [13]. Of these 34 applications, eight were marked as "high priority" based on the expected safety benefit when taking into account the expected rate of market penetration. For each application, the report lists "preliminary" communication requirements as shown in Table 2.3 on the following page. The selected set of applications covers a wide spectrum of V2X application characteristics, i.e., broadcast as well as unicast, V2V as well as V2I/I2V and periodic as well as event based communication. For most applications, the specified transmission rate is 10 Hz, the maximum latency is 100 ms and the required communication range is between 150 m and 400 m. However, the report leaves open how exactly these communication parameters were derived. An exception is the Forward Collision Warning (FCW) application, for which the report states that current radar based FCW systems have an update rate of 10 Hz and a detection range of 150 m, which the DSRC based system should match or exceed.

Vehicle Safety Communications - Applications (VSC-A) Project (2006-2009)

The results of the VSC project were taken up by a successor project named Vehicle Safety Communications - Applications (VSC-A) which implemented and tested prototypes of six selected safety applications[15]. The selection process was based on

[15] Note that while the VSC-A project focused on V2V applications, other projects, e.g., the Cooperative Intersection Collision Avoidance System for Violations (CICAS-V) project, implemented and tested V2I/I2V communications [1].

Application name	Direction	Mode	Trigger	Rate	Latency	Range
Emergency Electronic Brake Lights	V2V	Broadcast	Event	10 Hz	100 ms	400 m
Pre-Crash Sensing	V2V	Unicast	Event	50 Hz	20 ms	50 m
Cooperative Forward Collision Warning	V2V	Broadcast	Periodic	10 Hz	100 ms	150 m
Lane Change Warning	V2V	Broadcast	Periodic	10 Hz	100 ms	150 m
Traffic Signal Violation Warning	I2V	Broadcast	Periodic	10 Hz	100 ms	250 m
Curve Speed Warning	I2V	Broadcast	Periodic	1 Hz	1000 ms	200 m
Left Turn Assistant	V2I, I2V	Broadcast	Periodic	10 Hz	100 ms	300 m
Stop Sign Movement Assistance	V2I, I2V	Broadcast	Periodic	10 Hz	100 ms	300 m

Table 2.3: High-priority safety applications and their communication requirements as of the VSC project task 3 final report [13]

the analysis of crash statistics in the United States. The experts created a ranking of frequent crash scenarios and analyzed which DSRC-based safety applications were suitable to address the respective scenario. The result was a set of six safety applications as described in Table 2.4 on page 28.

The prototype implementations employed a *Target Classification* mechanism based on the reception of BSMs, creating a 360° view around the Host Vehicle (HV), i.e., the one running the safety application. As illustrated by Figure 2.6 on the facing page, each detected Remote Vehicle (RV) is assigned one of twelve categories depending on its relative position to the HV, e.g., ahead, ahead right, ahead far left, oncoming right, intersecting left, and so on. In the figure, categories involving the HV's *far left* and *far right* were omitted for better readability. The HV and RVs are indicated in dark gray and light gray, respectively. The figure further shows which classes of RVs the different safety applications are interested in. For example, the FCW application only monitors the behavior of RVs ahead and in the same lane as the HV [78]. To correctly classify RVs into categories, the HV requires the reception of BSM information, i.e., position information from Part I and path history and prediction data from Part II [78], from vehicles within application-dependent geographic zones. Table 2.5 on page 29 describes these zones for each application as defined in [78].

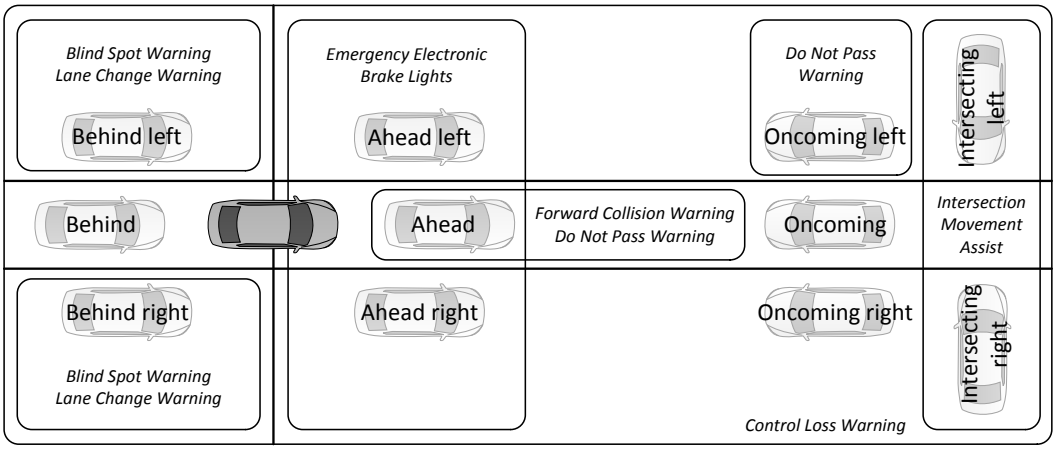


Figure 2.6: Target classification (simplified) and corresponding safety applications of the VSC-A project [6]

Safety Pilot Project (2011-2013)

Since 2011, the evaluation and testing of four of the six selected VSC-A safety applications, i.e., FCW, Blind Spot Warning (BSW)+Lane Change Warning (LCW), Intersection Movement Assist (IMA) and Emergency Electronic Brake Lights (EEBL), has been continued and extended in the *Safety Pilot*[16] project. The project consists of *Driver Clinics*, in which the reaction of naïve drivers and the resulting safety benefit are evaluated, and the aforementioned *Model Deployment* in Michigan, where safety applications are tested in a long-term, large-scale and everyday life scenario. The continued evaluation by the USDOT and leading car makers emphasizes the relevance of the selected safety applications.

2.2.2 Focus in Europe

In Europe, use cases for V2X communications have been described by different research initiatives, e.g., the EU FP6 IP project SAFESPOT[17] [79], the EU FP7 IP project PRE-DRIVE C2X[18] as well as the Car-to-Car Communications Consortium (C2C-CC)[19], a research organization of leading European vehicle manufacturers [80]. The results of these initiatives have been taken up in the ETSI Basic Set of Applications (BSA) document [8] which defines the use cases of the ETSI TC ITS application layer as described in the following.

[16] *http://www.its.dot.gov/safety_pilot*

[17] *http://www.safespot-eu.org*

[18] The project's website as well as the deliverable D4.1 "Detailed description of selected use cases and corresponding technical requirements" from 2008 are no longer publicly available.

[19] *http://www.car-to-car.org*

Application	Description
Forward Collision Warning (FCW)	Warns the driver of an "impending rear-end collision" with a vehicle "ahead in traffic in the same lane and direction of travel".
Blind Spot Warning + Lane Change Warning (BSW+LCW)	Warns the driver "during a lane change attempt if the blind-spot zone is ... or will soon be occupied by another vehicle traveling in the same direction." In addition, the application informs the driver at any time if a vehicle is detected in the blind spot zone.
Do Not Pass Warning (DNPW)	Warns "during a passing maneuver attempt when a slower vehicle, ahead and in the same lane, cannot be safely passed" because of oncoming traffic. In addition, the application informs the driver if such a maneuver would be safely possible at any time when a vehicle is detected ahead and in the same lane.
Intersection Movement Assist (IMA)	Warns if "it is not safe to enter an intersection due to high collision probability" with other vehicles. Initially, the focus is on "stop-sign controlled and uncontrolled intersections".
Emergency Electronic Brake Lights (EEBL)	Transmits a dedicated warning message if a vehicle is braking hard; receiving vehicles can warn their drivers in spite of Line of Sight (LOS) obstructions (other vehicles, fog, etc.)
Control Loss Warning (CLW)	Transmits a dedicated warning message if control loss is detected; receiving vehicles can warn their drivers.

Table 2.4: V2V safety applications evaluated by the VSC-A project [6]

ETSI TC ITS Basic Set of Applications (BSA)

The BSA document ETSI TR 102 638 presents a collection of more than 50 use cases for V2X communications, associated with seven (preliminary) applications which are divided into four application classes, i.e., active road safety, cooperative traffic efficiency, cooperative local services and global internet services [8]. An overview of the safety-related use cases and their communication requirements as of the BSA document is presented in Table 2.6 on page 30. It should be noted here that the European terminology is different from the U.S. American one. For example, EEBL is a *use case* in Europe, not an *application*. The use cases and communication requirements listed in the BSA document are similar to the ones described in the VSC final report, cf. Section 2.2.1 on page 25. CAM-based safety applications typically require

Application	Minimum performance requirements
FCW	Detection of vehicles within 300 m ahead and in the same lane.
BSW+LCW	Detection of vehicles within 20 m behind and of vehicles that will enter the blind spot zone within 5 s.
DNPW	Detection zone is defined by concept but not by concrete numbers.
IMA	Detection of vehicles within 300 m of an intersection.
EEBL	Detection of hard braking events from vehicles within 300 m ahead.
CLW	Detection of events from vehicles within 150 m in the same direction and 300 m in the opposite direction.

Table 2.5: Communication-relevant minimum performance requirements of V2V safety applications as defined by the VSC-A project [78]

a transmission rate of 10 Hz and a latency of 100 ms. The required communication range is not specified in the document.

ETSI TC ITS Applications

On top of the Facilities layer, the ETSI TC ITS architecture specifies an *Applications* layer for safety, traffic efficiency and "other" applications [52]. Three safety applications have been defined, i.e., Road Hazard Signaling (RHS), Longitudinal Collision Risk Warning (LCRW) and Intersection Collision Risk Warning (ICRW) [64]. The objective of the RHS application is to "make the driver aware of a road hazard which may become the cause of an accident" [64]. The application currently addresses ten different use cases among which are *emergency vehicle approaching, slow vehicle, emergency electronic brake lights* and *roadwork*. To achieve its objective, the RHS applications makes use of the CA and DEN Basic Services, cf. Section 2.1.3 on page 18. As of writing this thesis, the standards for the LCRW and ICRW applications are still to be published.

Figure 2.7 on page 32 illustrates the time horizons in terms of Time To Collision (TTC)[20] targeted by the applications defined so far by ETSI. While the RHS application is categorized as *driver awareness*, the LCRW and ICRW applications fall in the category *driver warning*. The major difference between the two categories is that the latter requires the driver to take action immediately, while the former does not [64]. Consequently, different QoS levels are foreseen for each category. If the *ego vehicle*[21], i.e., the vehicle running the application, is in warning or pre-crash state (and the channel is "not in a congested situation"), the RHS document specifies a maximum

[20] In the standard, the TTC is calculated as the sum of maximum end-to-end latency, maximum driver reaction time, maximum action time, i.e., the time required for crash countermeasures like braking, and a safety margin addressing, e.g., potential measurement errors [64].

[21] We use the terms *ego vehicle* and Host Vehicle (HV) synonymously in this thesis.

Use case	Rate	Latency
Emergency vehicle warning	10 Hz	100 ms
Slow vehicle warning	2 Hz	100 ms
Motorcycle warning	2 Hz	100 ms
Overtaking vehicle warning	10 Hz	100 ms
Lane change assistance	10 Hz	100 ms
Pre-crash sensing warning	10 Hz	50 ms
Co-operative glare reduction	2 Hz	100 ms
Across traffic turn collision risk warning	10 Hz	100 ms
Merging Traffic Turn Collision Risk Warning	10 Hz	100 ms
Co-operative merging assistance	10 Hz	100 ms
Intersection Collision Warning	10 Hz	100 ms
Co-operative forward collision warning	10 Hz	100 ms

Table 2.6: CAM-based use cases of the ETSI Basic Set of Applications (BSA) [8]

latency of 300 ms from the detection of an event to the initiation of a countermeasure by the receiver as well as a minimum transmission range of 300 m. Otherwise, a maximum latency of 1.5 s applies and a channel congestion based reduction of the transmission range is allowed [64].

2.2.3 Cooperative Awareness

To evaluate the performance of safety applications, a comparison of the number and severity of crashes with or without the respective application would be desirable. Experiments with real drivers in a controlled environment can provide valuable insight, e.g., by means of a real vehicle on a simulated road [81] or an actual crash situation using a dummy vehicle [82]. However, such experiments can be costly, require special equipment and typically have a small number of participating drivers. It would thus be beneficial to assess driver assistance systems in simulations.

Some recent works have integrated driver behavior and car following models into a communication network simulator, evaluating for example the vehicle speed at crash impact [83], the ratio of cars involved in a crash [84] and the predictability of crashes based on the received V2X messages [85]. However, obtaining realistic figures by

means of simulations is challenging, since crashes are rare and extraordinary events[22], influenced significantly by human behavior and thus difficult to model [86][87].

On the other hand, the ability of a safety application to correctly determine the likelihood of a crash with a neighboring vehicle depends on the availability and quality of information on this particular neighbor's position and dynamics, a concept often referred to as *(cooperative) awareness*. For example, the target classification introduced in the VSC-A project can be interpreted as a tool to manage the host vehicle's awareness of its neighbors' presence, cf. Figure 2.6 on page 27. In the following, we review different metrics for cooperative awareness which have been suggested in the literature. They each address one of more of the three dimensions of awareness formulated in [88]: *Dissemination area*, *latency* and *reliability*.

Packet Inter-Reception Time (IRT) and Related Metrics

From the HV's point of view, the up-to-dateness of the information on a RV's position depends on when the last beacon message has been received from this particular neighbor. This time difference between two successive and successful reception events for a sender-receiver pair was denoted as Packet Inter-Reception Time (IRT) and introduced[23] as a metric for awareness in [90]: "[T]he IRT metric quantifies the latency perceived by an application running on top of a given protocol suite. This suggests that IRT could be used for comparing different protocols and parameter settings even if the application latency requirements are not precisely known." The authors point out that IRT captures the channel access delays of the IEEE 802.11p MAC (cf. Section 2.1.2 on page 13) as well as consecutive packet losses caused by interference and channel obstructions.

Closely related to the IRT metric, Bai et al. introduced the concept of *T-window reliability*, an application level metric which is defined as "the probability of successfully receiving at least one single packet from neighbor vehicles during the tolerance time window T" [91]. In other words, an application is defined to be reliable if IRT is less than or equal to T. The authors study the relationship between *communication reliability* in the form of the Packet Delivery Ratio (PDR), i.e., the probability of receiving an individual message, and *application reliability* in the form of T-window reliability. They analyze the distribution of consecutive packet drops and come to the conclusion that, in a constant environment, T-window reliability can be calculated as

$$p_{app}(d) = 1 - (1 - p_{Rx}(d))^{T/t_{Tx}} \tag{2.1}$$

where d is the sender-receiver distance, p_{Rx} is the PDR, T is the defined time window and t_{Tx} is the beacon transmission interval.

Since the publication of these original two papers in 2006, the concepts of IRT and T-window reliability have been taken up and extended by a number of other authors.

[22] For example, in Germany 409 crashes with injuries or fatalities occurred per 1 billion vehicle kilometers traveled in 2010 [Source: Statistisches Bundesamt (Federal Statistical Office)].

[23] A similar concept to IRT was introduced as *delivery delay* in [89], but not discussed in detail.

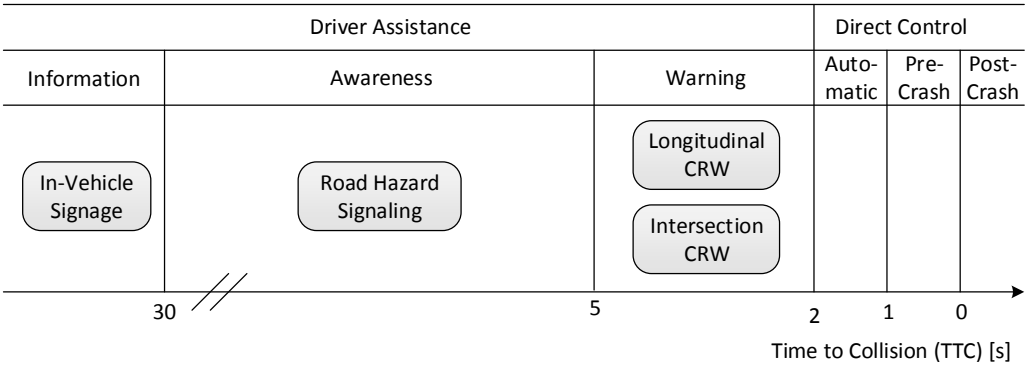


Figure 2.7: ETSI safety applications and corresponding time horizons [64]

In [92], Mittag et al. studied T-window reliability with $T = 1\,s$. In [93], Sepulcre et al. introduced the concept of a *critical distance* between two vehicles before which the T-window reliability requirement has to be fulfilled. In [94], An et al. generalized T-window reliability into the probability of receiving at least n packets within time window T. Additionally, they described the concept of an *awareness range*, i.e., the range up to which the desired reliability is fulfilled. In [95], Schmidt et al. introduced a distance-dependent degradation of T. Other works studied the inverse of T-window reliability, i.e., the probability that *no* packet was received within the time window T, e.g., [96], [97] and [98]. In [99], Kaul et al. introduced a metric called *system age* which is defined as the "average end-to-end delay observed in any vehicle's state within a certain cluster of nodes". While this metric is initially based on IRT, it provides a system-level view on the up-to-dateness of awareness information. An overview of IRT-related awareness metrics is presented in Table 2.7 on the facing page.

2.3 Scalability and Channel Congestion

Channel congestion control has been a major research topic in the ITS community, as we will discuss in detail in Chapter 3 on page 45. A simple back-of-the-envelope calculation illustrates why (cf. [71] and [19] for similar considerations). Let us assume that each vehicle transmits a beacon of 400 Bytes at a rate of 10 Hz and a data rate of 6 Mbit/s, cf. Section 2.1.3 on page 18. Then, it takes about 0.53 ms to transmit one packet[24] and the channel can hold a maximum of 1875 packets per second if all participants are coordinated perfectly. The channel can thus support a maximum of 187 vehicles under the given assumptions.

Let us now consider how many vehicles can realistically be expected on a road segment. According to [100], a density of 28 vehicles/km/lane is the "maximum density at which sustained flows at capacity are expected to occur" on a freeway, i.e., before a

[24] Note that the IEEE 802.11p header is always transmitted at the lowest available data rate of 3 Mbit/s. Thus, it would actually take longer to transmit the packet.

Metric	Description
T-window reliability [7]	Probability of receiving at least one packet within time window T
Neighborhood awareness [92]	Probability of having received at least one packet within the last second
Application reliability [93]	Probability of receiving at least one packet within time window T and distance d
Awareness probability [94]	Probability of receiving at least n packets within time window T
Awareness range [94]	Maximum distance at which the awareness probability is greater than or equal to a certain threshold
Awareness [95]	Ratio of neighbors whose last packet was received within a distance-dependent packet lifetime
Invisible neighbor [96]	A neighbor within the application's Region of Interest (ROI) from whom no message has been received within time window T
Blackout probability [97]	Probability that IRT exceeds 1 second
Update delay [98]	Complementary Cumulative Distribution Function (CCDF) of IRT
System age [99]	Average IRT over a certain time interval, averaged over a vehicle's neighbors, averaged over all vehicles.

Table 2.7: Awareness metrics related to Packet Inter-Reception Time (IRT)

traffic jam starts to form. The default value for a traffic jam density is specified as 120 vehicles/km/lane [100]. If we assume a freeway with three lanes per direction, we get a total density between 168 vehicles/km in the former and 720 vehicles/km in the latter case. The number of vehicles sharing the channel at a particular location depends on the applied transmit power and carrier sense threshold, cf. Section 2.1.2 on page 13. Let us assume a carrier sense range of 500 m. Since each vehicle can get triggered into channel busy mode by vehicles ahead and behind itself on the road, it shares the channel with all vehicles within 1 km, i.e., between 168 and 720 vehicles in our example. We can thus see that, even under non-extreme assumptions, the communication demand from CAMs/BSMs alone may exceed the available channel capacity. In addition, periodic CAMs and BSMs are expected to share the channel with other message types as well, in particular DENMs and event-driven BSMs, respectively,

cf. Section 2.1.3 on page 18. These messages contain safety-of-life critical information which has to be communicated with the highest priority. Thus, the objective of channel congestion control for VSC is typically to reserve a certain fraction of the available bandwidth for such emergency messages [1][20].

It should be noted that an uncontrolled rebroadcasting of event-driven messages may also congest the channel, a phenomenon known as the *broadcast storm* problem [101]. That is, there are actually two kinds of channel congestion control in VSC, one addressing the adaptation of the transmission parameters of single-hop CAM/BSM messages, which is the focus of this thesis, and one aiming at limiting unnecessary retransmissions of multi-hop emergency messages, e.g., [20], [102] and [103], which is a research topic of its own.

2.3.1 Degrees of Freedom

While the number of vehicles typically cannot be influenced, the amount of data injected into the network can be controlled by adapting the following parameters[25]:

1. *Data rate*, i.e., the number of bits conveyed per second.
2. *Transmit power*, i.e., the signal power emitted by the (isotropic) antenna.
3. *Packet size*, i.e., the total number of bytes transmitted over the air.
4. *Transmission rate*, i.e., the number of packets generated per second.

In addition, the temporal and spatial reuse of the channel as well as the likelihood of packet collisions are influenced by:

5. *Contention window*, i.e., the maximum number of backoff slots.
6. *Carrier sense threshold*, i.e., the energy level at the antenna above which the medium is considered busy.

Table 2.8 on the next page illustrates the ISO/OSI layer associated with each of the mentioned parameters as well as their corresponding range of allowed values as far as specified by different standards. We can see that a wide range of values is applicable to each degree of freedom.

2.3.2 Transmission Range, Carrier Sense Range and Interference Range

Out of the six mentioned degrees of freedom, the spatial reuse of the shared medium is determined by the combination of transmit power, data rate and carrier sense

[25] Note that packet size and data rate determine the *transmission time* of a packet over the channel.

Parameter	Layer	Value range
Data rate	PHY	{3, 4.5, 6, 9, 12, 18, 24, 27} Mbit/s [42]
Transmit power	PHY	≤33 dBm [35][42]
Contention window	MAC	[15;1023] for class AC_BK and AC_BE [42]
Carrier sense threshold	MAC	≤-65 dBm [42]
Packet size	APP[a]	≤~2342 Bytes[b]
Transmission rate	APP	[1 Hz;10 Hz] [62], ≤10 Hz [6]

Table 2.8: Degrees of freedom influencing the scalability of VSC and their allowed value ranges

[a] The size of lower layer headers and trailers is typically fixed. Thus, we expect the main influencing factor to be application payload. For scalability, the total number of bytes is decisive.

[b] The given value is the maximum size to be transmitted without segmentation, i.e., 2304 Bytes for the maximum frame body size of the 802.11p MAC (assuming no extension for security), 32 Bytes for MAC header and trailer and ~6 Bytes for PHY overhead [53].

threshold. In this context, three distances are typically of interest in an IEEE 802.11 network [104]:

1. *Transmission range* d_{Tx}, i.e., the range inside which a packet can be successfully received;
2. *Carrier sense range* d_{cs}, i.e., the range inside which nodes are able to sense the signal, even though they may not be able to decode it;
3. *Interference range* d_{int}, i.e., the range inside which a new transmission may interfere with the original transmission.

In the following, we discuss each of the three distances as well as the trade-offs in adapting each of the three degrees of freedom. For presentation clarity, we abstract from fading effects which are encountered in a real-world scenario and assume a deterministic degradation of transmit power over distance. We calculate path loss according to the power-law model used later in this thesis, cf. Annex A.1 on page 223.

We model interference as a *cumulative noise* at the receiver, i.e., the total interference and noise level corresponds to the sum of the Received Signal Strength (RSS) values of all overlapping transmissions and the noise floor, which we set to -99 dBm. For a successful reception, a certain Signal to Interference and Noise Ratio (SINR) has to be fulfilled throughout the duration of the packet. This SINR-based model corresponds to the *physical model* of Gupta and Kumar [105] and is also employed in the network simulator ns-2 after the overhaul of Chen et al. [106].

Transmission Range

The last column of Table 2.1 on page 14 lists the required SINR associated with each data rate in IEEE 802.11p, according to measurements by Jiang et al. [43]. A higher data rate has the advantage that it takes a shorter time to transmit a packet of a fixed size. Thus, the channel is busy for a shorter amount of time and more packets can be transmitted. However, a higher data rate also requires a higher SINR, which decreases the communication range at a fixed transmit power, as we discuss in the following. This trade-off is illustrated by Figure 2.8 on the facing page, which depicts the corresponding communication ranges without interference for each data rate compared to the maximum number of packets per second the channel can hold under the given assumptions. We can see that a high data rate significantly reduces communication range.

The interference-free deterministic communication range according to the power-law model can be calculated as

$$d_{Tx} = 10^{(P_{Tx} - P_{RxTh} - L)/10\gamma} \tag{2.2}$$

where P_{Tx} denotes the transmit power in dBm, P_{RxTh} denotes the minimum RSS required for reception in dBm, L denotes the reference loss in dB and γ is the path loss exponent.

Figure 2.9a on page 38 illustrates the resulting RSS with respect to sender-receiver distance for a transmission using 12 dBm in our configuration. The figure illustrates the resulting communication ranges for three different data rates, i.e., 318 m for 3 Mbit/s, 219 m for 6 Mbit/s and 91 m for 12 Mbit/s. That is, in order to maintain the same communication range, transmit power has to be increased accordingly. Figure 2.9b on page 38 illustrates that in our example, a transmit power of 15 dBm and 22 dBm is required for 6 Mbit/s and 12 Mbit/s, respectively, to achieve the same communication range as the combination of 12 dBm and 3 Mbit/s.

Figure 2.9c on page 38 illustrates the impact of an interfering transmission on the resulting communication range in our continued example. Note that according to the cumulative noise model, the interfering signal consists of the sum of the interferer's RSS and the noise floor. We observe that the communication ranges for 3 Mbit/s, 6 Mbit/s and 12 Mbit/s have been reduced significantly, i.e., to 156 m, 116 m and 54 m, respectively.

Interference Range

The previous example illustrates that a simultaneous transmission may significantly reduce the transmission range. The degree of interference thereby depends on the distances between sender, receiver and interferer. For uniform transmit power, the relationship between sender-receiver distance d_{sr} and interference range d_{int}, i.e., the distance between receiver and interferer, is typically defined as

$$P_{Rx}(d_{sr}) - P_{Rx}(d_{int}) = S_o \Leftrightarrow d_{int} = d_{sr} 10^{S_o/10\gamma} \tag{2.3}$$

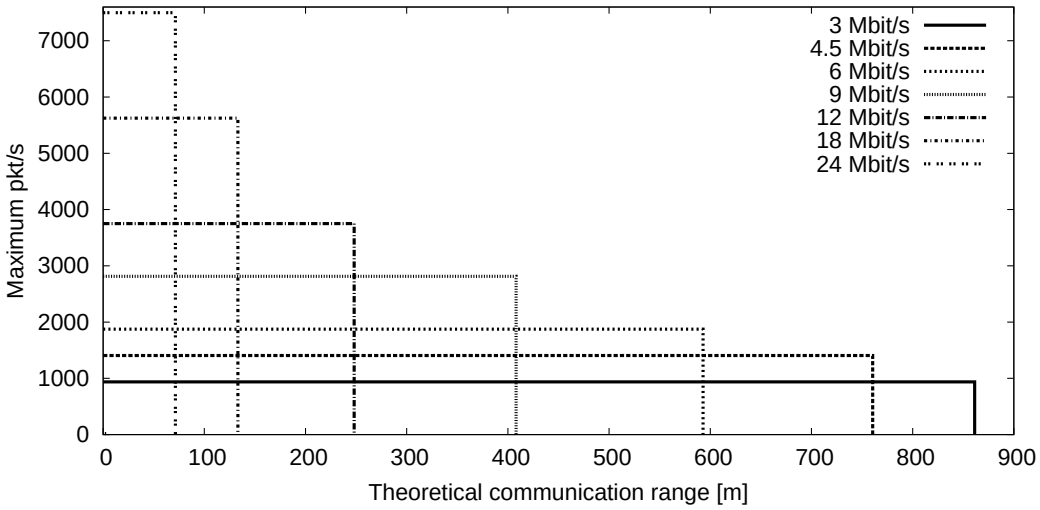


Figure 2.8: Trade-off between communication range and the maximum number of packets/s for different data rates (using $P_{Tx} = 20$ dBm)

where γ is the path loss exponent and S_o is the required Signal to Interference Ratio (SIR) in dB [107][108][109]. Note that this definition does not take into account the noise level and only compares the RSS values at the receiver.

If we take noise into account, we get

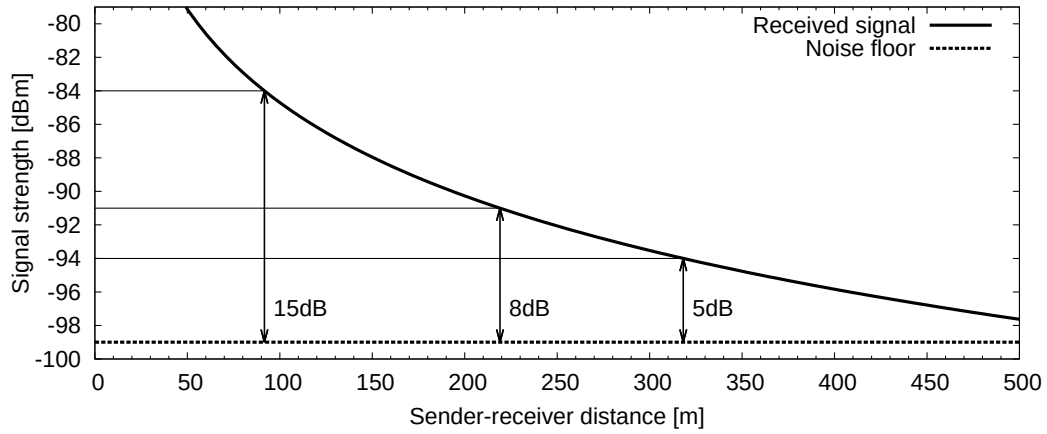
$$P_{Rx}(d_{sr}) - (P_{Rx}(d_{int}) + N_o) = S_N \Leftrightarrow$$
$$d_{int} = \left(-\frac{10^{-P_{Tx}/10} d_{sr}^{-n} \left(S_N d_{sr}^{\gamma} 10^{(N_o+L)/10} - 10^{P_{Tx}/10} \right)}{S_N} \right)^{-1/\gamma} \tag{2.4}$$

based on the considered power-law model, where N_o is the noise floor in dBm and S_N is the SINR required for reception. Note that when taking noise into account, it is not possible to achieve an SINR of exactly S_N at d_{Tx} in the presence of interference, since the interference signal theoretically propagates infinitely, cf. [110].

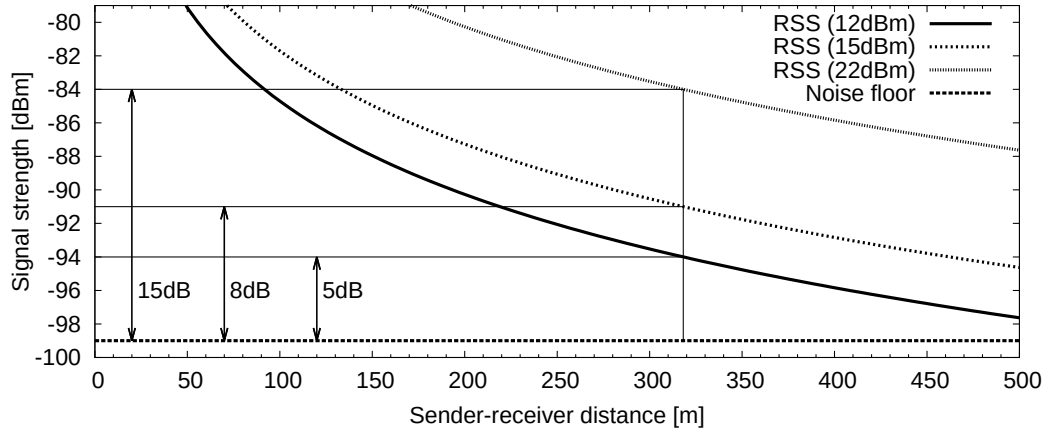
If we set $d_{sr} = d_{Tx}$, then the minimum distance between two transmitters which does not violate the SIR criterion can be calculated as

$$d_{min} = d_{Tx} + d_{int} \tag{2.5}$$

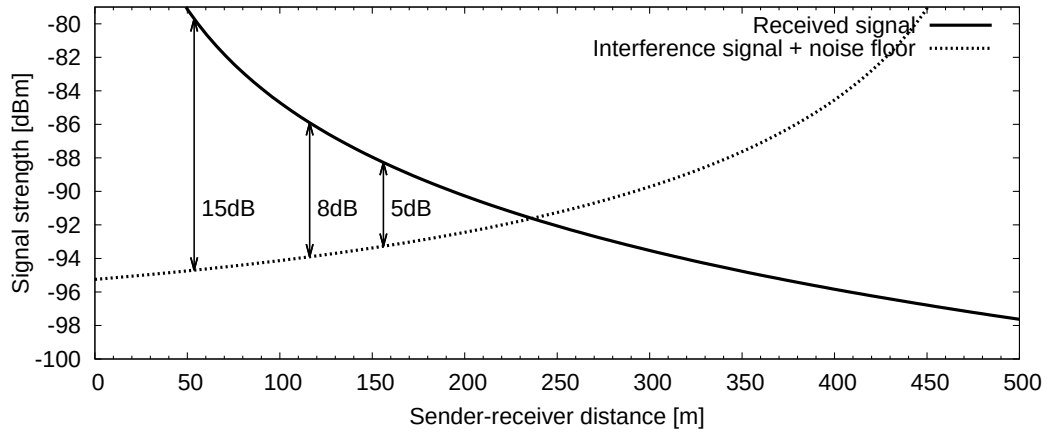
Figure 2.10 on page 39 illustrates the difference between both models in our continued example. The distance between sender and interferer was set to 812 m according to Equation (2.5) for a transmit power of $P_{Tx} = 12$ dBm. We can see that while the resulting SIR at the nominal communication range without interference is 8 dB, the resulting SINR is approximately 5 dB. That is, while the SIR model results in a communication range of 219 m, i.e., the communication range without interference, the SINR model results in a communication range of 158 m.



(a) Communication ranges (without interference) for different data rates



(b) Combinations of transmit power and data rate resulting in the same communication range



(c) Communication ranges (with interference)

Figure 2.9: Relationship of data rate, interference and communication range (example using $P_{Tx} = 12$ dBm under the given assumptions)

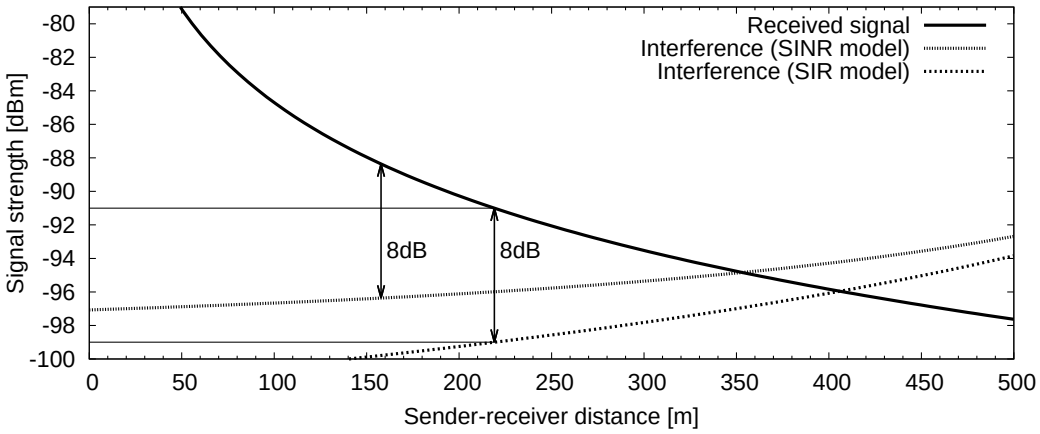


Figure 2.10: Interference range according to the SIR and SINR models

Carrier Sense Range

The carrier sense threshold determines the energy level above which a station considers the medium to be busy, cf. Section 2.1.2 on page 13. In combination with the transmit power, it determines how many simultaneous transmissions can occur in the network. In some works, the carrier sense range is thus defined as the minimum distance between two concurrent transmitters [111]. In other words, it is the distance up to which the RSS of a signal triggers a receiving radio into channel busy mode.

Without taking noise into account, the carrier sense range according to the power-law model can be calculated as

$$d_{cs} = 10^{(P_{Tx} - P_{cs} - L)/10\gamma} \tag{2.6}$$

where P_{Tx} is the transmit power in dBm, P_{cs} denotes the carrier sense threshold in dBm and L is the reference loss in dB.

To calculate the carrier sense range according to the cumulative noise model, we first need to calculate the difference between carrier sense threshold and noise floor in dB as

$$P_{csN} = 10log_{10}(10^{P_{cs}/10} - 10^{N_0/10}) \tag{2.7}$$

where N_0 is the noise floor in dBm. Then, we can apply P_{csN} to Equation (2.6) and get

$$d_{cs} = 10^{(P_{Tx} - P_{csN} - L)/10\gamma} \tag{2.8}$$

Figure 2.11 on the following page illustrates the resulting carrier sense ranges for three exemplary carrier sense thresholds according to the SIR model (received signal) and the SINR model (received signal + noise floor) in our continued example. We can see that for high carrier sense thresholds, e.g., 85 dBm, the difference between both models in terms of the resulting carrier sense range is negligible. For low carrier sense thresholds, e.g., -95 dBm however, we observe a significant difference. In the example,

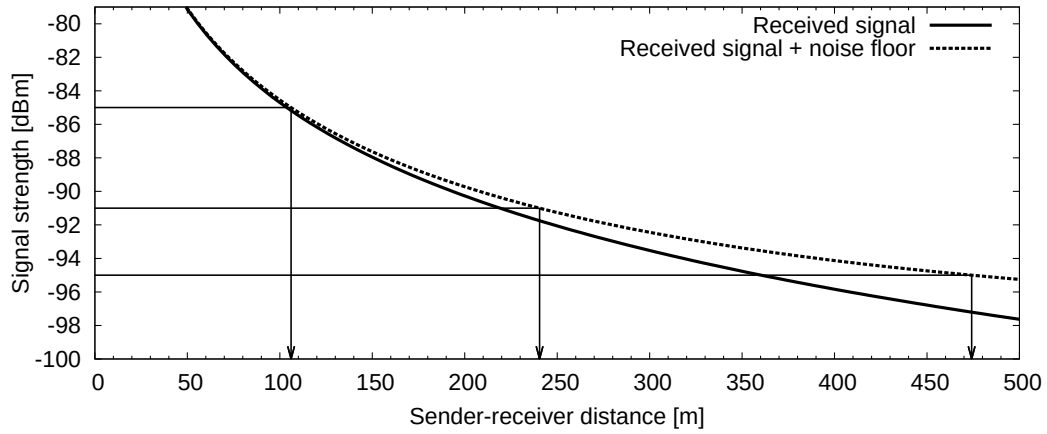


Figure 2.11: Exemplary deterministic carrier sense ranges for carrier sense thresholds -85 dBm, -91 dBm and -95 dBm

the SIR model results in d_{cs} = 360 m, while the SINR model results in d_{cs} =474 m. For P_{cs} = -85 dBm, we get d_{cs} = 104 m and d_{cs} = 106 m, respectively.

Comparing both results for the SINR model, the higher carrier sense threshold thus allows for more than four times as many concurrent transmissions compared to a carrier sense threshold of -95 dBm. However, as we have seen in Figure 2.9c on page 38, the resulting lowered sender-interferer distance results in a reduced communication range.

2.3.3 Metrics for Channel Load

In the literature, different metrics have been suggested to quantify channel load. In the following, we discuss two model-based metrics which are closely related to each other as well as one metric which can be directly measured by IEEE 802.11 hardware and is thus widely used.

Communication Density (CD) and Beaconing Load (BL)

In [112], Jiang et al. introduce Communication Density (CD) as "the number of carrier sensible transmissions per unit of time and per unit of road". The authors state that the CD is "proportional to, and therefore can be effectively calculated as"

$$CD = \rho \cdot r \cdot d_{Tx} \tag{2.9}$$

where ρ is the vehicle density in vehicles/km, r is the transmission rate in Hz and d_{Tx} is the communication range in km. By means of simulations, the authors show that scenarios with the same CD, e.g., a scenario with vehicle density ρ and transmission rate r and a second scenario with vehicle density $^1/_2\rho$ and transmission rate $2r$, experience nearly the same broadcast performance in terms of PDR over distance (for a fixed transmission power) and Channel Access Time (CAT) distribution. Another finding

of the authors is that "the total communication density of a system can be summed up from its subgroups' communication densities". For example, if one direction of a highway has a high CD of, e.g., 1000 messages/s due to a traffic jam and the other direction has a low CD of, e.g., 500 messages/s due to free-flow conditions, the road segment has a total CD of 1500 messages/s.

A similar concept as CD is introduced in [113], where Torrent Moreno defines Beaconing Load (BL) as "the average amount of load offered to the channel within a node's [carrier sense range]". The metric is based on a linear road topology and is calculated as

$$BL = \rho \cdot r \cdot 2d_{cs} \cdot M \tag{2.10}$$

where d_{cs} and M denote the carrier sense range in km and the message size in bytes, respectively. It is interesting to note that Equations (2.9) and (2.10) share two common factors, i.e., ρ and r. Since the CD is meant to be *proportional* to Equation (2.9), d_{Tx} can be substituted with $2d_{cs}$ for a fixed carrier sense threshold without loss of generality. Thus, in this thesis we define communication density as

$$CD = \rho \cdot r \cdot 2d_{cs} \tag{2.11}$$

where d_{cs} is calculated according to Equation (2.8).

Table 2.9 on the following page illustrates the resulting CD values for different vehicle densities, transmit power levels and transmission rates on a highway with three lanes per direction. The table further shows in brackets the corresponding BL values based on a packet size of 400 Bytes. The carrier sense range is calculated for a carrier sense threshold of -95 dBm according to the power-law model as described in Section A.1 on page 223. The CD and BL values exceeding a data rate of 6 Mbit/s are highlighted in gray shade. Note that the assumed maximum BL of 6 Mbit/s corresponds to a CD of 1875 messages/s, cf. the back-of-the-envelope calculation above. We can see in the table that for 5 vehicles/km/lane, the available channel capacity is not exceeded for any combination of transmit power and transmission rate. For 25 vehicles/km/lane and 50 vehicles/km/lane, however, at least one of the parameters has to be adapted in order not to overload the channel.

Channel Busy Ratio (CBR)

While CD and BL are useful tools for theoretical analysis, they are difficult to measure in a real-world environment. In practice, the Channel Busy Ratio (CBR), i.e., the fraction of time a radio perceived the channel as busy, is frequently used as a channel load metric. In the IEEE 802.11 standard, Channel Busy Ratio (CBR)[26] is defined as

$$U = \left\lfloor \frac{t_{busy}}{t_{total}} \times 255 \right\rfloor \tag{2.12}$$

[26] The term CBR is not used in the standard. Instead, the concept is denoted as *Channel Utilization*, *CCA Busy Fraction* and *Channel Load* in different clauses of the document [42]. Other works also use the term Channel Busy Time (CBT) as a synonym for CBR, e.g., [72].

		r [Hz]		
ρ [veh/km/ln]	P_{Tx} [dBm]	1	5	10
5	10	11.1 (0.04)	55.5 (0.18)	110.9 (0.35)
	20	154.0 (0.49)	192.5 (0.62)	385.0 (1.23)
	30	231.0 (0.74)	668.3 (2.14)	1336.7 (4.28)
25	10	55.5 (0.18)	277.3 (0.89)	554.5 (1.77)
	20	192.5 (0.62)	962.6 (3.08)	1 925.1 (6.16)
	30	668.3 (2.14)	3 341.7 (10.69)	6 683.4 (21.39)
50	10	110.9 (0.35)	554.5 (1.77)	1 109.0 (3.55)
	20	385.0 (1.23)	1 925.1 (6.16)	3 850.3 (12.32)
	30	1 336.7 (4.28)	6 683.4 (21.39)	13 366.9 (42.77)

Table 2.9: Exemplary Communication Density (CD) [msg/s] (Beaconing Load (BL) [Mbit/s]) values for different vehicle densities ρ, transmit power levels P_{Tx} and transmission rates r; combinations in grey cells exceed the assumed channel capacity of 6 Mbit/s.

where t_{busy} is the Channel Busy Time (CBT) in μs, i.e., the time "the [station] sensed the medium was busy, as indicated by either the physical or virtual carrier sense (CS) mechanism" and t_{total} is the total measured time in μs [42]. The linear scaling with 255 is used for the over-the-air transmission of channel information in feedback frames. Due to its definition in the IEEE 802.11 standard, CBR is widely available as a channel feedback in common hardware.

While the IEEE 802.11 standard requires the availability of CBR measurements, it does not specify how CBR is to be implemented. In [72], the ETSI standard specifies that CBR should be implemented by means of *channel probing*. That is, the CCA function (cf. Section 2.1.2 on page 13) is evaluated at fixed, uniform time intervals, resulting in a sequence of true (1) or false (0) values. When the CBR function is evoked, it returns the average of all samples collected within the measurement interval, i.e.,

$$U = \frac{\sum_{i=1}^{n} k_i}{n} \tag{2.13}$$

where n is the number of samples k_i and $k_i \in \{0, 1\}$.

Figure 2.12 on page 44 schematically illustrates a sample CBR calculation based on channel probing. In the figure, the measured signal strength at the antenna temporarily falls below the Carrier Sense Threshold (CSTh). During this IDLE time, the CCA function returns 0, else it returns 1. The dots along the CSTh line illustrate

discrete channel probing measurements. The figure further shows the calculation of CBR averaged over ten channel probes at three points in time, resulting in CBR values of 60%, 20% and 30%.

2.4 Summary of Characteristics and Challenges

Safety-oriented vehicular communication networks are expected to be dominated by single-hop broadcasts of CAM/BSM messages with the objective to establish a cooperative awareness among vehicles in order to avoid crashes and dangerous situations. With increasing vehicle density, the aggregate of these messages alone can exceed channel capacity, unless transmission parameters are adapted.

The *focus on cooperative awareness* distinguishes VSC from other communication networks, where typically the maximization of throughput is the optimization objective. While awareness is a receiver concept, it can only be controlled by the transmitter who needs to send out the right information at the right timing and with sufficient communication range.

Safety applications have *stringent timing requirements* which depend not only on the application itself but also on the current traffic situation. In other words, changes in the topology not only affect channel load, but also the optimization objective. The definition of safety applications' requirements is highly complex and a research topic of its own. In particular, if n vehicles run m safety applications each, up to $n \cdot m$ requirements have to be addressed by the system.

VSC takes place in a *pervasive broadcast environment* with unbounded signal propagation. The selected frequency band of 5.9 GHz comes with a high sensitivity of radio-wave propagation to changes in the environment, which makes it difficult to predict if a message can be received successfully. In addition, the broadcast nature of the foreseen communication makes message acknowledgments impractical. A particular challenge is thus how to set transmission parameters in order to achieve sufficient communication reliability over an unreliable wireless channel and without feedback on successful receptions.

For channel congestion control, *at least six degrees of freedom* have been identified, i.e., transmit power, data rate, carrier sense threshold, contention window, transmission rate and message size. Due to the high mobility of vehicles, the unbounded nature of radio-wave propagation, the infeasibility of central coordination, the lack of message acknowledgments and the real-time requirements of safety applications, designing channel congestion control for VSC is a challenging task and one of the major research topics in the field of V2X communication [1][19][29].

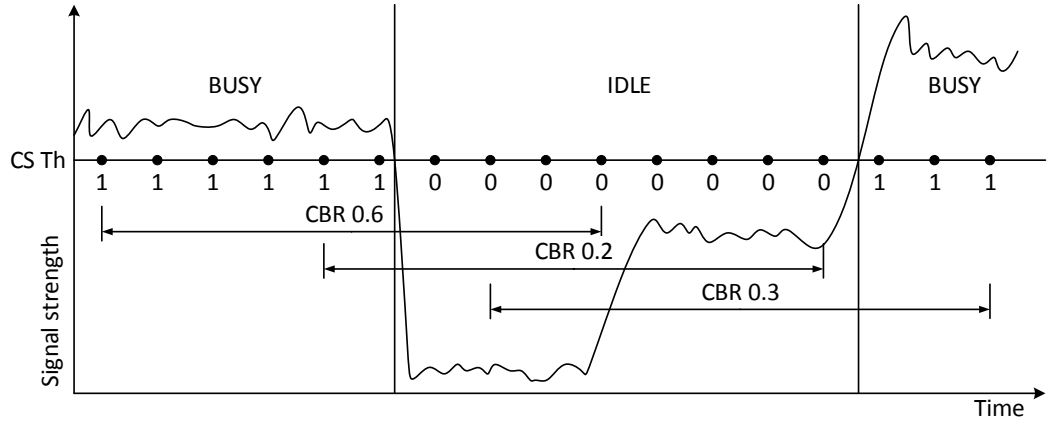


Figure 2.12: Schematic illustration of Channel Busy Ratio (CBR) calculation with channel probing

3

Congestion Control in Wired, Wireless and Vehicular Networks

The topic of congestion control plays a major role in many different aspects of today's communication infrastructure. The emergence of congestion control approaches began in the 1980s when packet switching networks, in particular the early Internet, became increasingly popular. In contrast to circuit switching networks such as telephone networks where a certain amount of bandwidth is reserved for each connection, packet switching networks use the available bandwidth more efficiently by allowing packets to take different routes through the network based on where capacity is available. However, while in circuit switched networks excess connections can be detected and refused, the dynamic bandwidth usage in packet switching networks can result in more packets being injected into the network than it can handle [114]. If this is the case, a drastic decrease in performance can be observed, known as a *congestion collapse*.

A famous series of such congestion collapses occurred in the early Internet of 1986, when a drop of three orders of magnitude in throughput (from 32 kbit/s to 40 bit/s) was observed. What had happened was that the original version of the Transmission Control Protocol (TCP) had reacted to packet loss at overflowing router queues by generating even more packets [22]. Today, about 25 years after the introduction of TCP congestion control by Jacobson and Karels in 1988 [22], TCP (along with its many flavors) is still the most prominent representative of congestion control protocols. However, a number of other approaches has emerged, custom tailored to the requirements of different applications and network types.

In the first part of this chapter, we review fundamental concepts of congestion control and existing taxonomies. Then, we briefly introduce the basic principles of TCP to illustrate the operation of Additive Increase Multiplicative Decrease (AIMD), the control mechanism also employed in Periodically Updated Load Sensitive Adaptive Rate control (PULSAR), the congestion control protocol developed in this thesis. In the next step, we discuss congestion control approaches in other wireless networks, i.e., cellular networks, Wireless Sensor Networks (WSNs) and Mobile Ad-Hoc Networks (MANETs). The last part of this chapter is dedicated to the review of related approaches in vehicular networks. We have structured the review according to the six control dimensions identified in the previous chapter.

3.1 Basic Concepts

In this section, we introduce basic concepts of congestion control based on pioneering works which have had a significant influence on research in this field. These concepts are independent of concrete algorithms and will be used later in this thesis.

3.1.1 Congestion Avoidance and Control

In [115], Jain and Ramakrishnan distinguish between

- *Flow control*, i.e., a scheme which protects the receiver of a flow of packets from being overwhelmed by too many packets sent from the source.
- *Congestion control*, i.e., a scheme which "protects the network from being flooded by its users (source and/or destination)".
- *Congestion avoidance*, i.e., a scheme which tries to keep the network at its optimal operation point.

While flow control concerns only one source-receiver pair and has the objective to prevent the buffer at the receiver from overflowing, congestion control and congestion avoidance address the input of any node into the network.

Figure 3.1 on the facing page illustrates the relationship between latter two concepts[115]. The figure shows a network's throughput, response time and power as a function of load, where *power* is the ratio of throughput and response time. While throughput first increases with increasing load, a saturation point is reached when the load approaches network capacity. If network load is further increased, throughput sharply drops to zero, a point the authors denote as *congestion collapse*. At this point, buffers overflow throughout the network, preventing any packets from coming through. The point where throughput breaks in abruptly is denoted as a *cliff*. The objective

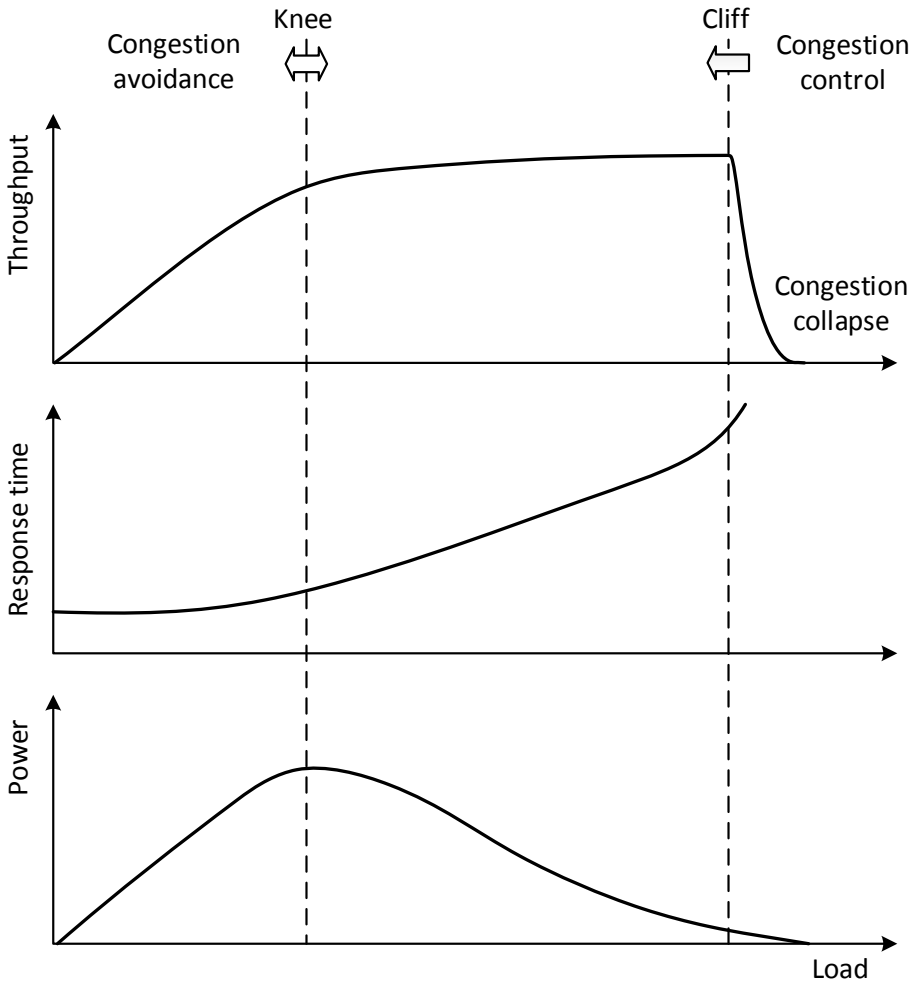


Figure 3.1: Congestion avoidance and control, adaptation of Figure 1 in [115]

of congestion control is to prevent the network from passing the cliff and help the network get back into operational mode if a congestion collapse occurs.

The network's response time also increases with load, since queues are building up. It approaches infinity after the cliff has been passed. It is easy to see that the objectives of minimizing response time and maximizing throughput are contradictory. The trade-off between the two metrics is depicted by the power curve, which shows a maximum before the cliff. This point, after which throughput increases little but response time increases significantly, is referred to as a *knee* by the authors. The objective of congestion avoidance is to maintain network load near the knee.

Basic Components of a Congestion Avoidance Scheme

In [115], Jain and Ramakrishnan describe the basic components of a congestion avoidance scheme, structured into a *feedback* and a *control* part.

The feedback part consists of

- *Congestion detection*, determining the congestion state based on, e.g., processor utilization, link utilization for queue lengths.
- *Feedback filter*, i.e., a low-pass filter to "to pass only those [congestion] states that are expected to last long enough for the user action to be meaningful.".
- *Feedback selector*, selecting the set of users to be notified of network congestion.

The control part consists of

- *Signal filter*, accumulating and interpreting the feedback received from the network, which may be contradictory.
- *Decision function*, deciding which action to take, i.e., to increase or to decrease, taking into account feedback signals for the last T seconds.
- *Increase/decrease algorithm*, determining the amount of the change in the determined direction.

3.1.2 Fairness and Efficiency

Jain et al. define an *optimal allocation* as one which is "100% fair and 100% efficient" [116]. *Efficiency* is defined as the maximization of *resource power*, which is given by

$$\text{Power} = \frac{\text{Throughput}^{\alpha}}{\text{Response time}} \tag{3.1}$$

where α is a weighting factor and set to 1 for equal weight of the two factors.

According to [116], efficiency is maximized at the knee. The efficiency of other allocations can be quantified using

$$\text{Efficiency} = \frac{\text{Power}}{\text{Power at knee}} \tag{3.2}$$

To quantify fairness, the authors introduced a *fairness index* given by

$$\text{Fairness} = \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} x_i^2} \tag{3.3}$$

where x_i originally denoted the throughput of user i [115]. In this interpretation, the equation assumes that maximal fairness is given for equal allocations to all users. In later works, x_i was defined as a_i/A_i^*, where a_i is the allocation of user i and A_i^* is the maximally fair allocation for user i [116]. While A_i^* could be set equal for all users, this definition allows to calculate the fairness index also for non-homogeneous allocation demands.

Max-Min Fairness and Proportional Fairness

In Jain's fairness index, A_i^* could be calculated according to different fairness concepts [116]. Two examples commonly encountered in computer networks are *max-min fairness* [117] and *proportional fairness* [118]. A max-min fair rate allocation is one for which "an increase of any rate within the domain of feasible allocations must be at the cost of a decrease of some already smaller rate"[119]. In other words, max-min fairness maximizes the minimum allocation. Note that this definition is strict and also excludes infinitesimally small changes in rate allocation.

Figure 3.2 on the next page illustrates a simple network with seven nodes, A to H, and four data flows, one each for AE and FG and two for GH, with rate allocations x_0, x_1 and x_2, respectively. The two bottleneck links BC and CD have a capacity of c each. To find the max-min allocation, the most heavily used bottleneck is addressed first, link CD in our case. Since three flows share this link, each is allocated $1/3c$, i.e., $x_0 = 1/3c$ and $x_3 = 1/3c$. The flow FG is allocated the same share of resources, i.e., $x_1 = 1/3c$. Since this allocation results in an underutilization of link BC, FG can increase its share to $x_1 = 2/3c$ without impacting the other flows.

In the above example, the total throughput of the network is $x_0 + x_1 + 2x_2 = 5/3c$, which is less than the available capacity of $2c$. An allocation which makes more use of the available bandwidth is, e.g., $x_0 = 1/4c$, $x_1 = 3/4c$ and $x_2 = 3/8c$, resulting in a throughput of $7/4c$. We observe that the flow AE gave up some bandwidth for the benefit of flows FG and GH. In this allocation, any change resulting in a feasible allocation would result in a negative *average change*, which is the definition of *proportional fairness* [119].

While proportional fairness makes better use of the available bandwidth, max-min fairness is considered more fitting for safety communications, as we will discuss in Section 5.3 on page 119.

3.1.3 Binary Adaptation Schemes

In [120], Chiu and Jain analyze binary linear control algorithms for congestion avoidance, i.e., each user's allocation x_i is adapted as

$$x_i(t+1) = \begin{cases} a_I + b_I x_i(t) & \text{if } y(t) = 0 \\ a_D + b_D x_i(t) & \text{if } y(t) = 1 \end{cases} \tag{3.4}$$

where $y(t)$ is a binary function of time t indicating increase (0) or decrease (1), a_I and b_I are increase parameters and a_D and b_D are decrease parameters.

In particular, the authors focus on the following four control functions:

- Additive Increase Additive Decrease (AIAD), i.e., $b_I = b_D = 0$
- Additive Increase Multiplicative Decrease (AIMD), i.e., $b_I = a_D = 0$

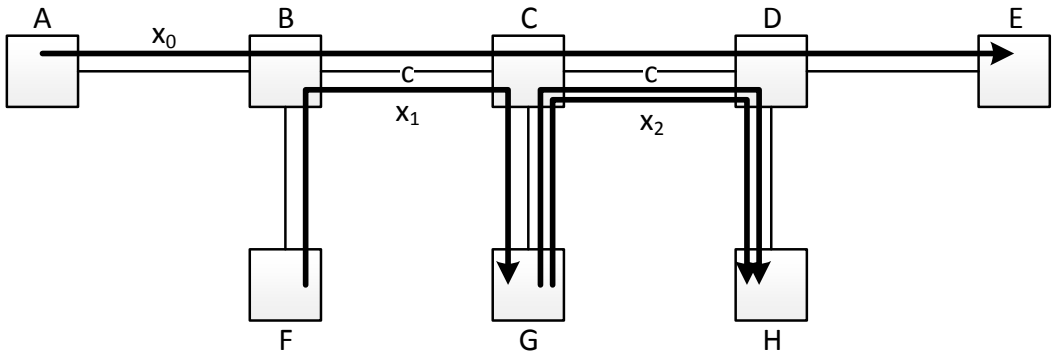


Figure 3.2: Example scenario based on [117][119] to illustrate different fairness concepts

- Multiplicative Increase Additive Decrease (MIAD), i.e., $a_I = b_D = 0$
- Multiplicative Increase Multiplicative Decrease (MIMD), i.e., $a_I = a_D = 0$

They show analytically that a distributed control has to fulfill the following requirements in order to converge to fairness and efficiency[1] the latter two concepts defined as explained above: $a_I > 0$, $b_I \geq 1$, $a_D = 0$ and $0 \leq b_D < 1$ [120]. In other words, the decrement should be multiplicative and the increment should be additive with an optional multiplicative component. The only one of the four considered control functions which fulfills these requirements is AIMD[2].

To illustrate their analysis, the authors use a vector representation of user's allocations x_i which is illustrated by Figure 3.3 on page 52 for the case of two users. Optimal efficiency is achieved if $x_1 + x_2 = c$, where c is the capacity at the knee, while optimal fairness is achieved for $x_1 = x_2$. The two optimality lines intersect in the optimal point which has 100% fairness and 100% efficiency. To the left of the optimal efficiency line, the network is unterutilized, while it is overutilized to the right. Lines parallel to both optimal lines represent equal levels of fairness and efficiency. That is, the points p_2 and p_3 have the same efficiency, while p_3 and p_4 have the same fairness. The arrows p_0p_1 and p_0p_2 illustrate the effect of a multiplicative and additive allocation change, respectively. Since both types of adaptation affect the allocations x_1 and x_2 simultaneously, a multiplicative adaptation takes place on a linear line between the origin and the starting point, while an additive adaptation corresponds to a shift on a 45° line. Using this representation, we can see three things [120]:

- Neither pure additive nor pure multiplicative adaptations can result in convergence to optimal fairness and efficiency, since the system oscillates on the respective linear lines.

[1] The authors further assume that there is no truncation, i.e., users do not check if the new allocation is out of the bounds of a feasible total allocation.

[2] In the remainder of this thesis, we use the symbols $\alpha_I = a_I$ and $\beta_D = 1 - b_D$ for the additive increase and multiplicative decrease parameters of AIMD, respectively. This notation is better suited to describe the *target rate* concept, cf. Section 6.3.3 on page 149.

- Multiplicative increments lead *away* from the optimal point, while multiplicative decrements lead *toward* the optimum.
- Additive increments are needed to move toward the optimum after a multiplicative decrement.

Responsiveness and Smoothness

As further analyzed in [120], the parameter choice of the linear control determines its convergence time to efficiency, denoted as *responsiveness*, as well as the amplitude of oscillation, denoted as *smoothness*. The two concepts are illustrated by Figure 3.4 on page 53, which is an adaptation of Figure 3 in [120].

For a linear control $x_i(t+1) = a + bx_i(t)$ with n users, the responsiveness is given by

$$t_e = \frac{log\left[(an + (b-1)X_{goal})/(an + (b-1)X_o)\right]}{log(b)} \tag{3.5}$$

where X_{goal} is the load level at the knee and X_o is the initial state [120].

Further, the smoothness in terms of maximum overshoot is given by

$$s_e = |an + (b-1)X_{goal}| \tag{3.6}$$

While the responsiveness t_e is a monotonically decreasing function of a and b, the smoothness s_e is a monotonically increasing function of both parameters [120]. Thus, the parameter choice is a trade-off between both objectives.

It should be noted that the analysis by Jain et al. has some limitations. First, it is based on a synchronized case in which all users make their adjustments at the same points in time. In [121], Gorinsky and Vin show that in some cases AIMD may not converge to fairness if adjustments are made asynchronously. We will get back to this issue in Section 6.3.2 on page 144. Second, it is assumed in [120] that all users receive the same feedback from the network and that the feedback is not delayed. Without these assumptions, AIMD may also not necessarily converge to fairness [122].

3.1.4 Taxonomy According to Yang and Reddy

Due to the complexity and diversity of congestion control approaches and their characteristics, their classification is a non-trivial task and different taxonomies can be envisioned. A frequently cited taxonomy for congestion control approaches is presented by Yang and Reddy in [123], as illustrated by Figure 3.5 on page 54. Taking on a control theoretic perspective, the authors distinguish between *open loop* and *closed loop* control. In open loop control, actuators make adjustments based on local knowledge only. In contrast, closed loop control makes adjustments based on feedback from the network.

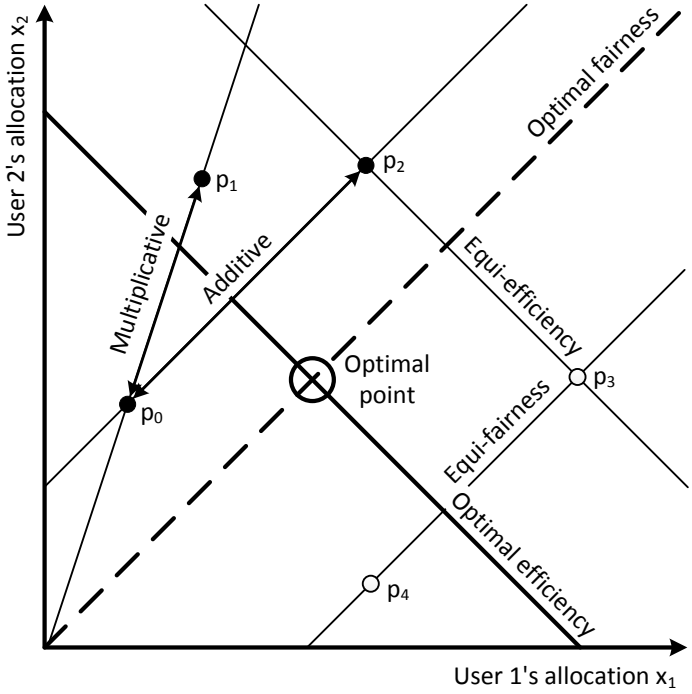


Figure 3.3: Vector representation of linear control algorithms according to [120]

Within open loop control, the authors distinguish between *source control* and *destination control*, depending on which end of the communication flow is targeted by the control. While the former tries to optimize the traffic flow through a switch or router, e.g., by allocating separate queues for each flow, the latter focuses on which packets to drop when buffers are full.

Within closed loop control, the authors distinguish between *implicit feedback* and *explicit feedback*. An example for implicit feedback control is the congestion control of TCP, which is based on measurements of round-trip delays, cf. Section 3.2 on the facing page. Within explicit feedback control, the authors further distinguish between *persistent feedback* and *responsive feedback*. An example for persistent explicit feedback is the *binary feedback scheme* introduced by Ramakrishnan and Jain, in which routers along the way can set a congestion flag in the packet header [124]. An example for responsive explicit feedback is a *choke packet*, sent back to the originating node by the receiving node or an intermediate router, asking it to reduce its rate [125]. Finally, the authors distinguish *local feedback* and *global feedback*, of which the former is exchanged between adjacent nodes and the latter is transferred from destination to source.

According to the taxonomy of Yang and Reddy, PULSAR applies a closed loop control with persistent explicit feedback to the transmission rate, cf. Chapter 6 on page 135. Optionally, transmit power can be adapted using an open loop source control.

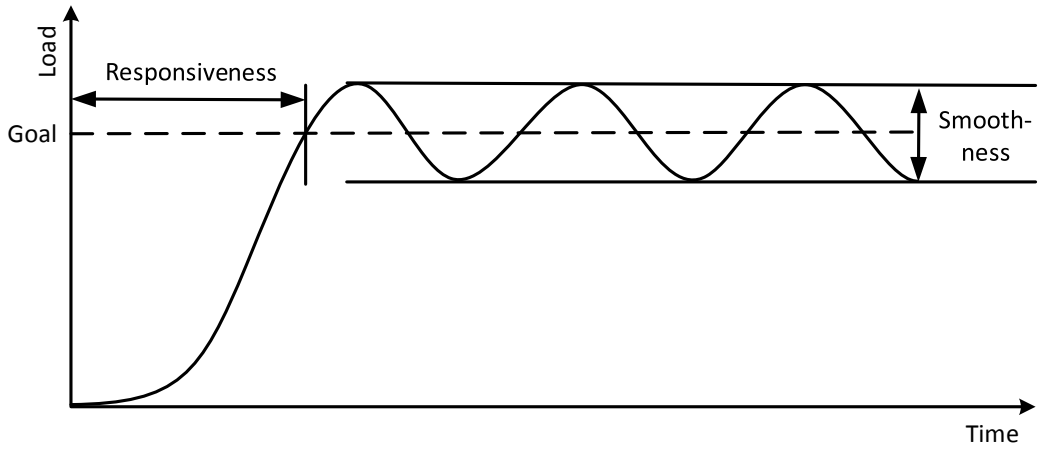


Figure 3.4: Smoothness and responsiveness according to [120]

3.2 Transmission Control Protocol (TCP)

In the Internet, millions of users share the available resources of links and routers. Since resource sharing and congestion control are two sides of the same coin [126], a significant amount of research effort has been dedicated to congestion control approaches to ensure the scalability of the Internet despite its continuously growing number of users. The most prominent protocol employing congestion avoidance and control mechanisms is the aforementioned Transmission Control Protocol (TCP). In the following, we briefly review congestion control in TCP and some of its many flavors. Note however, that within the scope of this thesis, we can only scratch on the surface of related work on this topic. Also, we do not discuss other mechanisms of TCP, e.g., connection management and flow control. For more information on TCP, please refer to networking textbooks, e.g., [49][114][126][127].

TCP is an end-to-end transport protocol which establishes a reliable connection-based service on top of an unreliable transport protocol, typically the well-known Internet Protocol (IP). It employs a host-centric, feedback-based and window-based congestion control scheme whose core is the adaptation of a variable called *congestion window*. The congestion window indicates how many bytes the node may transmit before an Acknowledgment (ACK) has been received and is measured in the unit of the Maximum Segment Size (MSS) of TCP. The adaptation is based on AIMD with $\alpha_I = 1$ and $\beta_D = 1/2$, cf. Section 3.1 on page 46.

3.2.1 Congestion Detection

TCP uses implicit feedback from the network to determine if congestion occurred. Since in a wired network such as the Internet transmission errors are rare, the main reason packets are lost is if they are dropped due to overflowing queues, i.e., if the network is congested. TCP thus uses packet loss, detected based on missing acknowl-

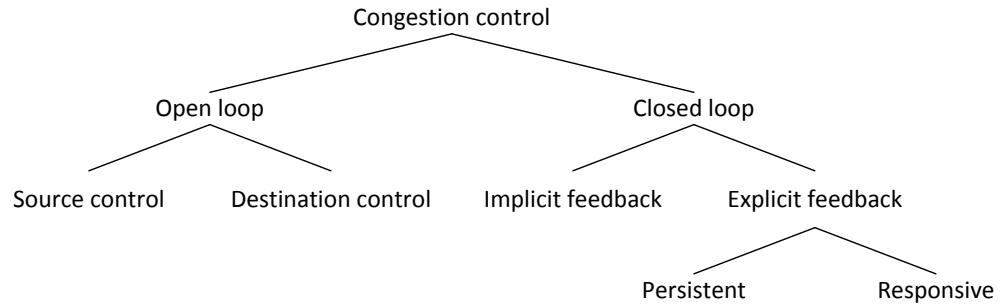


Figure 3.5: Taxonomy of congestion control algorithms according to [123]

edgments, as a congestion indicator. It uses a Retransmission Timeout (RTO) to decide if a packet has been lost and is to be retransmitted. The RTO is calculated using an Exponentially Weighted Moving Average (EWMA) of Round Trip Time (RTT) measurements, i.e., the time difference between the transmission of data and the reception of the corresponding ACK. In the original TCP algorithm[3], the RTO was calculated as

$$\text{RTO} = \beta \cdot \text{SRTT} \tag{3.7}$$

where β has a default value of 2 and the Smoothed RTT (SRTT) is given by

$$\text{SRTT} = \alpha \cdot \text{SRTT} + (1 - \alpha) \cdot \text{RTT} \tag{3.8}$$

where RTT is the new measurement and $\alpha \in [0, 1]$, with a recommended value of $\alpha \in [0.8, 0.9]$. This mechanism was later improved by Karn and Partridge, who excluded retransmitted packets from RTT averaging [128], and by Jacobson and Karels, who included RTT variance in the SRTT calculation [22]. Jacobson and Karels further introduced an *exponential backoff* mechanism, where the RTO is doubled for every retransmission.

TCP employs a second mechanism to detect packet loss, referred to as *fast retransmit*. When more than three *duplicate ACKs*, i.e., acknowledgments for data which has already been acknowledged before, have been received, TCP suspects that a packet has been lost and does not wait for the RTO to expire before retransmitting.

3.2.2 Rate Adaptation

The congestion window in TCP is adapted based on three mechanisms[4]:

- *Slow start*. At the beginning of a session or after a RTO event has occurred, TCP starts with a low congestion window, typically one MSS. The congestion window is increased exponentially, i.e., it is increased by one MSS for each

[3] IETF RFC 793, "Transmission Control Protocol - DARPA Internet Program Protocol Specification", available at *http://tools.ietf.org/html/rfc793*

[4] Based on RFC 5681, "TCP Congestion Control", available at *http://tools.ietf.org/html/rfc5681*

received ACK, until the slow start threshold *ssthresh* is reached. The slow start threshold is set to $min(F/2, 2\text{MSS})$ if congestion is detected based on RTO or three duplicate ACKs, where F is the *flight size*, i.e., the amount of data which has been sent but not yet been acknowledged[5].

– *Congestion avoidance*. After the slow start threshold has been reached, the congestion window is increased (approximately) linearly. That is, for each received ACK the congestion window is increased by $\text{MSS} \times (\text{MSS}/W)$, resulting in a total increment of approximately one MSS after all bytes in the congestion window have been acknowledged.

– *Fast recovery*. In contrast to a timeout, the reception of three duplicate ACKs means that some packets still got through the network. Thus, the fast recovery mechanism sets the congestion window to $min(F/2, 2\text{MSS})$ instead of resetting it to one[6]. Then, the algorithm continues with congestion avoidance. The fast recovery mechanism distinguishes two variants of TCP referred to as *Tahoe* and *Reno*, of which the latter employs the fast recovery mechanism [127].

Figure 3.6 on the next page illustrates the adaptation of the contention window for TCP Tahoe, resulting in a *saw tooth* shaped curve characteristic for TCP. The figure indicates slow start phases (white) and congestion avoidance phases (gray). At simulation times 1 s, 2.4 s, 5.6 s and 8.7 s, the slow start mechanism is invoked due to the reception of three duplicate ACKs. When the slow start threshold has been reached, TCP continues with congestion avoidance.

We present this example here in order to illustrate similarities and differences to the rate adaptation in PULSAR. Both protocols employ the AIMD mechanism, which leads to a similar shape of the rate adaptation curve over time, cf. Figure 3.6 on the following page and Figure 7.1 on page 169. However, while TCP is a window-based protocol with two rate control mechanisms, i.e., slow start and congestion avoidance, PULSAR adapts transmission rate directly and uses AIMD only. The intention of the slow start mechanism is to prevent an initial burst of data when the connection is starting up, which leads to packet loss and retransmissions [22]. In contrast, Cooperative Awareness Messages (CAMs) and Basic Safety Messages (BSMs) are expected to be transmitted periodically as a single packet and thus not to occur in bursts. PULSAR thus does not include a slow start mechanism.

A second aspect of TCP's slow start is that it accelerates AIMD's slow convergence to efficiency when starting from a low transmission rate, cf. Section 7.1.1 on page 168.

[5] The concept of flight size takes into account *flow control* in the form of the receiver's advertised window. Thus, F can be smaller than the congestion window W. When considering only the congestion control part of TCP, we can assume $F = W$.

[6] In addition, the congestion window is temporarily inflated by the number of duplicate ACKs.

[7] The figure is based on a simulation of a single TCP connection on a bottleneck link with 0.5 kbit/s capacity and an initial congestion window value of 64. The simulation scenario for the network simulator ns-2 and the applied evaluation script originate from the practical course *Simulation of Computer Networks* at the Karlsruhe Institute of Technology (KIT).

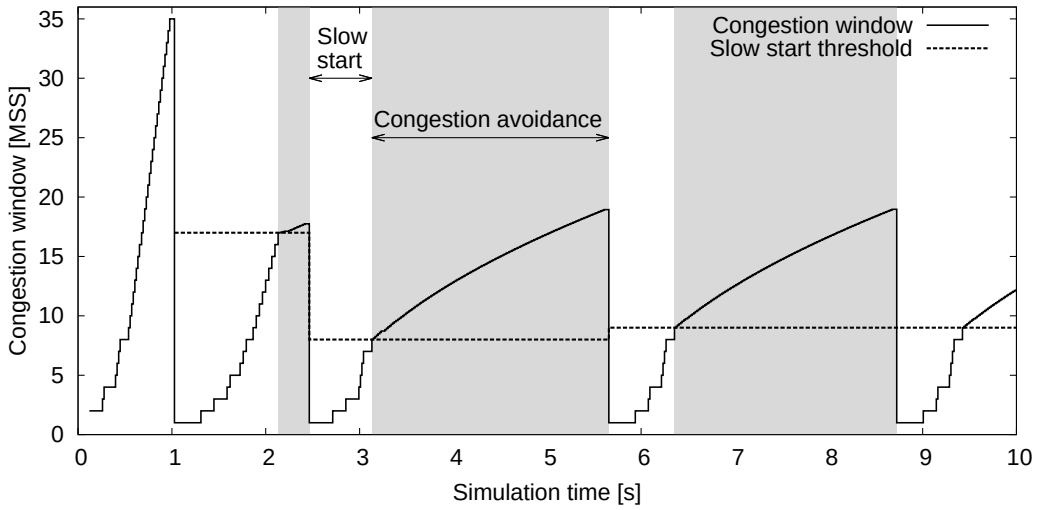


Figure 3.6: Example of congestion window adaptation in TCP Tahoe[7]

In future work, PULSAR could be extended by an additional mechanism similar to slow start which increases transmission rate faster while below a certain channel load threshold, e.g., by applying different AIMD parameters as described in Section 7.3.3 on page 195.

3.2.3 TCP over Wireless Links

Congestion control in TCP is based on the assumption that packet loss implies a congested network. In wireless networks, however, packets can be lost due to many effects such as interference from other transmissions and handoffs to other base stations, which do not necessarily imply that the network is congested. Thus, TCP in its basic form tends to reduce transmission rate unnecessarily over wireless and lossy links, resulting in suboptimal performance [129][130][131].

In [130], Balakrishnan et al. review different approaches for tackling this issue and distinguish between three categories. The first category is *end-to-end protocols* which use additional feedback to determine the congestion state of the network. Such approaches are typically based on a combination of Selective Acknowledgments (SACKs), which acknowledge the reception of individual packets, and Explicit Congestion Notification (ECN) to distinguish wireless losses from those caused by congestion. The second category is *link-layer protocols* which hide losses on the wireless link from the TCP connection by using local retransmissions and Forward Error Correction (FEC). The TCP connection then perceives the channel as having an improved quality with increased delay. Such approaches perform best if they take into account TCP's retransmission behavior. The third category is *split-connection protocols*, which terminate the TCP connection at the base station and open a second one within the wired network. However, the authors found this technique to perform suboptimally.

3.3 Wireless Networks

Wireless networks often employ congestion control solutions custom tailored to their respective characteristics and communication patterns. While some networks, e.g., cellular networks and typical wireless Local Area Networks (LANs), are centrally coordinated and typically single-hop, others, e.g., WSNs and MANETs, are self-organizing and employ multi-hop communication.

3.3.1 Cellular Networks

In cellular networks, congestion control is typically based on Transmit Power Control (TPC). For voice traffic, the objective is to maintain a sufficient Signal to Interference and Noise Ratio (SINR) for a successful connection in both downlink, i.e., from Base Station (BS) to Mobile Node (MN), and uplink, i.e., from MN to BS. Uplink power control is typically considered to be more challenging, since on the one hand the signals of all associated MNs overlap at the BS and on the other hand MNs have scarce energy resources and thus need to minimize their power output.

In [132], Yates studies under which conditions an iterative power update algorithm $p(t+1) = I(p(t))$ converges to a minimal power allocation $p = (p_1, ..., p_n)$ which allows each user $j \in [1, n]$ to overcome the interference $I_j(p)$ by other users, i.e., to achieve the required SINR. Yates shows that the algorithm converges if $I(p)$ is *positive* ($I(p) > 0$), *monotonous*, ($p_1 > p_2 \Rightarrow I(p_1) > I(p_2)$) and *scalable* ($\alpha I(p) > I(\alpha p) \forall \alpha > 1$). The conclusion also holds for uplink power control as well as for downlink if BS assignments are fixed. While the framework presented by Yates provides valuable insight into the convergence behavior of power control schemes, it has to be noted that this work, as well as others in the area of cellular networks, assume that the resulting SINR at the receiver is known for each power value [133]. Such feedback could be obtained, e.g., by piggybacking the observed Received Signal Strength Indicator (RSSI) at the destination in the packets sent back to the originator. In practice, the BS can use a single bit to tell the MN to increase or decrease its transmit power [134]. While such a practice seems feasible for unicast communications, it may be a challenge in a broadcast environment where each node would likely have to convey this information to all of its neighbors.

Surveys of other power control approaches for cellular networks can be found in, e.g., [133] and [135]. These works also review enhancements of power control by means of other techniques, e.g., admission control, beamforming and BS assignment.

3.3.2 IEEE 802.11 Wireless LAN

In a typical wireless Local Area Network (LAN), wireless nodes, e.g., laptops or smart phones, associate with an Access Point (AP) to connect to the Internet. If the node

is mobile and, e.g., moves away from the AP, link quality decreases, unless transmit power and/or data rate are adapted. APs thus typically employ an automatic data rate adaptation[8]. A commonly used technique is Auto Rate Fallback (ARF), which decreases data rate after a sequence of failed transmissions and vice versa [136]. However, ARF does not take into account the reasons for transmission failures. In case of frame collisions, it also decreases data rate, which is counterproductive for alleviating channel congestion [137][138]. CARA [137] therefore aims at distinguishing channel errors from collisions by employing the Ready To Send (RTS)/Clear To Send (CTS) mechanism after a failed transmission, cf. Section 2.1.2 on page 13. Since RTS/CTS packets are small and sent with robust modulation rate, a successive transmission failure is likely due to channel conditions. In this case, an ARF-like scheme is applied. A survey of data rate adaptation schemes can be found in [139].

A second degree of freedom influencing link quality is transmit power, which is adapted with similar motivation as in cellular networks, i.e., to save energy at the mobile station in the uplink, e.g., [140], and to mitigate interference at neighboring APs in the downlink, e.g., [141]. In [142], Ramachandran et al. introduce *Symphony*, a scheme which jointly adapts transmit power and data rate on a per-link basis, i.e., uplink or downlink, with the objective to achieve a link performance at lower transmit power which is at least as good as performance at maximum power.

3.3.3 Wireless Sensor Networks

Wireless Sensor Networks (WSNs) can be deployed to collect environmental data over large geographical areas, e.g., to detect forest fires or weak points in a dam. They are typically structured in a tree-like topology where sensor nodes transfer information, e.g., measurement data, toward the root of the tree, i.e., the *sink*, by means of multi-hop wireless communication. On the way to the sink, data is often aggregated to save energy and bandwidth. A major constraint in WSNs is that sensor nodes have very limited energy resources, which often cannot be replenished. The focus of congestion control is thus to minimize packet collisions, since collided transmissions mean a waste of energy resources. Thereby it has to be taken into account that the different "branches" can interfere with each other.

Congestion control in WSNs is typically based on rate control, often employing the AIMD mechanism, e.g., [143][144][145]. Channel congestion is detected either by leaf nodes based on Channel Busy Ratio (CBR) measurements and/or buffer queue lengths [143][144][146]. Alternatively, it is detected centrally by the sink [145]. CODA [143] employs backpressure messages to notify upstream nodes to reduce their rates as well as an end-to-end flow control based on the reception of ACKs. In WRCP [147], nodes estimate the available per-flow capacity at their location and share this information with each other, aiming for a convergence to a smooth transmission rate without oscillations. IFRC [144] has some parallels to the work presented in this

[8] E.g. Alcatel-Lucent AP 124, *http://enterprise.alcatel-lucent.com/docs/?id=11483*

thesis, since the authors develop a congestion information sharing mechanism aiming at a max-min fair rate allocation and use AIMD for rate control. However, the work addresses flows of data between source and sink and is tailored toward a tree topology.

While transmit power influences channel load as well, it is typically not considered a degree of freedom for congestion control in WSNs. Instead, transmit power is chosen as low as possible in order to maintain a connected network, an approach referred to as *topology control* [148]. Examples for topology control approaches are ATPC [149], where transmit power is adapted individually for each neighbor based on link quality feedback, DTPC [150], where each node adjusts its transmit power until it hears the desired number of neighbors, and COMPOW [151], which converges to the lowest common transmit power providing network connectivity.

3.3.4 Mobile Ad-Hoc Networks (MANETs)

As the name says, Mobile Ad-Hoc Networks (MANETs) consist of mobile nodes communicating *ad hoc*, i.e., without central coordination. Potential deployment scenarios include battlefields and disaster areas [101]. While Vehicular Ad-Hoc Networks (VANETs) can be seen as a special form of MANETs, there are significant differences. For example, VANETs have a higher mobility than other MANETs and topologies based on roads rather than open field. In addition, MANETs generally address Internet-like applications using multi-hop unicast, while VANETs target traffic safety and efficiency applications and rely on broadcast and Geocast [152]. Consequently, congestion control in MANETs is typically end-to-end based rate control based on TCP, using explicit feedback from the network to overcome link and route failures [153]. In addition, transmit power control is applied in MANETs with similar objectives as in WSNs, i.e., to save energy and to maintain network connectivity [154].

3.4 Vehicular Networks

In this section, we review dedicated congestion control approaches for vehicular communications. We structure the review based on the six available control dimensions, i.e., message size, transmission rate, carrier sense threshold, contention window, data rate and transmit power. For each control dimension, we structure our review based on the optimization objective, e.g., to maximize throughput. We thereby explicitly focus on the control of single-hop broadcast messages rather than on multi-hop broadcast schemes targeted toward avoiding the *broadcast storm problem* [101]. While the latter schemes, e.g., EMDV [20] and ZCOR [103], implicitly aim at reducing channel load as well, their primary objective is to efficiently disseminate information while minimizing the number of redundant broadcast messages. Such approaches can be seen as complementary to the control of single-hop awareness messages.

In the field of Vehicle Safety Communications (VSC), some approaches have studied the adaptation of transmit power and transmission rate with the objective of meeting the requirements of specific safety applications rather than of controlling channel load. In [19], Sepulcre et al. introduce the term *awareness control* to distinguish between both optimization objectives. They define awareness control as "those techniques aimed at ensuring each vehicle's capacity to detect, and possibly communicate with the relevant vehicles and infrastructure nodes present in their local neighborhood, through the dynamic adaptation of their transmission parameters"[19]. That is, awareness control has similarities with *topology control* whose objective is to maintain network connectivity. In contrast, "congestion control aims to limit the observed load on the wireless channel for all nodes in order to provide fair and harmonized access to the wireless medium"[19]. Both concepts thus have different objectives, control scopes and performance metrics. While the success of congestion control is measured in terms of channel load, the success of awareness control is typically quantified using a metric based on Packet Inter-Reception Time (IRT), cf. Table 2.7 on page 33. The similarities and differences between both concepts are summarized by Table 3.1 on the facing page. Sepulcre et al. conclude that a joint congestion and awareness control would be required. However, they acknowledge that such a task would be highly complex in light of potentially conflicting requirements of multiple safety applications.

In this thesis, our objective is to take a step towards a hybrid congestion and awareness control. In that respect, our work differs from many previous congestion control approaches which often did not address how their adjustments would impact the performance of safety applications. However, our main focus is on controlling channel load rather than on defining safety applications' requirements.

3.4.1 Packet Size Control

The IEEE 802.11 standard allows for a flexible frame body size between 0 and 2304 Bytes, with optional extensions for high-speed wireless LANs [42]. In [155], Yin et al. show that the optimal frame size for the Distributed Coordination Function (DCF) in terms of throughput is a trade-off between overhead and frame loss. Longer frames have a better ratio between overhead and payload than shorter frames, but they are also more likely to be lost due to transmission errors or collisions. Yin et al. conclude that an optimal frame length exists for every channel condition, independent of the number of contenting stations.

In the context of VSC, the scope for adapting packet size is limited, since the structures and header formats of CAMs and BSMs have been standardized and a CAM/BSM is expected to fit into a single frame, cf. Section 2.1.3 on page 18. Existing approaches typically focus on omitting redundant information and reducing security overhead.

Criterion	Congestion control	Awareness control
Objective	Limit channel load	Meet safety applications' requirements
Control scope	All vehicles	Individual vehicle or subset
Desired outcome	Homogeneous Tx parameters	Individual settings
Fairness concept	yes	no
Metrics	Measured (CBR) or estimated (beaconing load)	Reliability of awareness range
(Main) degrees of freedom	Transmit power and transmission rate	Transmit power and transmission rate

Table 3.1: Congestion and awareness control according to Sepulcre et al. [19]

Message Dispatcher

In [156], Robinson et al. introduce the concept of a *message dispatcher* which includes the minimum number of data elements in a BSM based on which applications are run by the vehicle itself and its neighbors. For example, if application *A* requires data element *a* at a rate of 2 Hz, *a* is only included in every 5th packet, given a transmission rate of 10 Hz. In [157], Robinson et al. further extend the concept based on *predictive coding*, i.e., data elements are only included "when the error in estimating the state is 'sufficiently large' ". In other words, data elements are sent to update a receiving vehicle's model of the sending vehicle's state. The authors show that their scheme is able to reduce channel load by over 80% compared to the baseline of transmitting every data element in every packet. However, in light of the already small size of the mandatory elements of CAMssBSMs, the saving potential of the message dispatcher appears limited compared to the size of the security overhead.

Certificate Omission

Recent works have thus studied the potential of reducing channel load by omitting certificates. That is, while a signature is required in every CAM/BSM to verify its authenticity and integrity, certificates could be cached by a receiver and thus may not have to be attached to every packet.

In [158], Feiri et al. compare three certificate omission schemes, i.e.,

1. Periodic Omission of Certificates (POoC) [159], which includes the certificate in every *n*th CAM;
2. Neighbor based Certificate Omission (NbCO) [67], which includes the certificate if there is a change in the neighbor table, i.e., if a new neighbor has been discovered;
3. Congestion based Certificate Omission (CbCO) [158], which includes the certificate in every *n*th CAM, where *n* depends on the number of vehicles in the neighbor table.

The authors argue that POoC may perform too restrictively in situations with low channel load, while NbCO tends to attach more certificates in situations with high vehicle density, which is counterproductive for scalability. They study the trade-off between *network packet loss*, i.e., due to collisions or queue drops, and *cryptographic packet loss*, which occurs if a signature cannot be verified due to a missing certificate and the packet consequently has to be discarded. They show that CbCO can lead to a reduction in *total packet loss*, i.e., network and cryptographic loss combined, of 300% when applying the CbCO scheme compared to when no certificates are omitted at all. However, their results also show that up to 70 CAMs/BSMs can be received in a row without the possibility of verification, which corresponds to an IRT of up to 7 s due to cryptographic packet loss alone and might be problematic for safety applications.

We conclude that, while certificate omission holds potential for the reduction of channel load, the reception of unverifiable CAMs/BSMs is undesirable. Thus, it would be preferable to first exploit other means of reducing channel load. We thus do not further consider the adaptation of message size in this thesis.

3.4.2 Data Rate Control

As discussed in Section 3.3 on page 57, data rate control plays an important role to maintain connectivity when link quality changes. In vehicular communication networks, link quality is prone to change due to the high mobility involved. Consequently, dedicated algorithms have been developed, typically targeting infotainment applications like video downloading and web browsing which employ unicast flows between vehicles and/or infrastructure.

Maximization of Throughput for Unicast Communication

In [160], Camp et al. propose to perform an estimation of the current *channel coherence time*, i.e., the time during which the impulse response of the channel does not change significantly, in order to select the best data rate for the current environment. If the transmission time of a packet is longer than channel coherence time, the char-

acteristics of the channel may change too much to decode the packet successfully. In [161], Shankar et al. introduce CARS, a hybrid of proactive and reactive data rate control based on an estimation of Packet Error Rate (PER). With increasing vehicle speed, more weight is given to a proactive estimation of PER based on vehicle density than to a reactive EWMA of past transmission statistics.

BRAVE [162] is motivated by the idea of offloading mobile data generated by smart phones to urban wireless LAN hotspots in order to save the costs of using cellular networks. In a nutshell, BRAVE monitors the RSSI values of received ACKs over 500 ms and uses a lookup table to select the data rate for the next four transmission attempts.

Maximization of Broadcast Performance

In contrast to the aforementioned approaches developed for infotainment applications, a fixed data rate is typically assumed for safety communications. Since VSC is based on single-hop broadcast, there is not one but many receivers for whom to optimize reception performance. Also, the primary objective is to establish a mutual awareness among vehicles rather than to transfer the maximum amount of data.

In [43], Jiang et al. study the question of optimal data rate selection for VSC. By means of simulations, the authors evaluate the reception performance for three intermixed groups of vehicles on a long circular highway. Two groups serve as reference groups to ensure that the Communication Density (CD) of the scenario does not change, cf. Section 2.3.3 on page 40. That is, for different data rates the transmit power of the third group is set such that the two reference groups experience the same Packet Delivery Ratio (PDR) and thus the same amount of interference. The authors conclude that a data rate of 6 Mbit/s results in the highest PDR in most scenarios, with the exception of situations with very low or very high channel load, where 4.5 Mbit/s and 9 Mbit/s perform slightly better, respectively. Their results are consistent for packet sizes of 100 Bytes, 200 Bytes and 500 Bytes.

The optimality of the 6 Mbit/s data rate has been confirmed in [163], where Bai et al. measured the PDR between two moving vehicles in different highway and suburban environments. Their findings indicate that a data rate of 6 Mbit/s results in a significantly higher PDR than a data rate of 18 Mbit/s, which can be explained based on the more robust modulation scheme. However, they also find that 6 Mbit/s outperforms 3 Mbit/s in terms of PDR, even though its modulation is less robust. They attribute this result to the chosen packet size of 300 Bytes, which results in a transmission time of 0.4 ms and 0.8 ms for 6 Mbit/s and 3 Mbit/s, respectively. The coherence time in a rural or highway environment being about 0.3 ms to 0.4 ms [36], Bai et al. conclude that the benefit of being able to decode a packet during channel coherence time outweighs the benefits of a more robust modulation. They do not, however, present results for different packet sizes.

The problem of transmission time and channel coherence time results from the fact that in the IEEE 802.11p standard, the channel is estimated only once for each packet.

In [164], Fernandez et al. thus introduce an enhanced scheme which updates channel estimates *during* the reception of a packet. They study the performance of their scheme compared to the default channel estimator for different packet sizes and data rates using calculations based on real-world measurements. Their results indicate that the new design outperforms the default estimator in terms of PDR while using one data rate step higher. For example, 6 Mbit/s in the new design outperforms 3 Mbit/s in the old design. They find that, for a fixed packet size, a higher data rate always performs worse than a lower one, with the exception of 6 Mbit/s which outperforms 4.5 Mbit/s. In terms of throughput, however, they find that a data rate of 12 Mbit/s performs best up to a packet size of about 600 Bytes.

Due to the findings of the aforementioned studies as well as its usage as a default value in the VSC-A project [6], 6 Mbit/s has become widely accepted as a default data rate for VSC, e.g., [47][71][165][166]. However, evaluations so far were typically only based on the packet-level reliability metric PDR, which does not necessarily mean that the *application* level reliability is also optimized at this data rate. The results of a recent measurement campaign by Sepulcre et al. indicate that a data rate of 3 Mbit/s may result in a higher application reliability than 6 Mbit/s [21]. However, the study does not state which channel estimator and packet size were used. In addition, Sepulcre et al. focused on an urban intersection scenario rather than a highway environment like the other studies. In a heavily non-line of sight dominated environment, the higher robustness of 3 Mbit/s may outweigh other effects and thus lead to different conclusions.

Preliminary results using different data rates in the optimization study presented in Section 4.2.2 on page 89 indicate that the optimality of the 6 Mbit/s data rate still holds based on the metric of minimizing IRT at a certain sender-receiver distance in a controlled highway-based simulation environment and for a fixed packet size. However, the examples above show that the selection of a data rate requires a consideration of details beyond the granularity of a packet-based network simulator as employed in this thesis. For this reason, we adhere to the commonly accepted value of 6 Mbit/s and leave a further exploration of this topic for future work which could, e.g., apply a detailed physical layer simulator like *PhySim* [48].

3.4.3 Contention Window Control

As discussed in Section 2.1.2 on page 13, the choice of the contention window size is a trade-off between packet collision probability and channel access delay. In addition, the dynamic adaptation of the contention window based on binary exponential backoff is not applicable to broadcast communications.

Optimal Contention Window in a Saturated IEEE 802.11 Network

According to Bianchi's analysis in [167], the probability of a frame collision in a saturated IEEE 802.11 network, i.e., if every station has a frame to transmit at every point in time, is given by

$$p_{coll} = 1 - (1 - \tau)^{n-1} \tag{3.9}$$

where n is the number of contending stations and the probability τ that a station transmits in a randomly chosen time slot is given by

$$\tau = \frac{2}{W + 1} \tag{3.10}$$

for a fixed contention window size W. Figure 3.7 on the following page illustrates p_{coll} for different fixed contention window sizes. We observe that a contention window of 15, the maximum default value of CW_{min} in IEEE 802.11p, results in a collision probability of near 100% for 40 or more competing stations.

According to [167], the optimal contention window size (in terms of system throughput) is given by

$$W_{opt} = n\sqrt{2T_c/T_s} \tag{3.11}$$

where T_c is the average duration of a frame collision and T_s is the duration of a time slot, i.e., 13 μs for IEEE 802.11p. For a fixed payload size P, T_c is given by

$$T_c = T_H + T_P + T_{DIFS} + T_{tx} \tag{3.12}$$

where T_H is the transmission time of the Physical layer (PHY) and Medium Access Control (MAC) header, T_P is the transmission time of the payload, T_{DIFS} is the duration of the DCF Interframe Spacing (DIFS), i.e., 58 μs for IEEE 802.11p, and T_{tx} is the transmission time, which is approximately 0.8 μs for 250 m at the speed of light. Assuming a total packet size of 400 Bytes in a 6 Mbit/s channel, we thus get $W_{opt} = n\sqrt{2(533\mu s + 58\mu s + 0.8\mu s)/13\mu s}$, i.e., $W_{opt} = 382$ for 40 vehicles and $W_{opt} = 3816$ for 400 vehicles[9].

The example illustrates why contention window adaptation generally plays an important role in IEEE 802.11 networks. However, while for typical wireless LANs, the initial assumption of non-empty transmission queues at every point in time can be reasonable, it corresponds to a worst-case assumption for VSC. First of all, CAMs and BSMs are generated at a maximum rate of 10 Hz, i.e., once every 100 ms [16][29]. Second, since the information contained in safety messages is only useful for a very short period of time, it does not make sense to queue them if a new message has been generated, cf. [99][168][169]. Third, the optimization in [167] did not take into account how long it may take a station to access the medium. Backoff times of 5 ms

[9] A derivation very similar to this one has been previously presented by Kaul et al. in [99].

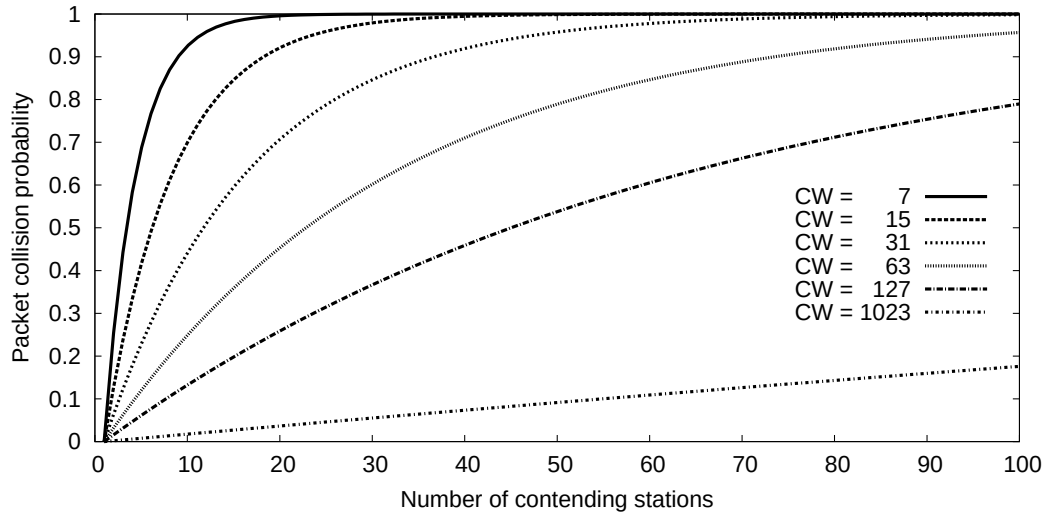


Figure 3.7: Packet collision probability for fixed contention window (CW) size, assuming that every station has a packet to send at every point in time [167]

and 50 ms for contention window sizes of 382 and 3816, respectively, are significant compared to a transmission interval of 100 ms. Different works thus came to the conclusion that a contention window size of 15 performs reasonably well, e.g., [99] and [170]. In the following, we review these and other approaches for vehicular communications in greater detail.

Maximization of Per-Packet Reception Probability

A number of approaches suggest to optimize contention window in order to minimize the number of packet collisions due to simultaneous transmissions, i.e., to maximize the PDR. In [171], Balon and Guo suggest to adapt the contention window according to the current PDR, which is estimated based on the number of missing sequence numbers in received packets. If the PDR has increased or decreased by a certain threshold, the contention window is decreased or increased, respectively.

In [172], Mertens et al. calculate the optimal contention window according to Bianchi's analysis, setting n in Equation (3.11) on the previous page to the number of entries in a vehicle's neighbor table. Additionally, they adapt data rate proportionally to channel load to compensate for the increasing delay with higher vehicle density. However, the authors did not address that the number of neighbor table entries does not necessarily correspond to the number of contending nodes, depending on the ratio of reception range and carrier sense range.

In [173], Stanica et al. introduce a *reverse back-off mechanism* which starts with a large contention window and halves it for every expired CAM, i.e., every CAM that could not be transmitted before a new one was generated. In their evaluation, they study a scenario which results in 24 and more consecutive packet losses in a row, indicating

that the channel is highly congested. In [174], the same authors use simulations to determine an optimal contention window for each vehicle density based on the maximization of PDR at close distances. The only considered delay constraint is the generation of the next CAM, at which the old one expires and is counted as lost in the PDR. Like Bianchi, they find that the optimal contention window scales linearly with the number of nodes, the linear factor depending on packet size. For example, a vehicle density of 50 nodes results in an optimal contention window of about 300. However, their results indicate that the PDR is similar, i.e., between 0.65 and 0.73, across a wide range of contention window sizes, i.e., between about 15 and 1400. The authors thus conclude that while adapting the contention window can improve PDR, it is not sufficient as a sole measure to solve the scalability problem in vehicular communications [175].

It should also be noted that the aforementioned studies typically focused on highly congested scenarios in which packet collisions due to simultaneous backoff countdowns are a frequent event. In [47], Mittag and Hartenstein analyze the performance of Carrier Sense Multiple Access (CSMA) for a contention window of 15 in a highly detailed physical layer simulator. They conclude that *Packet Level Incoordination (PLI)*, which includes collisions due to simultaneous countdowns, is effectively controlled by CSMA unless the channel is saturated.

Minimization of End-to-End Delay

A major drawback of IEEE 802.11p in the context of safety communications is that it cannot give delay guarantees due to its probabilistic channel access strategy. To mitigate this issue, a lower contention window can be assigned to higher priority messages in order to give them a higher probability of a short channel access delay. For example, this approach is applied in the Enhanced Distributed Channel Access (EDCA) mechanism, cf. Section 2.1.2 on page 13. Some authors have suggested alternative or complementary algorithms. In [176], Eichler and Schroth suggest to set the contention window individually for each packet based on its *utility* value. Similarly, in [177] Taleb et al. calculate a vehicle's contention window based on its safety-related *emergency level*. In [178], Zang et al. suggest to adapt the contention window of each EDCA access category based on a CBR-related channel load estimation, i.e., to double or halve the contention window if channel load is above an upper threshold or below a lower threshold, respectively. The objective is to reserve a fraction of bandwidth for event-driven emergency messages at the cost of increasing channel access delay for lower-priority messages, e.g, CAMs.

Application-Layer Optimization

While the aforementioned studies suggested optimizations on a per-packet basis, few studies have so far addressed the impact of contention window adaptation on safety applications. In [99], Kaul et al. present a study on the minimization of *system age*, i.e., "the average end-to-end [...] delay observed in any vehicle's state within a

certain cluster of nodes", cf. Table 2.7 on page 33. In other words, the system age reflects the average age of the received status information of a vehicle's neighbors. The authors show that for each vehicle density, there is an optimal point minimizing information age and that this point cannot be reached by adapting contention window alone. Kaul et al. therefore use a contention window of 15 and present an algorithm adapting message generation rate control.

In [170], Reinders et al. study the performance of different contention window sizes on PDR and IRT, cf. Section 2.2.3 on page 30. Their results indicate that PDR is similar for different contention window sizes as long as the channel is not congested. On the other hand, they show that a high contention window has a detrimental impact on IRT due to the increasing channel access delay. In the end, the authors conclude that, contrary to their expectations, an increased contention window does not help to improve the reception performance of CAMs/BSMs.

Conclusion

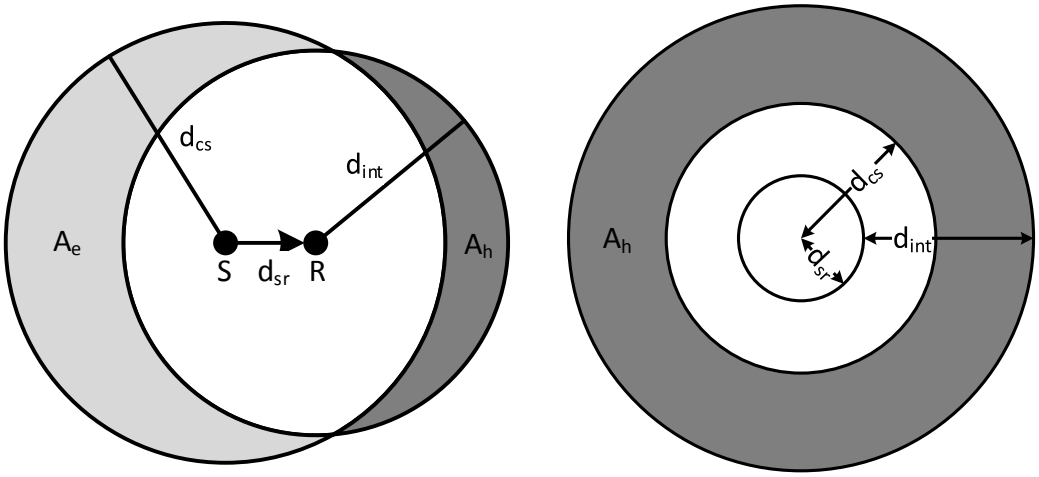
To summarize, different studies have shown independently of each other that contention window adaptation has a limited potential to improve PDR and IRT as long as the channel is not used beyond its capacity [48][99][174][170]. Since in this thesis, our objective is to keep the system near the *knee* rather than near the *cliff*, we thus use a fixed contention size of 15 according to the IEEE 802.11p standard and do not further investigate the topic of contention window adaptation. In particular, one of our objectives is to *prevent* the network from getting to a state where CAMs/BSMs expire.

3.4.4 Carrier Sense Threshold Control

The carrier sense threshold, i.e., the level of measured energy above which the radio considers the medium to be busy, has been studied with different motivations in the related work. While in common wireless LANs, the objective typically is to maximize network throughput and to avoid asymmetric links, approaches for vehicular networks typically focus on avoiding excessive channel access delays for safety-critical information.

Maximization of Throughput

In [110], Ma et al. study the influence of the carrier sense threshold on the trade-off between hidden terminals and exposed terminals in unicast communications, cf. Section 2.1.2 on page 13. Their objective is to find a carrier sense threshold which maximizes overall throughput by balancing exposed stations and hidden stations. While the former ones reduce throughput due to an inefficient channel usage, the latter ones reduce throughput due to packet collisions. Figure 3.8a on the facing page illustrates the relationship between sender-receiver distance d_{sr}, carrier sense range d_{cs}, and interference range d_{int} in the case of a transmission from node S to node R,



(a) Unicast, cf. Figure 1 in [110]

(b) Broadcast, cf. Figure 1 in [179]

Figure 3.8: Relationship between sender-receiver distance d_{sr}, carrier sense range d_{cs}, interference range d_{int}, hidden node area A_h and exposed node area A_e

cf. Section 2.3.2 on page 34. Ma et al. show that the hidden node area A_h is minimized for $d_{cs} = d_{sr} + d_{int}$, while the exposed node area A_e is minimized for $d_{cs} = d_{sr} - d_{int}$. The optimal carrier sense range balancing both areas thus lies in between.

Note that in a broadcast environment, exposed nodes in the classical sense do not exist, cf. [44]. The corresponding relationship between d_{sr}, d_{cs} and d_{int} is presented in Figure 3.8b. In [109], Yang et al. conclude that the throughput of a multi-hop wireless network is maximized if $d_{cs} = d_{sr} + d_{int}$ and transmit power is set as low as possible to maintain connectivity. Figure 3.8b illustrates that this setting theoretically eliminates the hidden node area A_h.

In the following, we transfer this finding to the context of VSC and the propagation model applied in this thesis. To maximize the communication range of a broadcast transmission, it would be desirable to have $d_{sr} = d_{Tx}$, i.e., to allow for a sender-receiver distance equal to the theoretical communication range without interference. Following the reasoning of Yang et al., we thus get $d_{cs} = d_{Tx} + d_{int}$. By plugging in Equation (2.6) on page 39 for d_{cs}, Equation (2.2) on page 36 for d_{Tx} and Equation (2.3) on page 36 for d_{int}, we get an optimal carrier sense threshold of

$$P_{cs}{}^{*} = P_{RxTh} - 10\gamma log_{10}(1 + 10^{S_o/(10\gamma)}) \tag{3.13}$$

according to the Signal to Interference Ratio (SIR) model, where P_{RxTh} is the required Received Signal Strength (RSS) for packet reception in dBm, γ is the path loss exponent and S_o is the SIR required for reception in dB. For $P_{RxTh} = -91$ dBm, $S_o = 8$ dB and $\gamma = 1.85$, we thus get $P_{cs}{}^{*} \approx -101$ dBm, which is below the assumed noise floor of -99 dBm and thus infeasible. We conclude that for uniform transmit power levels, the carrier sense threshold should be set low in order to maximize throughput.

Avoidance of Link Asymmetry

In [141], Mhatre et al. study the adaptation of the carrier sense threshold with the motivation of avoiding asymmetric links which can result from Transmit Power Control (TPC). The authors argue that with TPC, a low-power node may become a permanently exposed station to a high-power node if both use the same carrier sense threshold, since the latter would always sense the medium to be idle and would thus continue to transmit, while the former one would back off. To ensure that the network stays *symmetric*, i.e., that each node is able to "generate sufficient interference to suppress the transmission of [a] reference node", the authors show that a joint adaptation of transmit power and carrier sense threshold is required. More specifically, they suggest to select the carrier sense threshold inverse proportionally to the transmit power used. Intuitively speaking, this means that "if you want to shout, you need to listen more carefully so as not to disturb those who are whispering" [141]. Next to the avoidance of starvation in nodes, the authors further find that the suggested algorithm significantly improves network throughput compared to a uniform configuration. A similar observation was previously made in [180], where Fuemmler et al. find that the product of carrier sense threshold and transmit power should be constant in order to maximize throughput.

Note that [141] and [180] focus on typical IEEE 802.11 networks which are dominated by unicast communication between a station and its associated AP, where the station likely has a full queue of packets to send. In contrast, VSC is based on one-hop broadcasts of short periodic messages which are not expected to be queued due to their short lifetime. Due to the lack of exposed terminals in the classical sense and the expected gaps between packet transmissions from a single node, we assume that the starvation problem described by Mhatre et al. does not apply to VSC. While we did not observe fairness issues due to different transmit power levels in the simulations conducted in this thesis, future work might look into this topic more closely to determine if fairness issues may arise in certain scenarios where a low-power node is surrounded by many high-power nodes, in particular if for some reason high-power nodes would keep a queue full of packets instead of only one queued packet. Future studies might also consider the question if network throughput may be increased by adapting the carrier sense threshold as suggested in [141] and [180]. However, since Fuemmeler et al. observed fairness issues in their approach, such a solution would have to make sure that the fairness principles described in Section 5.3 on page 119 are still met.

Avoidance of Excessive Channel Access Delay

In the context of vehicular networks, the main focus of adapting the carrier sense threshold has been on controlling channel access delays. In a saturated channel, the accumulated Channel Busy Time (CBT) sensed by a node with a low carrier sense threshold may exceed the CAM/BSM generation interval, which leads to the undesired event of either sending outdated information or having to drop a packet with safety relevant information.

In [181], Schmidt et al. compare the performance of three fixed carrier sense thresholds in the context of periodic CAM transmissions. They observe an improvement of up to 200% in PDR when decreasing the carrier sense threshold from -65 dBm, i.e., the default value as of the IEEE 802.11 standard, to -85 dBm. However, they also observe an increase in average end-to-end delay of two orders of magnitude, i.e., from 0.2 ms to up to 20 ms. In [182], Schmidt et al. introduce an algorithm to adapt a vehicle's carrier sense threshold based on channel conditions. The objective of the control is to limit the channel access delay experienced by a vehicle. The scheme starts with a low carrier sense threshold, i.e., -95 dBm, and increases the carrier sense threshold by one step for every exponentially decreasing time interval the CAM could not be sent. The authors show that in a non-saturated scenario, a fixed carrier sense threshold of -95 dBm provides the best *awareness quality*, i.e., the highest ratio of neighbors from whom a CAM was received within a distance-dependent timing requirement. For a scenario with higher channel load, the results indicate that the adaptive scheme provides better awareness quality for close distances. However, it should be noted that the evaluated scenario with "high" channel load results in an average channel access delay of up approximately 50 ms for a carrier sense threshold of -95 dBm, which indicates that the injected channel load is significantly beyond the channel capacity, as we will see in Section 5.2 on page 111.

In [183], Stanica et al. introduce an algorithm which adjusts the carrier sense threshold linearly with vehicle density between a minimum and maximum value. The basic idea is to control the number of vehicles within carrier sense range in order to limit packet collisions due to simultaneous backoff countdowns on the one hand and to avoid channel access delays due to long channel busy times on the other hand. The authors show that their scheme has an improved PDR at close distances compared to a fixed carrier sense threshold of -95 dBm. However, like Schmidt et al., they focus primarily on highly saturated channel conditions. In this thesis, our objective is to adapt transmission parameters in order to *prevent* the system from getting to such a highly congested situation. Thus, we consider approaches like [182] and [183] to be complementary to our work.

Conclusion

The hidden terminal problem plays a major role in VSC. Figure 3.9 on the following page illustrates the relationship of carrier sense range and hidden nodes schematically in a scenario where two highways cross on top of each other. The figure makes the simplifying assumption that $d_{int} = d_{Tx}$ and considers three settings of the carrier sense range, i.e., smaller than the communication range (case A), equal to the communication range (case B) or twice the size of the communication range (case C). From the point of view of the vehicle in the center of the intersection, vehicles which are no hidden terminal in any setting are indicated in white. Black vehicles are hidden terminals in cases A and B, while grey vehicles are hidden terminals only in case A.

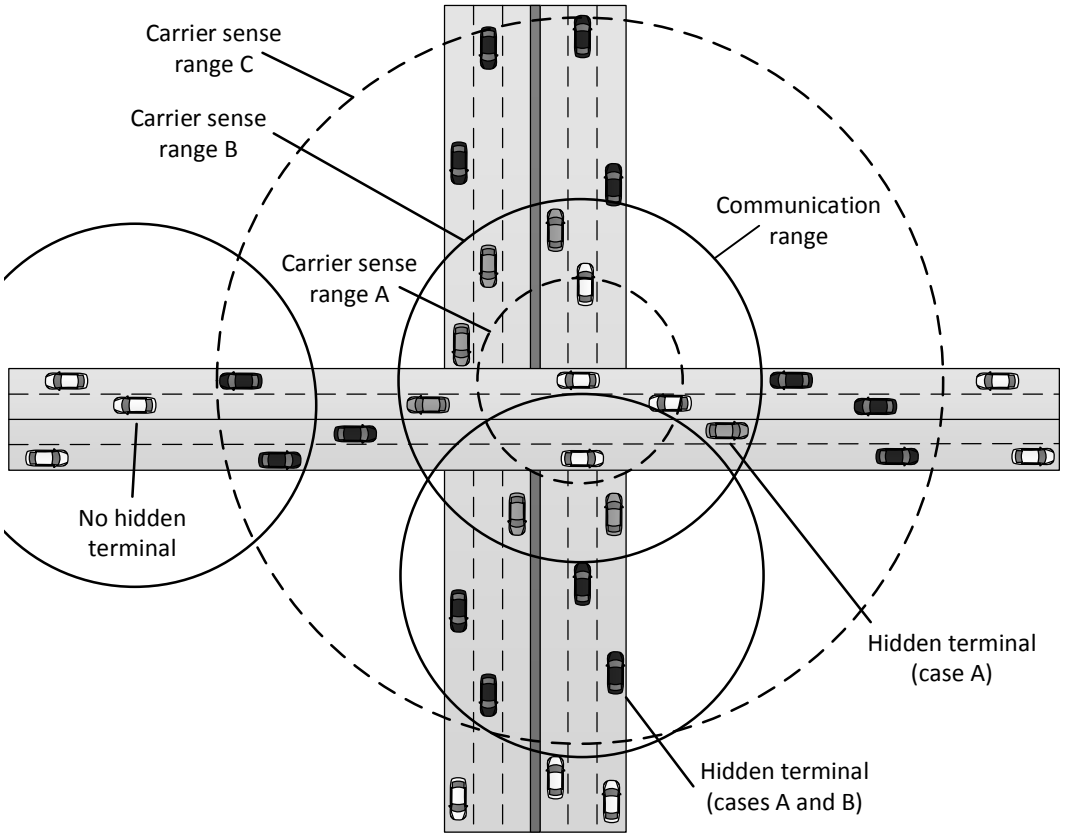


Figure 3.9: Schematic illustration of the impact of the carrier sense range on hidden terminals in a highway crossing scenario

The example illustrates that with a small carrier sense range, i.e., a high carrier sense threshold, the number of hidden terminals can become significant in a traffic scenario. In this thesis, we thus apply a low carrier sense threshold, which is in line with the findings of Yang et al. [109] and Schmidt et al. [182] as discussed above. More specifically, we set the carrier sense threshold to -95 dBm, which was chosen to be reasonably close to the noise floor while accounting for the capabilities of current hardware. Using Equations (2.2) on page 36 and (2.8) on page 39, we can calculate that for the propagation model and the default settings used in this thesis (cf. Appendix A on page 223) we get $d_{cs} \approx 2.16 d_{Tx}$. That is, our chosen carrier sense threshold corresponds approximately to configuration C in Figure 3.9.

Note that we apply this value also in the case of heterogeneous transmit power. While we did not encounter node starvation as described by Mhatre et al. [141] in our simulations, future work could explore if similar fairness issues might arise in other scenarios. In addition, future work might look at the question if the joint adaptation of carrier sense threshold and transmit power similar to [141] and [180] could help to improve spatial reuse in VSC compared to a fixed carrier sense threshold while not compromising fairness.

3.4.5 Transmit Power Control (TPC)

As discussed in Section 3.3 on page 57, Transmit Power Control (TPC) plays a major role in many different wireless networks. One of the objectives discussed so far, i.e., to minimize energy consumption, is typically assumed to play a minor role in vehicular networks due to the availability of a vehicle's propulsion system as an energy source. TPC therefore typically focuses on the mitigation of interference and/or on the maintenance of connectivity, i.e., topology control. In the following, we discuss related approaches based on their optimization objective.

Topology Control

In [184], Artimy introduces DTRA, an algorithm which adapts a vehicle's transmit power based on its local estimate of vehicle density. The basic idea is that if vehicle density increases, the distance between vehicles decreases and thus less transmission range is required to maintain connectivity. The local density estimation is based on the ratio between the vehicle's stopping time and the total trip time and thus does not require information exchange.

Similar ideas are expressed, e.g., in [185], where Caizzone et al. estimate vehicle density based on overheard messages and use Additive Increase Additive Decrease (AIAD) to adapt transmit power in order to keep the number of a vehicle's neighbors within an upper and a lower bound.

In [92], Mittag et al. study the benefits of relaying CAMs over two hops using lower transmit power compared to single-hop broadcast with higher transmit power in order to cover the same distance. They find that under perfect channel conditions, two-hop beaconing can indeed reduce channel load. Under realistic channel conditions in which packets can be lost, however, their results indicate that the theoretical benefits cannot be attained. They emphasize, however, that in urban non-line-of-sight conditions, a multi-hop relaying of CAMs might still be beneficial.

Compensation of Slow Fading

In [186], Guan et al. introduce a TPC algorithm which aims at maintaining a target communication range while compensating link quality changes, e.g., changes in path loss and shadowing due to line-of-sight obstructions. In a nutshell, each vehicle adapts its transmit power based on whether or not a sufficient number of neighbors within the desired target range has received its previous transmission. To provide feedback about the achieved transmission range, each node piggybacks two pieces of information in its periodic messages, i.e., its own target range as well as a subset of the node IDs from whom it has heard of since the last transmission and of whom it has been an intended receiver.

In [187], Cheng and Shakya introduce a two-stage algorithm with similar objective, i.e., to compensate link quality changes due to mobility and changes in the environment. In the first stage, probe packets are sent with increasing transmit power to determine the closest neighbor. In the second stage, transmit power is set such that it is just greater than the receiver threshold.

Congestion Control

In [188], Torrent-Moreno et al. introduce D-FPAV, a TPC algorithm whose objective is to keep the Beaconing Load (BL), cf. Section 2.3.3 on page 40, at any location on a highway below a maximum value. In essence, the algorithm follows a distributed water-filling approach to achieve max-min fairness, cf. Section 3.1 on page 46. By sharing information on the transmit power values used not only by a node itself but also by its neighbors, each node can locally calculate the maximum transmit power it can use in order not to violate the maximum BL requirement within its carrier sense range. Each node then sets its transmit power to the minimum of its own calculated value and the corresponding values calculated by its neighbors. The information sharing mechanism of D-FPAV causes a significant overhead, which is reduced by Mittag et al. in [189] by relaying the number of vehicles per road segment rather than their exact positions.

The D-FPAV algorithm has two similarities with the approach introduced in this thesis, i.e., the focus on max-min fairness and the sharing of congestion information over two hops. However, while PULSAR adapts transmission rate to control channel load and adapts transmit power to meet awareness requirements, D-FPAV adapts transmit power only and does not address safety applications' requirements. Further, it is a proactive solution rather than a reactive one, i.e., it relies on the accuracy of the employed radio-wave propagation model to correctly predict the channel load resulting from a transmit power allocation. In contrast, PULSAR is a reactive approach based on CBR measurements.

In [190], Lu and Poellabauer introduce ETPC, an algorithm trying to establish a common transmit power throughout a region. The basic idea is to prevent nodes located farther away from a congested location from increasing their transmit power, since they might become new hidden terminals. In ETPC, a vehicle detecting channel congestion calculates a maximum transmit power which is propagated throughout the neighborhood. Transmit power is increased if no lower value has been received for a certain time. While implemented differently, ETPC basically shares the objective of *participation fairness* with the approach presented in this thesis. However, the solution of Lu and Poellabauer does not provide a limit for the *participation range*, cf. Section 5.3 on page 119.

In [191], Khorakhun et al. introduce a TPC scheme in which they compare the measured CBR, cf. Section 2.3.3 on page 40, against a threshold. If the CBR is too high, transmit power is reduced and vice versa. The new transmit power is calculated

using Multiplicative Increase Multiplicative Decrease (MIMD), cf. Section 3.1 on page 46. For fairness, a node may only increase its transmit power if its current value is below the average transmit power used by its neighbors. Like the approach introduced in this thesis, the protocol by Khorakhun et al. uses a binary adaptation based on CBR measurements. In 7.4.1 on page 205, we thus compare our solution against the transmission rate control variant of Khorakhun et al.'s approach which is presented in [191] as well.

In [192], Nasiriani et al. study the convergence and fairness behavior of two algorithms adapting transmission range, i.e., Linear Range Control (LRC) and Gradient descent Range Control (GRC), of which the former is used in the joint power and rate control scheme by Huang et al. [23], which we discuss in the next subsection. We present the power control part here because it is independent of the rate control part of [23]. The LRC scheme scales transmission range D linearly between a minimum value D_{min} and a maximum value D_{max}, depending on where the current CBR U lies between a minimum value U_{min} and a maximum value U_{max}, i.e.,

$$D_{k+1} = \begin{cases} D_{max} & \text{if } U_k < U_{min} \\ D_{min} + \frac{U_{max} - U_k}{U_{max} - U_{min}}(D_{max} - D_{min}) & \text{if } U_{min} \leq U_k < U_{max} \\ D_{min} & \text{if } U_{max} \leq U_k \end{cases} \tag{3.14}$$

The GRC scheme, on the other hand, calculates D as

$$D_{k+1} = min(D_{max}, max(D_{min}, D_k + \eta(U^* - U_k))) \tag{3.15}$$

where U^* is the targeted CBR value optimizing reception performance, i.e., $U^* = 0.7$ CBR [193] and η is a scaling factor with a default value of 50. Nasiriani et al. find that the LRC scheme is sensitive to the choice of the corresponding minimum and maximum values and might not converge otherwise, in which case they recommend the GRC scheme.

Awareness Control

As already mentioned at the beginning of this section, the objective of awareness control is to adapt transmission parameters in order to meet the requirements of safety applications. The resulting channel load is typically observed but not specifically controlled.

In [194], Gozalvez and Sepulcre introduce OPRAM, an algorithm which computes a *critical distance* before which a vehicle approaching an intersection needs to receive a CAM/BSM from another approaching vehicle in order to prevent a potential crash. The algorithm uses a low default transmit power to communicate with vehicles on the same road in Line of Sight (LOS) conditions. One second before the vehicle reaches the critical distance, the transmit power is set to a maximum value, then decreased gradually to the default value as the vehicle approaches the critical distance.

In [195], Sepulcre et al. consider a highway scenario and calculate the critical distance for the Lane Change Warning (LCW) application, cf. Table 2.4 on page 28. They find that a number of combinations of transmit rate and power can fulfill the requirement of receiving at least one CAM before the critical distance and within a time window T. In the evaluation, they use a fixed transmission rate and adapt transmit power until the desired application reliability is reached. They show that channel load can be reduced if additional knowledge about the traffic situation is exploited and unnecessary CAMs are not transmitted.

In [196], Kloiber et al. suggest to use a random transmit power, e.g., following a Gaussian distribution. The underlying assumption is that the required *awareness quality*, i.e., the up-to-dateness and accuracy of the status information of neighboring vehicles, decreases with distance. Using a random transmit power, vehicles close to the transmitter receive frequent position updates, while more distanced vehicles receive updates less frequently. Compared to a scenario using maximum transmit power, the authors observe an improvement in IRT for close distances as well as a reduction in channel load.

Conclusion

While transmit power plays an important role in congestion control as well as in awareness control, there is no apparent consensus in the related work as to how and why it should be adapted, especially in order to meet both control objectives. In Chapter 4 on page 83, we analyze the impact of transmit power on IRT at a certain distance. Our results indicate that for each target distance, there is an optimal transmit power in terms of minimizing IRT. This result can help to significantly reduce the complexity of congestion control. We discuss in Chapter 5 on page 103 how the Transmission Rate Control (TRC) approach suggested in this thesis can be further augmented by means of an open loop TPC.

3.4.6 Transmission Rate Control (TRC)

In contrast to other communication networks, the periodic messages generated in VSC are expected to be of similar size and to be transmitted in a single packet, cf. Section 2.1.3 on page 18. For this reason, the term *transmission rate* is typically used to refer to the generation rate of CAMs/BSMs in units of Hertz rather than in the unit of Mbit/s. In the context of awareness control, it has been suggested to transmit CAMs/BSMs aperiodically as required by the driving context rather than periodically. We review these approaches at the end of this subsection.

Utility-Based Approaches

Some approaches have been suggested to alleviate the congestion problem in vehicular networks by prioritizing messages based on an *utility function*, e.g., [197] [198] [199]

[200]. Typically, these approaches target non-safety applications which disseminate information in a certain area by means of one-hop broadcasting. Message priorities are determined based on factors like the distance to an event, message age, vehicle speed and the new area to be covered by a (re-)transmission. In [197], [198] and [200], messages with high priority arriving in the transmission queue get transmitted first. If the channel allows it, low priority messages are transmitted later. While such an approach can help to control channel load, it seems difficult to apply in the case of VSC where CAMs/BSMs are only valid for a short time and thus delaying them seems counterproductive.

In Adaptive Traffic Beacon (ATB) [199], the transmission interval is adapted based on a utility function which consists of message priority and channel conditions. This approach thus has some similarities with the *relative transmission rate* concept introduced in this thesis, which adapts the transmission rate based on a safety-related utility function as well as on channel load. However, ATB has a different objective and controls the transmission *interval* rather than transmission rate. We discuss the implications of adapting the transmission interval in Section 5.4.2 on page 126. In addition, [199] does not discuss a fairness concept, which is most likely not as relevant in the context of non-safety communications.

Congestion Control

Congestion control approaches for VSC typically adapt the periodic transmission rate of CAMs/BSMs based on channel conditions. In [201], He et al. adapt the BSM generation rate based on MAC blocking, i.e., by stopping all transmissions if the CBR exceeds a certain threshold. Based on congestion events, vehicles adapt their transmission rate using AIMD. The protocol aims to maximize traffic generation fairness but does not take into account the awareness requirements of safety applications.

In [191], Khorakhun et al. introduce a binary control scheme which adapts the transmission *interval* of periodic messages based on CBR measurements. In [202], Busche et al. introduce SOURC, an extension of the work by Khorakhun et al., which has similar protocol elements as PULSAR, e.g., a two-hop piggybacking mechanism to disseminate CBR measurements. We discuss both approaches in greater detail in Sections 7.4.1 on page 205 and 7.4.2 on page 206, when we compare their performance against PULSAR's.

In [71], Kenney et al. introduce LIMERIC, a linear message rate control algorithm using CBR measurements as a feedback. The objective of LIMERIC is to improve the convergence behavior of binary control algorithms such as AIMD. We discuss the algorithm in greater detail in Section 7.4.3 on page 211, when we evaluate whether we can improve PULSAR's performance by simply replacing AIMD by LIMERIC without any further modifications to the protocol. Note that similar to the relative transmission rate based on AIMD in this thesis, LIMERIC supports the convergence to a weighted fair transmission rate for different (types of) vehicles [203]. However,

it is not clear if this weight can be adapted based on the driving context and how it could be translated into a safety benefit. In [204], Bansal et al. introduce a further extension of LIMERIC named EMBARC, which allows to transmit a message before the expiration of the periodic timer in the case that it is required to control the *tracking error* as perceived by other vehicles. That is, EMBARC is a hybrid of LIMERIC and an asynchronous awareness control as we discuss in the following.

Awareness Control

In the context of awareness control, the periodicity of CAMs/BSMs is often replaced by an aperiodic transmission of messages based on the driving context. In [205], Fukui et al. introduce DITRAC, which was likely the first scheme to adapt transmission rate based on vehicle dynamics. The algorithm transmits a CAM for every constant distance traveled, i.e., proportionally to vehicle speed[10].

In [23], Huang et al. present a rate control scheme which transmits BSMs based on the *suspected* tracking error neighboring vehicles have toward the transmitting vehicle. The basic idea is to transmit BSMs more frequently when the dynamics of the vehicle change. In between BSM receptions, vehicles use a model to estimate the current position of their neighbors. Based on its own transmitted BSMs, a vehicle can thus infer where its neighbors estimate its own position. If the resulting *suspected tracking error* $\tilde{\varepsilon}$, exceeds a certain threshold ε_{th}, the vehicle probabilistically transmits a new BSM in order to update its neighbors' knowledge. That is, the vehicle periodically invokes a Bernoulli experiment to decide whether or not to transmit a new BSM, of which the success probability is given by

$$p = 1 - e^{-\alpha|\tilde{\varepsilon}-\varepsilon_{th}|^2} \tag{3.16}$$

where α is a sensitivity parameter with a default value of 2 and the estimated tracking error $\tilde{\varepsilon}$ at time step k is calculated as

$$\tilde{\varepsilon}_k = (1-\xi)\tilde{\varepsilon}_{k-1} \tag{3.17}$$

where ξ is a second Bernoulli trial with the success probability $1-q$. That is, if the Bernoulli trial is successful, $\tilde{\varepsilon}_k$ is reset, i.e., a message reception is assumed. Otherwise, the estimated error accumulates according to the employed mobility model. The probability q corresponds to the estimated PER. The estimation of q is based on gaps in the sequence numbers of received packets. It should be noted that the approach of Huang et al. introduces a positive feedback loop, i.e., if channel quality degrades, more packets are generated, which in turn can congest the channel and thus degrade channel quality. As a countermeasure, the protocol employs a second feedback loop adapting transmit power based on CBR as given by Equation (3.14) on page 75.

In [165], Schmidt et al. explore *adaptive beaconing*, a scheme which transmits CAMs based the ego vehicle's own movement, i.e., similar to the approach of Huang et al., but

[10] A similar condition has been included in the CAM generation rules, cf. Section 2.1.3 on page 18

additionally takes into account vehicle density as well as the movement of neighboring vehicles. They argue that without taking other vehicles' movement into account, two vehicles with constant dynamics on a collision course may not receive a warning in time. Thus, they recommend a higher transmission rate with increasing crash probability, i.e., a lower Time To Collision (TTC).

In [206], van Eenennaam et al. suggest *reactive beaconing*, an approach which transmits a CAM with a distance-dependent delay after the reception of a CAM from a vehicle in front. The idea is to create cascades of transmissions which move against the flow of traffic, separated by a certain distance. However, the authors do not discuss the impact of their scheme on safety applications.

Conclusion

The number and timing of CAMs/BSMs transmissions influences not only channel load but also the up-to-dateness of the information available to safety applications. In this thesis, we focus on the periodic transmission of awareness messages, but introduce a concept how to adapt transmission rate within the boundaries provided by safety applications based on the current driving context. In future work, an additional aperiodic transmission could be integrated, following the example of the EMBARC approach [204].

3.4.7 Joint Control of Multiple Dimensions

While the majority of related approaches focused on the adaptation of a single control dimension, some works have introduced a joint control, typically focusing on transmit power and transmission rate. In [207], Baldessari et al. propose a combination of TPC and TRC designed to achieve fairness by allocating the same amount of channel resources to each node. That is, if a node increases its transmit power, it has to reduce its transmission rate and vice versa. However, the protocol does not take into account the impact on safety applications.

In [23], Huang et al. introduce a message transmission scheme which generates messages based on the movement of the transmitting vehicle and adapts transmission range with respect to channel load in terms of CBR. The two control loops run independently of each other and have thus been described in Sections 3.4.5 on page 73 and 3.4.6 on page 76, respectively. While the approach of Huang et al. shares the same objectives as the approach presented in this thesis, i.e., to control channel load while optimizing reception performance for safety applications, the authors have come to a different conclusion with respect to a control strategy.

The theoretical background of the approach by Huang et al. is described in [193], where Fallah et al. evaluate the Information Dissemination Rate (IDR), i.e., the number of packets received successfully by a vehicle's neighbors within a certain

range, with respect to transmission rate and transmit power. In a similar study to the one presented in this thesis, they find that the maximum IDR achievable for different combinations of transmit power and transmission rate is always the same, which confirms the findings in this thesis in terms of an optimal communication density. From this result, the authors conclude that IDR can be maximized by a separate control of both dimensions. Fallah et al. further study the relationship between IDR and CBR and find that it is a suitable feedback metric to maximize IDR.

Based on these conclusions, the authors present a design methodology for congestion control which adapts transmission rate based on the required tracking error and adapts transmit power with respect to channel load. However, the authors face the problem that the optimal choice of the transmit power depends not only on the current transmission rate but also on vehicle density. Since vehicle density is difficult to estimate in a real-world scenario, the authors "resort to a robust but suboptimal" design by adjusting the channel load between a minimum value of 0.4 CBR and a maximum value of 0.8 CBR, even though they identified the optimal value to be 0.65.

In this thesis, we take the opposite approach of selecting transmit power based on the driving context and adapting transmission rate based on channel load. In addition, we allow the possibility to set a vehicle's minimum and maximum transmission rate based on the current traffic situation. In contrast to the approach of Fallah et al., our approach supports the optimization for different distances individually for each vehicle. For example, the last vehicle in a traffic jam is likely to require a high transmit power as well as a high transmission rate in order to warn a fast approaching vehicle from behind. Since this vehicle would experience little change in its position, it would choose a low transmission rate based on the approach of Huang et al. In addition, its transmit power and therefore its communication range might be limited due to channel congestion. The combination of both might result in a dangerous situation in which a fast vehicle approaching the traffic jam does not receive position information in time for the driver to react. However, more information regarding the precise requirements of safety applications is required to make a final decision regarding which control strategy is better suitable to combine congestion and awareness control.

3.4.8 Summary and Conclusions

In this section, we reviewed a number of different congestion and awareness control approaches for vehicular networks which we selected as representatives for different optimization objectives and fairness criteria. An overview of selected related approaches is presented in Table 3.2 on the next page. The table shows the control dimension adapted by each approach as well as whether the approach addresses a fairness concept and/or the awareness requirements of safety applications. We will get back to the corresponding fairness definitions in Section 5.3 on page 119.

	Control dimension						Principles	
Approach	Tx power	Message rate	Modulation	Message size	CSTh	CW_{min}	Awareness	Fairness
D-FPAV [20]	X							X
SPAV [189]	X							X
Khorakhun et al. [191]	X							X
GRC [192]	X							X
OPRAM [208]	X						X	
Sepulcre et al. [93]	X						X	
ETPC [96]	X							
Baldessari et al. [207]	X	X						X
Huang et al. [23]	X	X					X	
DITRAC [205]	X	X					X	
Khorakhun et al. [191]		X						X
SOURC [202]		X						X
LIMERIC [71]		X						X
EMBARC [204]		X					X	X
Adaptive beaconing [165]		X					X	
Reactive beaconing [206]		X					X	
Kaul et al. [99]		X						
Camp et al. [209]			X					
CARS [161]			X					
Robinson et al. [157]				X			X	
PoOC [159]				X				
CbCO [68]				X				
CTA [182]					X		X	
Stanica et al. [183]					X			
Balon et al. [171]						X		
Mertens et al. [172]						X		
Stanica et al. [173]						X		
Eichler et al. [176]						X		
Taleb et al. [177]						X		
Zang et al. [178]						X		

Table 3.2: Classification of selected related approaches

Based on the findings of the related work as well as our own considerations expressed in this chapter, we focus on the control of transmission rate and transmit power in this thesis and do not further consider the adaptation of the other potential control dimensions. We leave the potential integration of further degrees of freedom into congestion control for future work.

4

Analysis of Optimization Problem and Solution Space

As illustrated by the previous chapters, the design of congestion control for Vehicle Safety Communications (VSC) is a multi-dimensional problem with at least six degrees of freedom. In Chapter 3 on page 45, we discussed why we focus on two degrees of freedom in this thesis, i.e., transmission rate and transmit power. Both dimensions directly influence not only the resulting channel load but also the performance of safety applications. We also saw that so far, existing approaches typically focused either on the control of channel load *or* on the adaptation of transmission parameters to meet safety applications' requirements. In addition, existing congestion control approaches generally did not define what the optimal outcome of the adaptation would be and why their protocol would converge to the desired result.
While typically congestion control is tackled from a sender's perspective, the functionality of the envisioned safety applications depends on what has been received by the relevant neighboring vehicles. In this chapter, we thus take on a receiver's point of view when addressing the following research questions:

1. What is a suitable optimization criterion to compare different combinations of transmission rate and transmit power?
2. Which parameter combinations optimize the selected criterion?
3. How much is lost without optimization?

Parts of the results presented in this chapter have been previously published in [26]. They lay the foundation for the design methodology introduced in Chapter 5 on page 103 and the resulting congestion control protocol described in Chapter 6 on page 135.

4.1 Optimization Criterion

In every optimization study, the choice of the optimization criterion plays an essential role. In this work, our objective is to prevent the channel from getting into an overload situation while optimizing reception performance for safety applications. On second thought, the first objective follows from the second one, since reception performance deteriorates in an oversaturated channel.

Intuitively speaking, the degree of "awareness" a vehicle has towards its neighbors has both a spatial and a temporal aspect. On the one hand, it depends on which neighbors matter for the safety of the particular vehicle. On the other hand, it depends on how long ago an update from a particular neighbor has been received. Due to the high mobility of vehicles, both aspects of awareness are constantly subject to change. In addition, different safety application can have different requirements and a merging of these requirements may be required, cf. [93].

While the specification of the requirements of different safety applications is still an ongoing field of research, e.g., [21][24], our focus in this chapter is on understanding general tendencies when adapting communication parameters rather than on defining which setup is optimal for a particular application in a particular traffic situation.

4.1.1 Spatial Aspect: Target Distance

In order to assist the driver in avoiding a potential collision, a safety application needs to issue a warning in time for the driver to react and initiate a countermeasure. That is, the application has to take into account the driver's perception-reaction time as well as the time it takes to execute the required action itself, i.e., to decelerate to the speed of the conflicting vehicle and to potentially come to a full stop.

From a transmitting vehicle's perspective, this means that it has to make sure that its messages are received within a certain distance, which we denote as *target distance* d_t. Assuming a monotonically decreasing probability of message reception with respect to distance, reception performance *within* target distance will be equally good as or better than *at* target distance.

In a highway situation for example, a vehicle's target distance could be approximated as

$$d_t = max \begin{cases} \frac{v^2}{2a} + (t_{rct} + t_{sys})v \\ \frac{{v_{rel}}^2}{2a} + (t_{rct} + t_{sys})v_{rel} \end{cases} \tag{4.1}$$

where v is the transmitting vehicle's current speed, a is a maximum deceleration value, t_{rct} is the perception-reaction time of the driver, t_{sys} is the system latency and v_{rel} is the maximum speed differential to a neighboring vehicle [26]. Note that Equation (4.1) is meant as an example how a target distance could be calculated. Very similar calculations have been previously presented in related works, where the resulting distance is typically denoted as a *warning distance*, e.g., [24][210]. The first part of the equation corresponds to the transmitting vehicle's *braking distance* d_{brk} assuming a constant deceleration and reflects the idea that all neighbors ahead and within this distance should be aware of the transmitting vehicle's presence, e.g., to prevent unsafe lane changes. The second part of equation primarily applies to neighbors behind the transmitting vehicle and reflects the *deceleration distance* d_{dec} of the neighbor with the highest relative speed, i.e., the distance it takes this neighbor to decelerate to the transmitting vehicle's speed, again assuming a constant deceleration.

Figure 4.1 on the following page schematically illustrates a German freeway situation with a forming traffic jam in one driving direction and free-flow conditions in the other driving direction. For some selected vehicles labeled A to J, the figure indicates their current speed, maximum relative speed as well as the resulting target distance according to Equation (4.1). The example illustrates two properties of the target distance concept:

1. High target distances can occur in areas with high vehicle density. For example, vehicle G requires a high target distance while passing a traffic jam.

2. A vehicle's target distance depends not only on its own movement, but also on the movement of its neighbors. For example, vehicles A and B have the same speed, but since B is located at the end of the traffic jam, it needs a higher target distance in case another vehicle approaches at high speed.

While Equation (4.1) is meant as an example of how a target distance could be derived from basic kinematic equations, other works have studied this topic in greater detail for different safety applications, e.g., [24] and [210]. For the purpose of this thesis, it is less important how exactly a vehicle's target distance is calculated. Rather, it matters that such a distance can be derived individually for each vehicle depending on its driving context and that is has the properties discussed above.

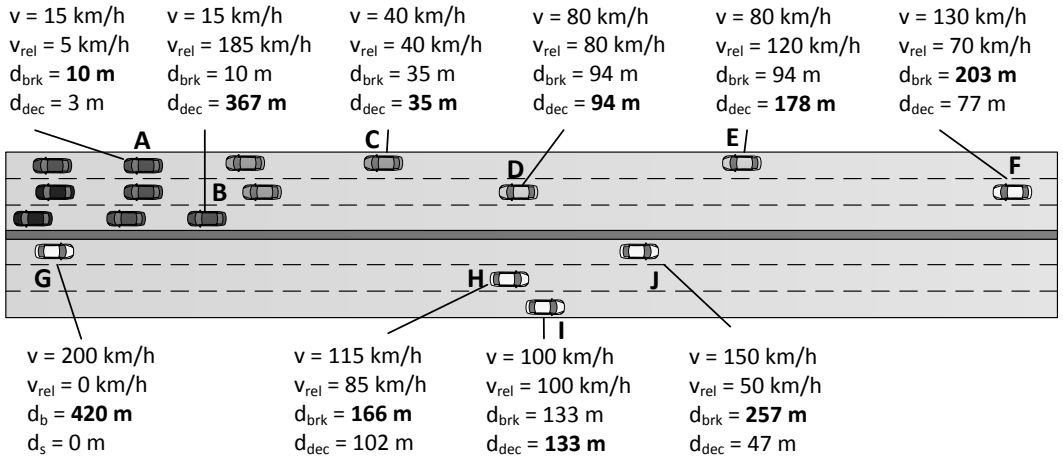


Figure 4.1: Schematic illustration of an a German freeway situation with resulting target distances (bold font); calculated based on Equation (4.1) using v_{max} = 200 km/h, a = -5 m/s², $t_{rct} + t_{sys}$ = 2 s

4.1.2 Temporal Aspect: Inter-Reception Time

To assess the probability of a collision with a particular neighboring vehicle, the vehicle running a particular safety application, i.e., the *ego vehicle*, needs to receive regular updates on this neighbor's position, speed and heading. A frequently employed metric for the up-to-dateness of such information is the Packet Inter-Reception Time (IRT), which denotes the time between successive message receptions from a particular transmitter, cf. Section 2.2.3 on page 30.

For a reliable communication channel, the resulting IRT would simply correspond to the transmission interval. However, since wireless communication is inherently unreliable, the IRT observed by the receiver depends on the number of subsequently lost packets N. If the probability of packet reception p_{Rx} is assumed to be constant, N follows a geometric distribution:

$$P(N = n) = p_{Rx}(1 - p_{Rx})^n, n \in \mathbb{N} \tag{4.2}$$

If we additionally assume a constant transmission interval t_{Tx} (and thus a constant transmission rate $r = 1/t_{Tx}$), it follows that the observed IRT is $t_{Rx} = (n + 1)t_{Tx}$. Given these assumptions, the IRT t_{Rx} follows a geometric distribution as well with the following Probability Mass Function (PMF):

$$P(I = t_{Rx}) = p_{Rx}(1 - p_{Rx})^{t_{Rx}/t_{Tx}-1} \tag{4.3}$$

For the purpose of this analysis, we would like to compare the IRT distributions at each target distance resulting from different transmission parameter setups and to select the configuration with the "best" reception performance for safety applications. In the following, we discuss three different ways of doing so, i.e., to compare Cumulative Density Function (CDF) values, percentiles and averages.

Comparing CDF Values

Assuming a constant probability of reception p_{Rx} and transmission interval t_{Tx}, the Cumulative Density Function (CDF) of IRT is given by

$$P(I \leq t_{Rx}) = 1 - (1 - p_{Rx})^{\lfloor t_{Rx}/t_{Tx} \rfloor} = 1 - (1 - p_{Rx})^{\lfloor t_{Rx} r \rfloor} \tag{4.4}$$

Note that while derived differently, the T-window reliability metric introduced by Bai et al. in [7] corresponds to above CDF with $t_{Rx} = T$, cf. Equation (2.1) on page 31. T-window reliability is defined as the probability of receiving at least one message from a particular transmitter within a time window T. While not shown explicitly in [7], the floor function in Equation (4.4) is required to account for the atomicity of packets if $t_{Rx}/t_{Tx} \notin \mathbb{N}$.

As illustrated by Table 2.7 on page 33, many state of the art awareness metrics are related to T-window reliability and thus require the definition of a time period T. This dependence, however, makes it difficult to apply them (and the CDF of IRT in general) to our study, since configurations that maximize Equation (4.4) for one value of t_{Rx} can result in a low CDF value for another value of t_{Rx}.

Comparing Percentiles

To eliminate the dependency on the required IRT t_{Rx}, an alternative to maximizing the CDF value of IRT could be to minimize its inverse, i.e., its kth percentile:

$$P_k = t_{Tx} \left\lceil \frac{log(1-k)}{log(1-p_{Rx})} \right\rceil \tag{4.5}$$

Minimizing the (kth percentile of) IRT at the selected target distance reflects the maximum freshness of neighbor information and thus seems like a suitable optimization criterion. However, since P_k is a step function, it is very sensitive to small changes of p_{Rx}. Using $k = 99$ for example, $p_{Rx} = 0.86467$ results in $P_{99} = t_{Tx}$, while $p_{Rx} = 0.86466$ results in $P_{99} = 0$. As a result, the metric can be prone to noisy behavior, especially if p_{Rx} is taken from simulations.

Comparing Averages

As an alternative to the kth percentile, the lowest IRT at a particular target distance could be determined based on the distribution's mean

$$\mu = \frac{t_{Tx}}{p_{Rx}} \tag{4.6}$$

While comparing averages has the advantage of eliminating the need to compare step functions and thus reduces noise, it has two disadvantages. First, since the skewness of the geometric distribution is nonzero, its mean does not correspond to a fixed percentile. From a safety point of view, it is thus difficult to give any

"guarantees". Second, the comparison of averages does not take into account variances. A distribution with a higher mean but a lower variance may be preferable from a safety application's perspective, while the one with lower average would be selected according to this criterion.

Conclusion

The previous discussion shows that there appears to be no ideal metric for our purpose of comparing IRT distributions at a particular distance with respect to their benefit for safety applications. Thus, we conducted the optimization study for all three variants, i.e., maximizing CDF values, minimizing percentiles and minimizing averages. We present our findings in the following.

4.2 Optimized Parameter Configurations

The analysis in Section 4.1.2 on page 86 has shown that IRT depends on two factors, i.e., transmission rate and reception probability. While the former is straightforward to control, the latter depends on many factors, e.g., transmit power, radio-wave propagation and the level of interference. The level of interference, in turn, depends on channel load, which among other things depends on transmission rate and transmit power. A priori, it is thus difficult to say how transmission rate and transmit power should be configured in order to optimize IRT at the intended target distance.

In this section, we study the impact of a wide range of combinations of transmission rate and transmit power on the resulting IRT at a certain distance. Our objective is to find common characteristics of the identified optimal parameter configurations in order to derive a strategy for congestion control. For presentation clarity, we thereby focus on the minimization of average IRT at a certain distance as the optimization criterion, since it led us to the same conclusions as the minimization of the kth percentile of IRT, albeit with less noisy results. We additionally present the results for the kth percentile where appropriate. Note that the maximization of the CDF for a given IRT leads to different conclusions with respect to a control strategy, which we discuss in Section 5.1.4 on page 109.

4.2.1 Methodology

To analyze and understand the fundamental implications of adapting either control dimension, we focus on a simple scenario consisting of a long road with a single lane of equidistant vehicles. To exclude boundary effects, we model distance calculation and radio-wave propagation in a pseudo-circular setup, i.e., the last vehicle of the road considers the first one to be its neighbor. To prevent interference resulting from the circular setup, we use a road length between 10 km and 25 km, depending on

the applied transmit power. While the modeled nodes are static, we account for the impact of vehicle movement on channel load by simulating three different vehicle densities. Our objective is to find a control strategy that works independent of vehicle density. Table 4.1 on the next page summarizes the considered parameter space.

The results presented in this section are based on simulations using the network simulator ns-2 in the configuration described in Appendix A on page 223. While analytical models for the behavior of Carrier Sense Multiple Access (CSMA) in a chain topology exist, e.g., [193][211], they typically do not account for heterogeneous transmission power levels for different nodes which we apply in Section 4.2.4 on page 93. The major challenge in calculating the resulting IRT at a particular distance from the sender is to analytically determine the probability of message reception, which depends on many factors such as the placement of nodes, their transmit power levels, fading, etc. We leave an analytic confirmation of the results presented in this section for future work.

4.2.2 Optimization for One Group of Vehicles

In the first part of our search for the set of transmission parameters optimizing reception performance at a certain target distance, we focus on a uniform configuration of nodes in order to understand the underlying principles. In Section 4.2.4 on page 93, we cross-check whether these observations hold if two intermixed groups of vehicles optimize for different target distances.

Figure 4.2 on page 91 illustrates the observed average IRT for two different vehicle densities ρ (rows of subfigures) and target distances d_t (columns of subfigures). The x-axis and y-axis of each plot reflect the evaluated parameter space of transmission rate and transmit power. The arrows drawn on the xy-plane point toward the minimum observed average IRT for this particular vehicle density and target distance. We can see that for each ρ and d_t, a different combination of transmit power and transmission rate optimizes average IRT.

For the purpose of this study, we have repeated this selection process for all combinations of ρ and d_t within the considered parameter space. The result is depicted in Figure 4.3 on page 92, which illustrates the identified optimal parameter combinations minimizing average IRT at a certain distance, plotted in separate diagrams. It is important to note that both subfigures are meant to be interpreted together. For example, in Figure 4.2b on page 91, we can see that the parameter combination minimizing average IRT for ρ = 50 vehicles/km and d_t = 300 m is 24 dBm and 11 Hz in our setup. This result is represented by an entry of 24 dBm at 300 m distance and 50 vehicles/km in Figure 4.3a and by an entry of 11 Hz at 300 m distance and 50 vehicles/km in Figure 4.3b.

The representation as separate diagrams facilitates a characterization of the identified optimal combinations across vehicle densities. Comparing Figures 4.3a and 4.3b, we

Parameter	Evaluated range
Transmit power [dBm]	0, 2, ..., 30
Transmission rate [Hz]	0.5, 1, 1.5, ..., 4, 5, 6, ..., 20
Vehicle density [vehicles/km]	50, 100, 200
Simulation runs	10

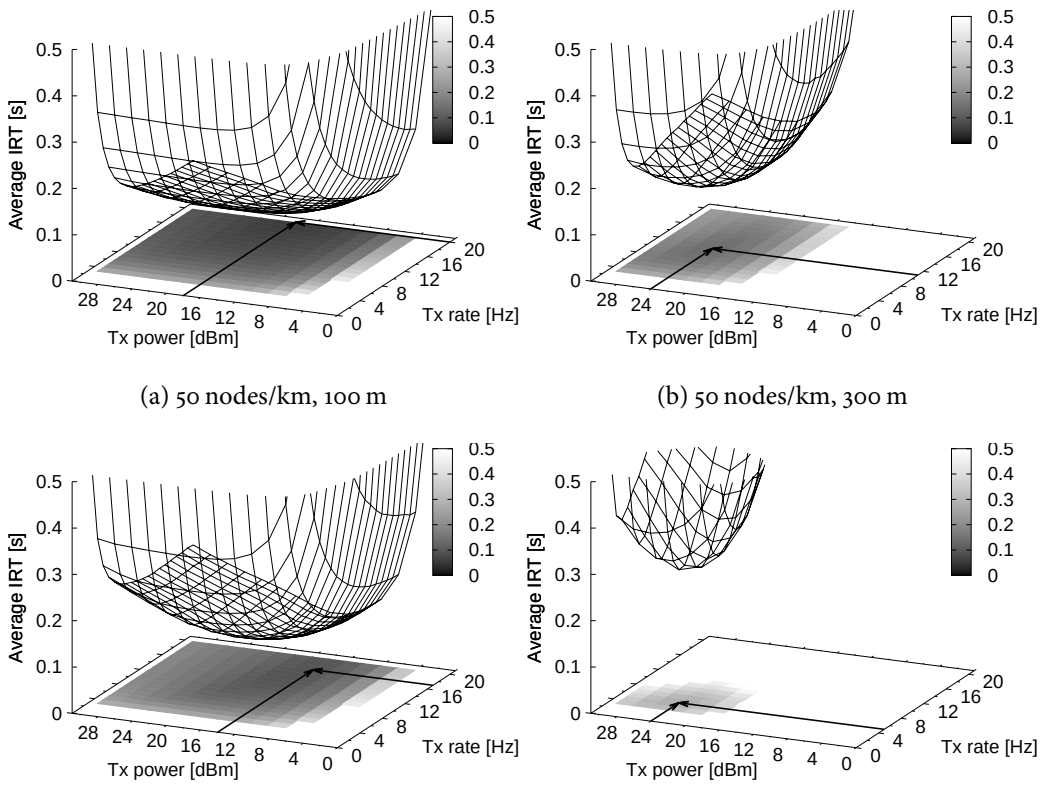
Table 4.1: Simulation parameters and their evaluated range

can see that *while the identified optimal transmit power at each distance is very similar across vehicle densities, the corresponding transmission rates typically do not overlap.* Note that the outliers in the transmit power curve at close distances for 50 vehicles/km and for far distances for 200 vehicles/km can be traced back to the limitations of our parameter space. In the former case, a higher transmission rate than 20 Hz combined with a lower transmit power would further reduce the observed average IRT. In the latter case, a higher granularity in transmission rate would be required.

The sensitivity of the identified optimal parameter combinations is illustrated by Figure 4.4 on page 93, which additionally depicts all parameter configurations which resulted in an average IRT of up to 100 ms above the identified minimum for each target distance and vehicle density. In Figure 4.4a on page 93, we can see that the sensitivity of the transmit power is generally lower for close distances. However, it should be noted that at close distances, the fixed offset of 100 ms corresponds to a higher fraction of the global optimum than for farther distances, where the observed minimum IRT is higher. A similar trend of a lower sensitivity for closer distances can also be observed for the corresponding transmission rates in Figure 4.4b. However, for the transmission rate the vehicle density appears to be the dominating factor in terms of sensitivity. We generally observe a lower sensitivity for lower vehicle densities.

Figure 4.5 on page 94 is a counterpart of Figure 4.3 on page 92 using the minimization of the 99th percentile of IRT as the optimization criterion. We observe the same trends as for average IRT in the sense that the identified optimal transmit power values for each distance are very similar across vehicle densities, while the identified optimal transmission rates do not overlap. However, we also observe some differences. First, the identified optimal transmit power values per distance are generally higher than with average IRT, while the transmission rates are generally lower. Second, there are more outliers, which we attribute to the fact that the percentile of IRT is a step function. In a nutshell, we conclude that both optimization criteria generally lead to the same conclusions.

Figure 4.6 on page 95 illustrates the Packet Delivery Ratio (PDR) experienced at each optimal parameter configuration. We can see that many combinations result in an Packet Delivery Ratio (PDR) of 0.6 to 0.7. For the 99th percentile, we observed a similar outcome, but at a higher level of 0.8 to 0.9. This finding might indicate that



(a) 50 nodes/km, 100 m

(b) 50 nodes/km, 300 m

(c) 100 nodes/km, 100 m

(d) 100 nodes/km, 300 m

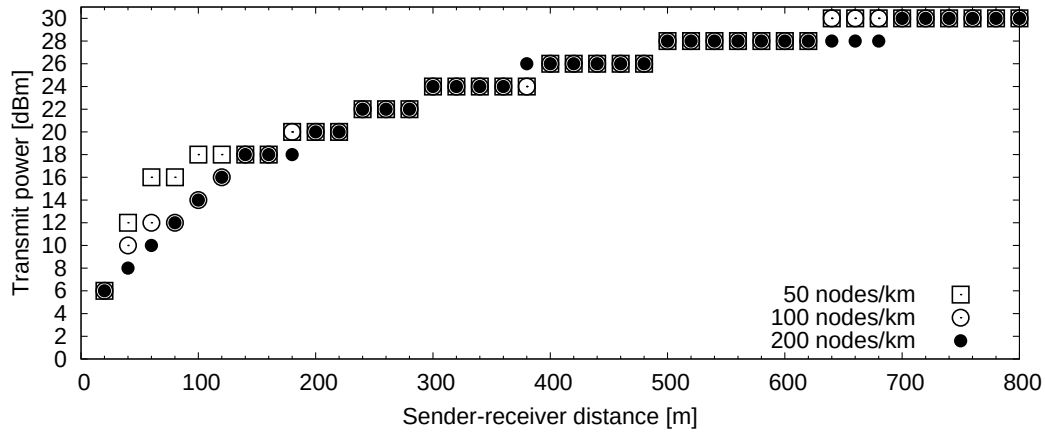
Figure 4.2: Examples of the observed average IRT values for a certain vehicle density and target distance; the arrows in the subfigures point to the optimal value [26]

there exists something like a *globally optimal PDR*. We leave a further exploration of this topic based on an analytical model for future work.

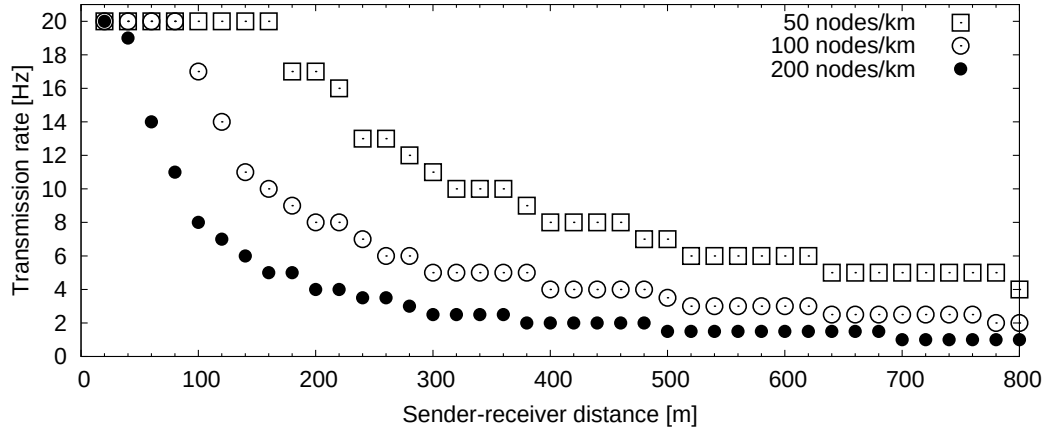
4.2.3 Optimization for a Certain Channel Load

Figure 4.7 on page 96 illustrates the channel load in terms of Channel Busy Ratio (CBR) resulting from the identified optimal parameter combinations minimizing average IRT at a certain distance. We can see that except for the outliers discussed earlier, all optimized combinations result in a CBR near 0.9. However, in practice it may be desirable to restrict channel input further, e.g., to keep the system near the "knee" rather than the "cliff" or to reserve a certain fraction of the available bandwidth for event-driven emergency messages. We discuss the selection of a convergence target for congestion control in Section 5.2 on page 111.

To find out how such a restriction would impact the characteristics of the parameter combinations optimizing reception performance, we repeated the analysis for three exemplary CBR values of 0.3, 0.5 and 0.7. In the first step, we determined the transmis-



(a) Transmit power

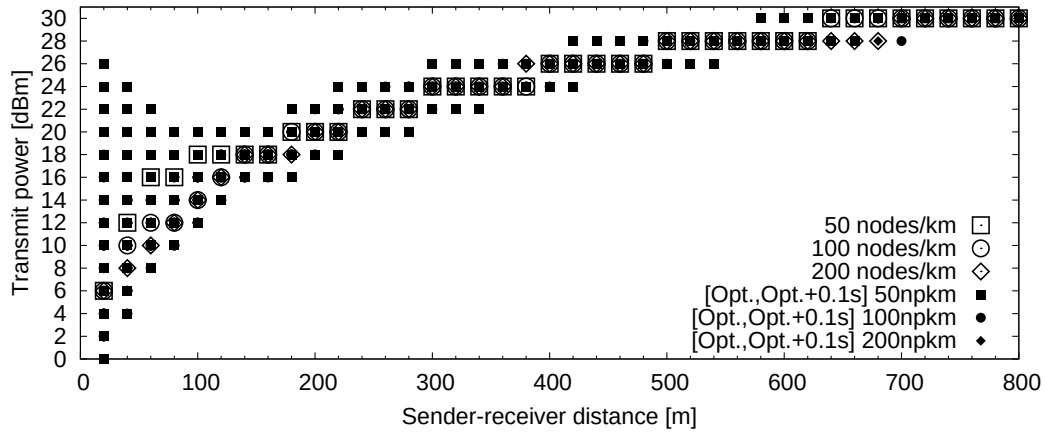


(b) Transmission rate

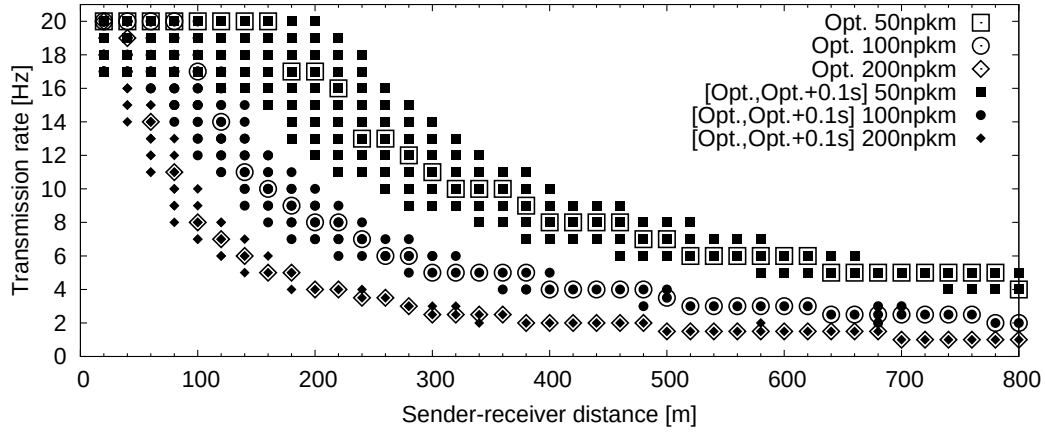
Figure 4.3: Parameter combinations optimizing average IRT

sion rate which resulted in the desired CBR for each of our 16 transmit power levels in the three considered vehicle densities. Then, we evaluated which setup minimized the average IRT for each target distance and vehicle density. While our results indicate that the independence of the optimized transmission power from vehicle density still holds, we observed differences in the optimized values across CBR targets.

Figure 4.8 on page 97 illustrates the transmit power and transmission rate combinations minimizing average IRT for a vehicle density of 100 vehicles/km. We can see that the optimized transmit power curve has a negative offset for each lowered CBR value. In an environment with a lower channel load, i.e., less interference, less received signal strength is necessary to achieve the required Signal to Interference and Noise Ratio (SINR) for reception. Thus, a lower transmit power is required. We conclude here that the optimized transmit power not only depends on the target distance, but also on the targeted channel load.



(a) Transmit power



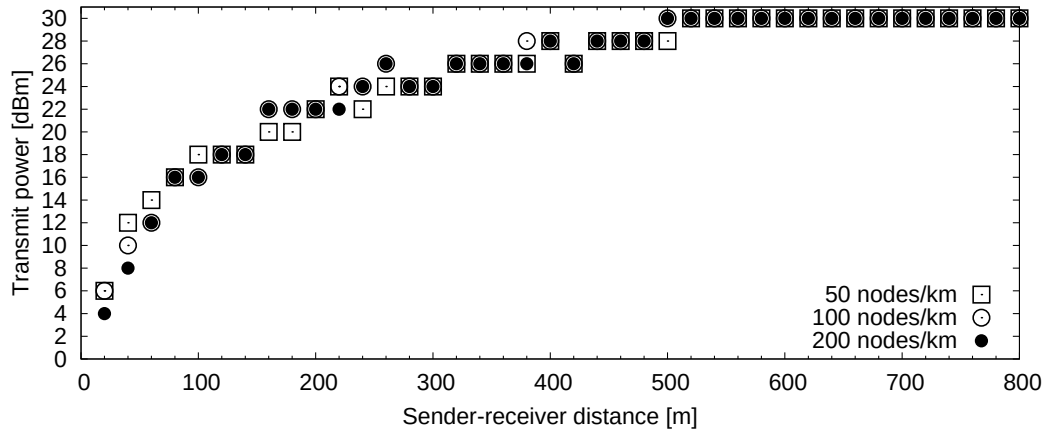
(b) Transmission rate

Figure 4.4: Sensitivity of the parameter combinations optimizing average IRT

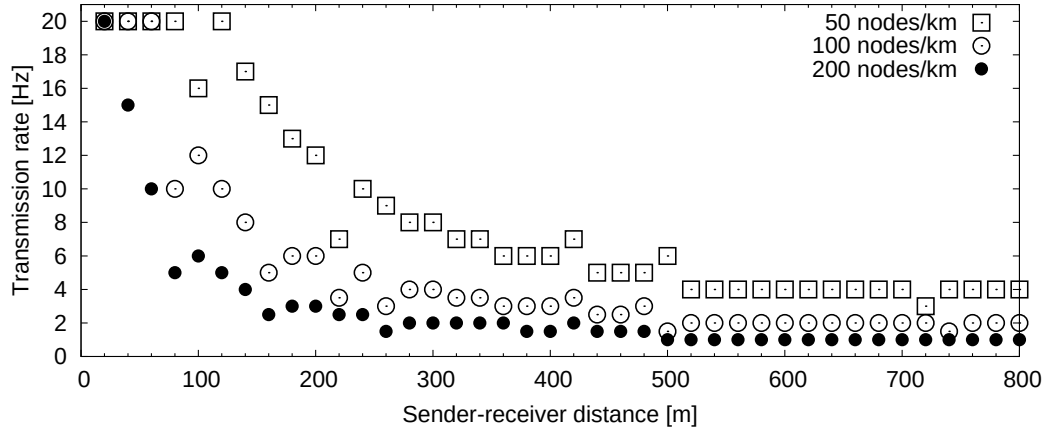
4.2.4 Optimization for Two Intermixed Groups of Vehicles

So far, by assigning uniform transmission parameters to all vehicles, we implicitly assumed that all vehicles share the same target distance d_t for which reception performance should be optimized. However, as discussed in Section 4.1.1 on page 84, this is typically not the case. In this section, we therefore extend our analysis to two (groups of) vehicles optimizing for different target distances at the same time. The question is: Can the optimized reception parameters obtained for the case of one group of vehicles be transferred to the case of two groups with different optimization objectives?

To answer this question, we split the vehicles in the previously applied circular road scenario into two intermixed groups configured with different transmission parameters, every other vehicle along the road belonging to one group. The size ratio of 1:1 was chosen to avoid bias. Since an evaluation of the entire parameter space would have resulted in more than 100,000 possible combinations, we used a fixed CBR of



(a) Transmit power



(b) Transmission rate

Figure 4.5: Parameter combinations optimizing the 99th percentile of IRT

0.7 to reduce complexity. More specifically, we conducted the study in two steps. In the first step, we obtained the configurations to be evaluated. For this purpose, we selected a transmission rate and transmit power for the first group as well as a transmit power for the second group according to Table 4.1 on page 90. Then, we adjusted the transmission rate of the second group to reach a CBR of 0.7. In the second step, we conducted the actual simulation and evaluated the resulting IRT of the messages sent by each group and received by any vehicle located at the sending group's target distance.

In contrast to the scenario with only one optimization objective, the question of which parameter configuration optimizes reception performance for both groups is not straightforward to answer, since it depends on a fairness definition. Depending on the driving context, it may be reasonable to assign a higher average IRT to one of the groups. For this reason, we focus on the Pareto optimal transmission parameter configurations, i.e., those for which the average IRT of one group cannot be further lowered without increasing the average IRT for the other group.

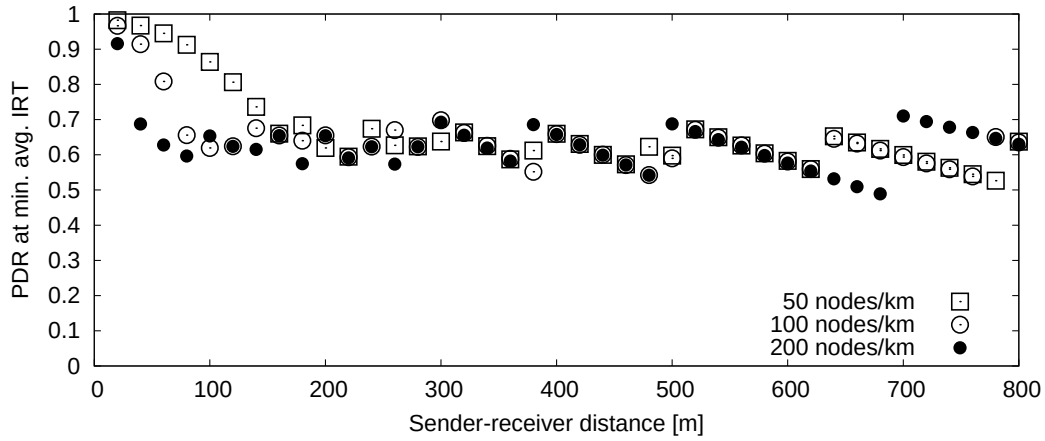


Figure 4.6: Packet Delivery Ratio (PDR) corresponding to the parameter configurations optimizing average IRT

Figure 4.9 on page 98 illustrates the resulting average IRT values for our two groups using exemplary target distances of 100 m and 300 m. Note that due to the group ratio of 1:1, it does not matter which group uses which target distance and the figure thus summarizes both variants. In the following, we refer to the group with d_t = 100 m as group 1 and to the group with d_t = 300 m as group 2. The figure consists of three subfigures which correspond to different vehicle densities. Each dot corresponds to the outcome in terms of average IRT at the respective target distance of each group and represents a quadruple of transmission parameters, i.e., the transmission rate and transmit power of each of the two groups. To give an example, in Figure 4.9c the combination of 2 Hz and 12 dBm for group 1 and 4.24 Hz and 20 dBm for group 2 results in an average IRT of 0.938 s for group 1 and of 0.381 s for group 2. The Pareto optimal combinations are indicated as medium sized solid black dots, located along the lower left border of the cloud of dots.

For the purpose of our study, the question arises which transmission parameter combinations resulted in the identified Pareto optimal points. Figure 4.10 on page 99 illustrates a parallel coordinate plot of the transmission parameter quadruples resulting in a Pareto optimal result for our exemplary target distances of 100 m and 300 m. The figure consists of four axes, connecting each quadruple by a thin line. The three subfigures correspond to different vehicle densities. In the figure, we observe a wide range of transmission rates for each group, while the transmit power values are clustered near 12 to 14 dBm for group 1 and near 20 dBm for group 2, independent of vehicle density. Approximately 80% of all observed Pareto optimal parameter quadruples use these transmit power values. Taking a look back at Figure 4.8 on page 97, we observe that for a CBR of 0.7, we had identified transmit power values of 12 dBm and 20 dBm to be optimal for d_t = 100 m and d_t=300 m, respectively.

In Figure 4.9 on page 98, the combinations using 12 dBm and 20 dBm for the respective groups are highlighted using large empty circles. We can see that for all three vehicle

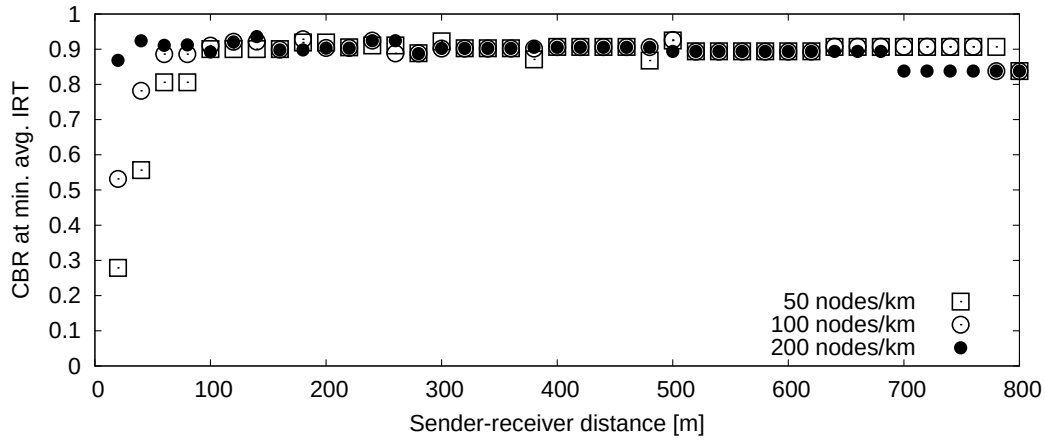


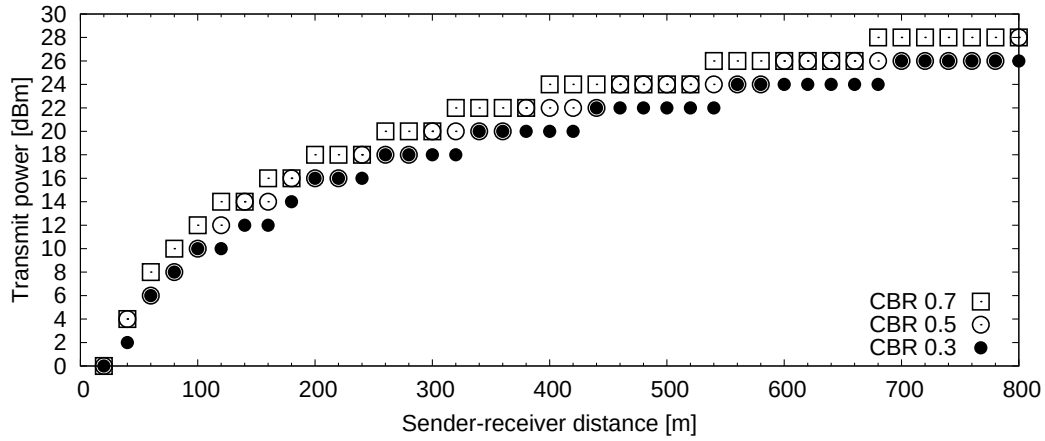
Figure 4.7: Channel Busy Ratio (CBR) corresponding to the parameter configurations optimizing average IRT

densities, these points are located near the lower left border of the cloud of dots. By interpolating them, we could approximate the Pareto frontier, i.e., the border between feasible and infeasible allocations. The example in Figure 4.9c illustrates how the adaptation of transmission rate can be used to reach different points along the Pareto frontier. We further notice in the figure that some points have been classified as Pareto optimal, even though they are not located directly at the Pareto frontier. This result is an artifact of the limited granularity of our setup and is one of the main reasons why 20% of the identified Pareto optimal configurations use other transmit power values. Another reason is illustrated by Figure 4.9a, where due to the maximum transmission rate of 20 Hz, no combination using 12 Hz and 20 Hz extends to the upper part of the figure.

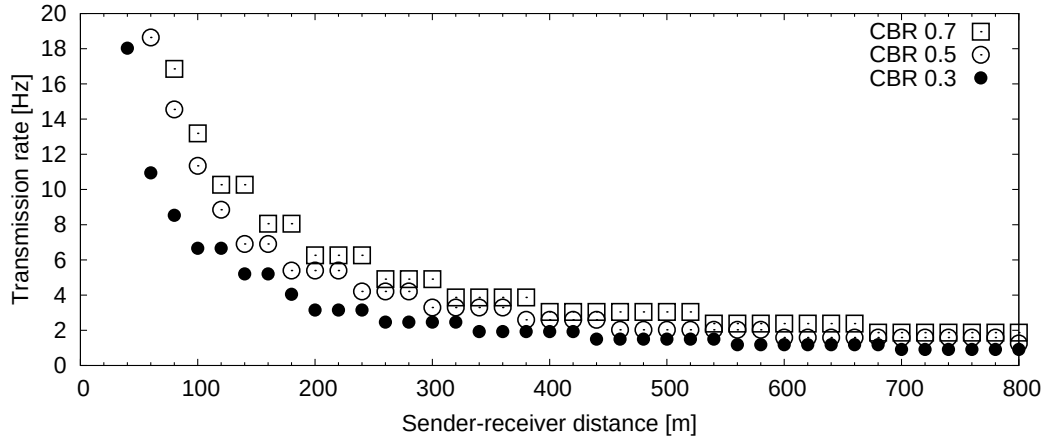
Figure 4.11 on page 100 illustrates the fraction of the identified Pareto optimal combinations using the previously determined optimal transmit power values for different combinations of target distances. We thereby allowed a tolerance of ±2 dB, which corresponds to the granularity of our setup. We can see that typically, the majority of Pareto optimal points, up to 100% in some cases, uses a similar transmit power as previously identified as optimal in the case of uniform transmission parameters. The discrepancies can largely be explained in the context of the limitations of the considered parameter space as laid out above.

4.3 Reception Performance Without Congestion Control

Having identified the transmission parameter combinations optimizing reception performance in our setup, the question arises how transmission rate and transmit power should be adapted to converge to these values. To motivate the design of congestion control in the next chapter, we conclude the current chapter by quantifying



(a) Transmit power

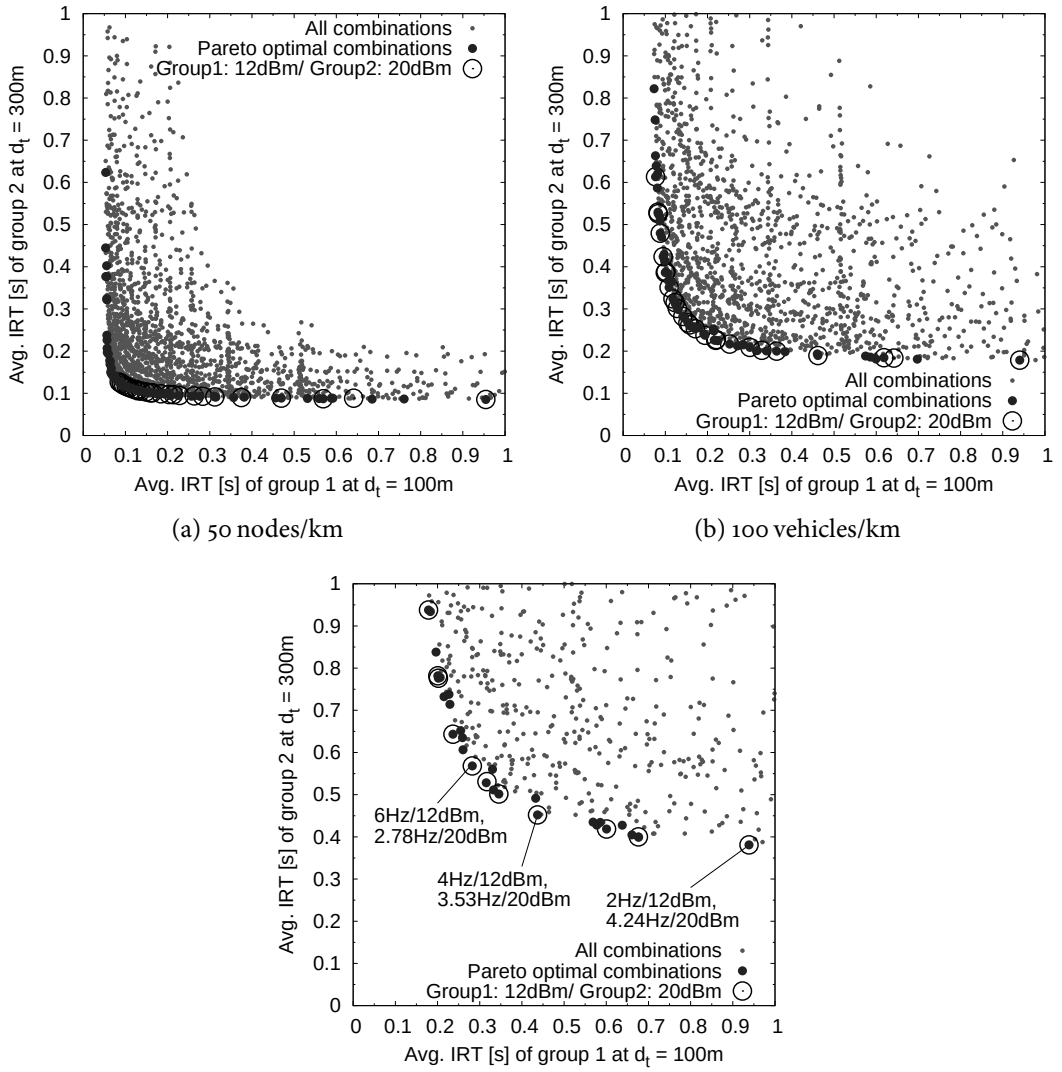


(b) Transmission rate

Figure 4.8: Parameter combinations optimizing average IRT at different CBR levels (vehicle density 100 vehicles/km)

the difference in terms of average IRT if no adaptation at all is performed, i.e., if a fixed set of transmission parameters is applied.

Figure 4.12 on page 101 illustrates the observed average IRT per distance compared to the global minimum for three different vehicle densities and three different fixed transmission parameter combinations. The figure is complemented by Figure 4.13 on page 102, which illustrates the corresponding *difference* in average IRT between the respective configuration and the global minimum. We observe that each combination optimizes average IRT for a different combination of vehicle density and target distance. For example, the combination of 10 Hz and 24 dBm results in a near-optimal outcome for a vehicle density of 50 vehicles/km and a target distance between 300 m and 400 m. For higher vehicle densities, however, reception performance at this distance is significantly impaired due to channel congestion. At 200 vehicles/km, the



(a) 50 nodes/km

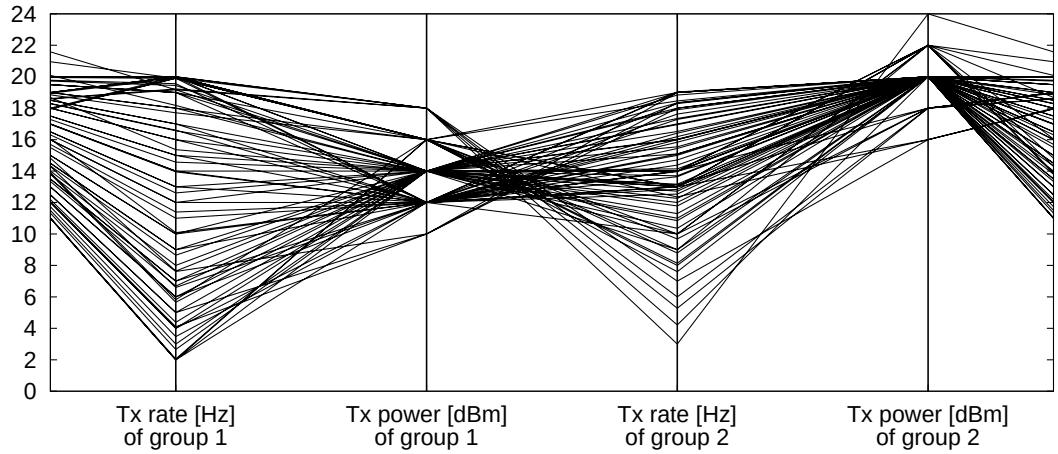
(b) 100 vehicles/km

(c) 200 vehicles/km

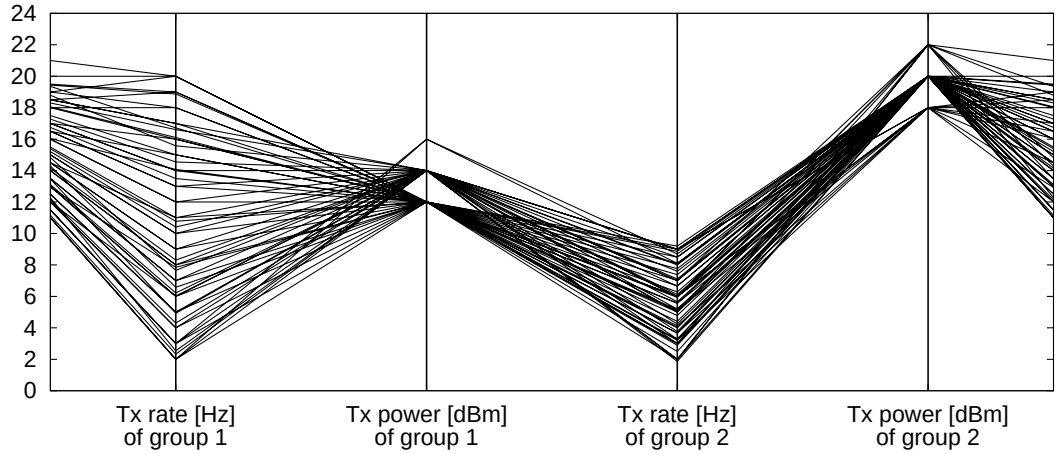
Figure 4.9: Average IRT results for two groups of vehicles optimizing for target distances 100 m and 300 m, including Pareto optimal points and combinations using the transmit power previously identified as optimal for the respective target distance

channel is so congested that the minimum difference between the observed value and the optimum exceeds 100 ms.

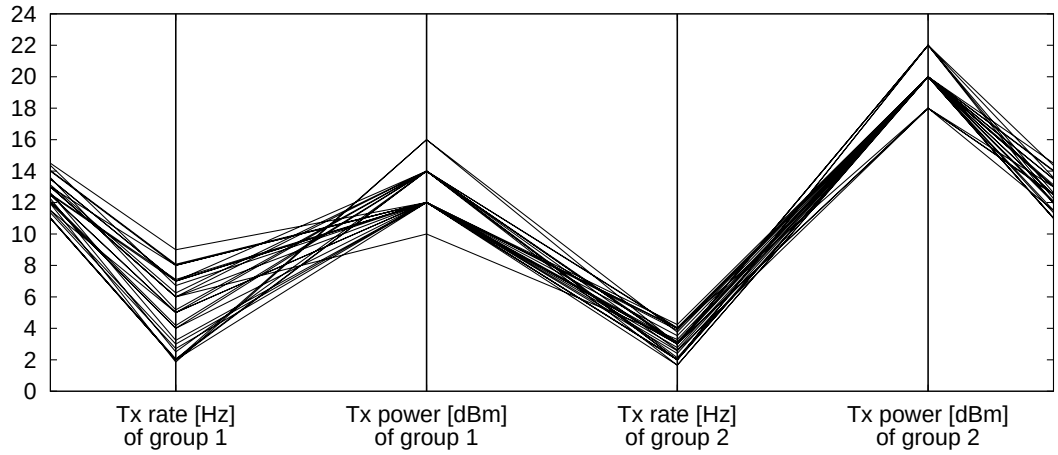
The other two examples of fixed transmission parameters illustrate that neither a reduction of the (fixed) transmit power nor a reduction of the (fixed) transmission rate can provide a near-optimal outcome for all target distances and vehicle densities. The suboptimal outcome is caused by an overutilization of the channel for high vehicle densities as well as by an underutilization of the channel for low vehicle densities. We conclude that, in order to use the channel efficiently, an adaptation



(a) 50 vehicles/km



(b) 100 vehicles/km



(c) 200 vehicles/km

Figure 4.10: Parameter combinations resulting in Pareto optimal points

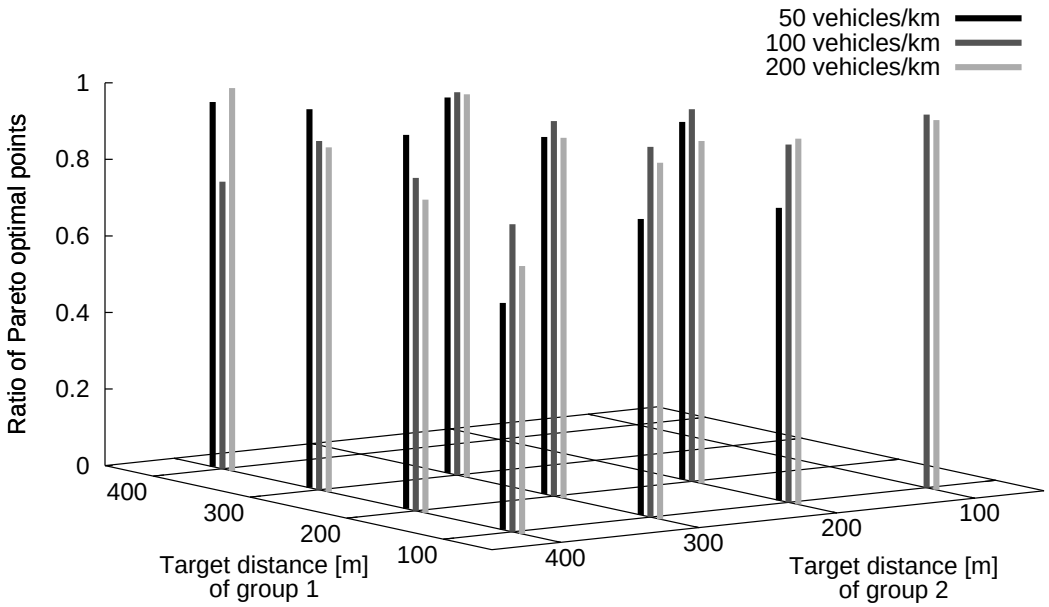
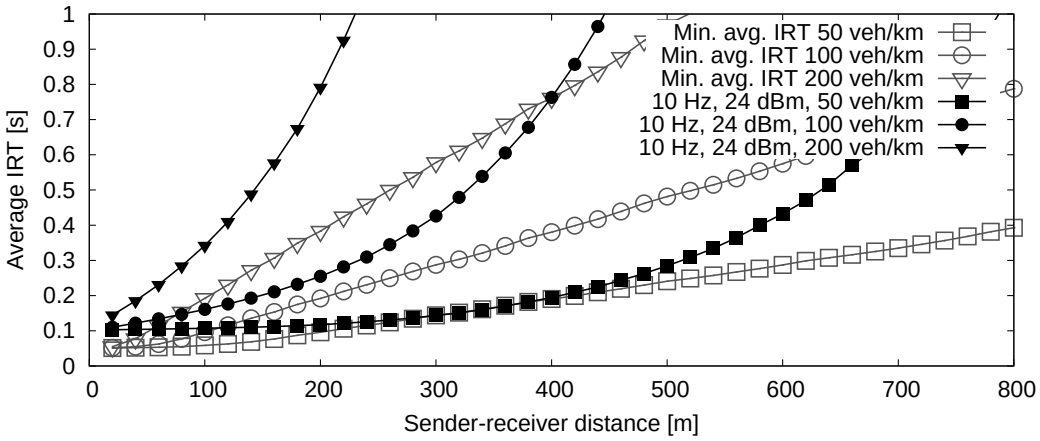
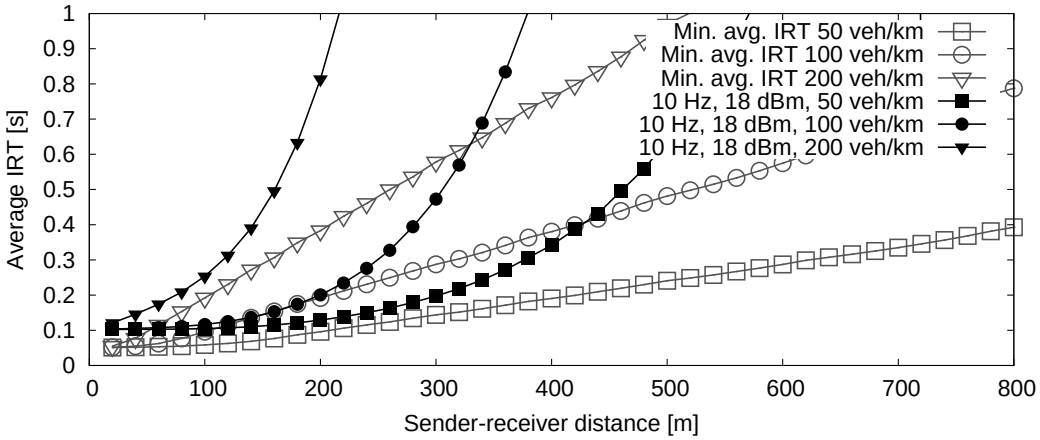


Figure 4.11: Ratio of Pareto optimal setups using the transmit power (±2 dB) identified as optimal in the homogeneous case [26]

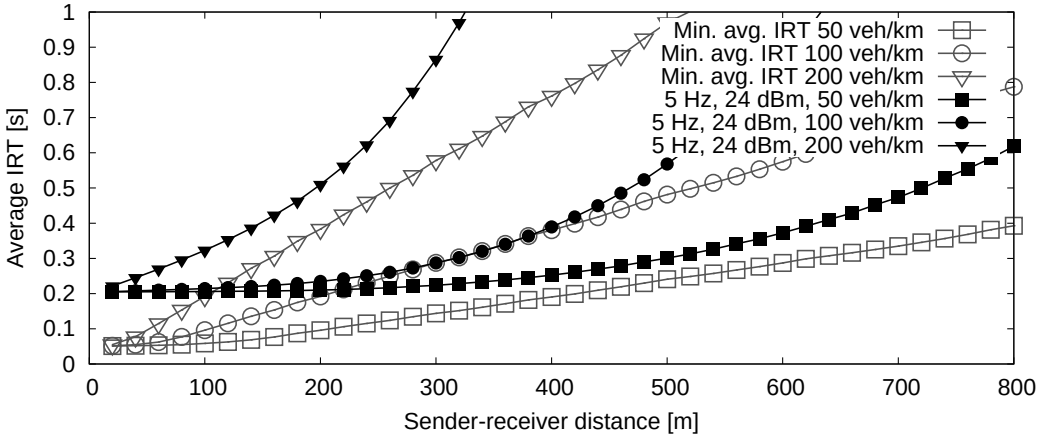
of transmission parameters is necessary to keep channel load at its optimal level when vehicle density changes.



(a) 10 Hz, 24 dBm

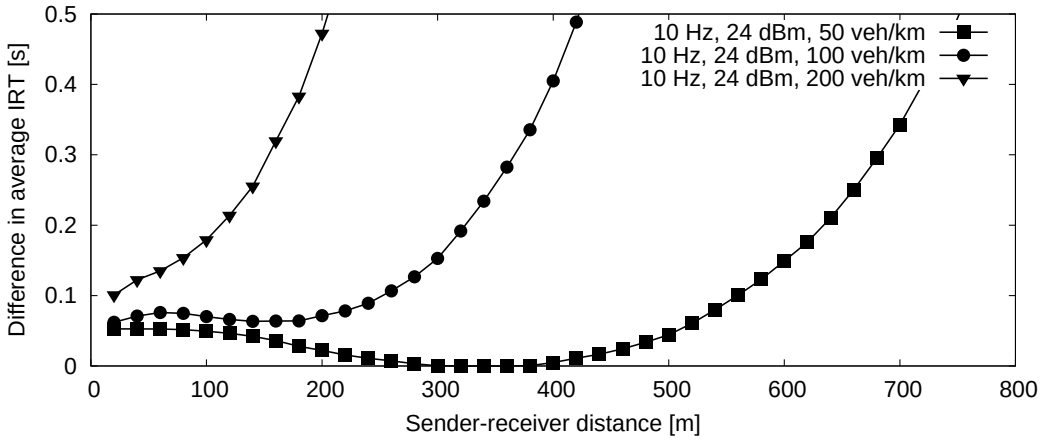


(b) 10 Hz, 18 dBm

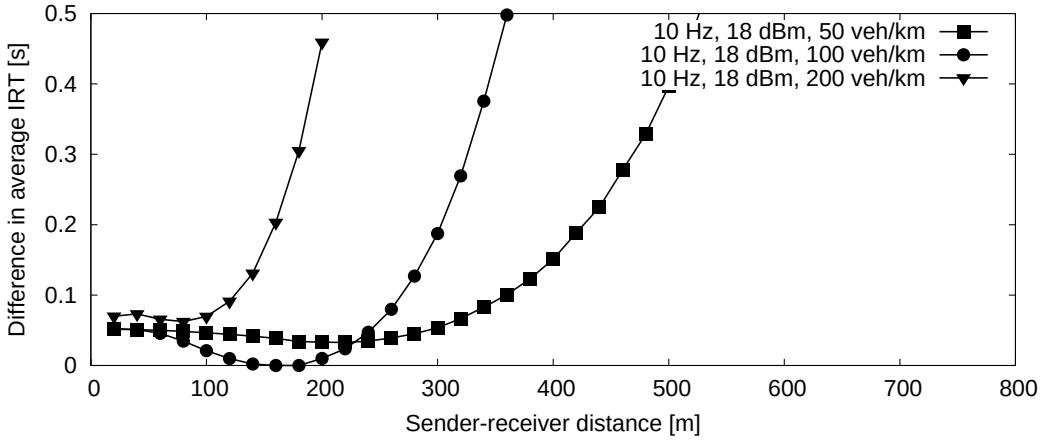


(c) 5 Hz, 24 dBm

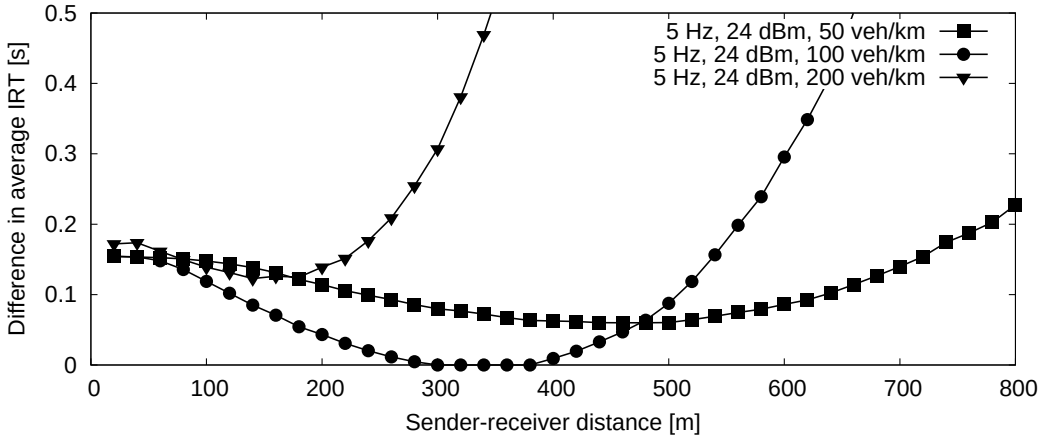
Figure 4.12: Average IRT for different fixed transmission parameter settings compared to minimum values



(a) 10 Hz, 24 dBm



(b) 10 Hz, 18 dBm



(c) 5 Hz, 24 dBm

Figure 4.13: Difference to minimum average IRT for different fixed transmission parameter settings

5

Design of Congestion Control for VSC

As discussed in Chapter 3 on page 45, so far no consensus has been established in the literature as to how and why transmission rate and transmit power should be adapted to control channel load in the context of Vehicle Safety Communications (VSC). While some authors advocate Transmit Power Control (TPC), e.g., [20][190], others suggest Transmission Rate Control (TRC), e.g., [165][202], while still others developed a combination of both, e.g., [23][207]. In addition, some authors have pointed out the necessity to integrate the consideration of safety applications' requirements into congestion control, e.g., [23][93][165].

In this chapter, we systematically derive a congestion control methodology for VSC which addresses both objectives of keeping the channel in working condition while operating within the limits defined by safety applications. We thereby exploit the commonalities of the optimized transmission parameter combinations identified in the previous chapter.

In particular, we address the following research questions:

1. How can transmit power and transmission rate be adapted to converge to the identified optimal parameter combinations?
2. What is the relationship between transmit power, transmission rate and channel load?
3. Which channel load should congestion control converge to?
4. How do we define fairness?

5. How can the requirements of safety applications be integrated into congestion control?

5.1 Control Strategies

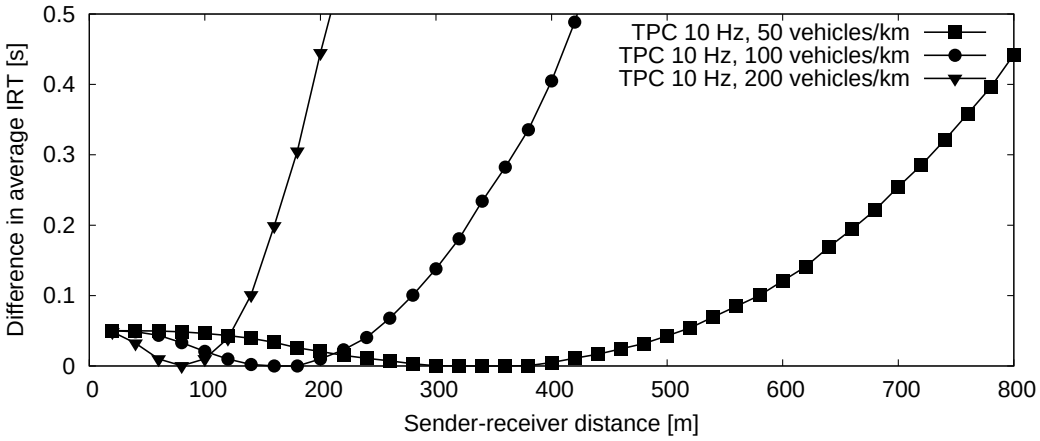
In this section, we evaluate the potential of different control strategies to converge toward the minimum average Packet Inter-Reception Time (IRT) identified in Section 4.2.2 on page 89 for each vehicle density and target distance. We first show that neither TPC nor TRC alone supports optimizing reception performance individually for each vehicle based on its target distance. We then introduce a joint control strategy which is able to overcome this limitation. Note that the results presented in this section are based on global knowledge from the simulator and thus correspond to the optimal outcome for each strategy. The outcome for a decentralized protocol may vary. The results presented in this section have been previously published in [26].

5.1.1 Transmit Power Control (TPC)

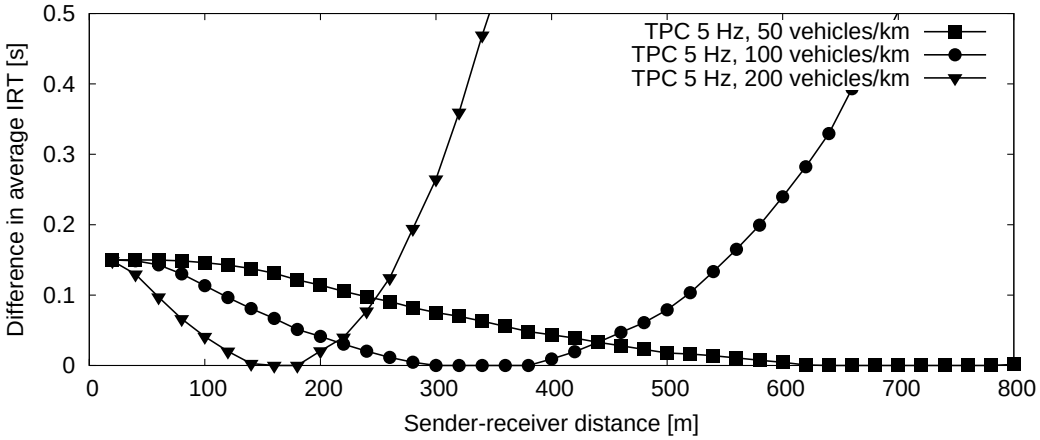
In TPC, while transmit power is adapted with respect to channel load, transmission rate is typically set to a fixed value for all vehicles, e.g., 10 Hz. In the previous chapter, we saw that each transmit power optimizes reception performance for a certain distance, cf. Figure 4.3a on page 92. That is, by adapting transmit power based on channel load, we implicitly adapt the distance at which reception performance is optimized based on channel load as well. Since the transmission rate is fixed, the optimization distance thus depends on vehicle density.

Figure 5.1 on the facing page illustrates the difference in average IRT between the global minimum for each distance and the optimal outcome of a TPC strategy with a fixed transmission rate. The two subfigures correspond to different fixed transmission rates, i.e., 10 Hz in Figure 5.1a and 5 Hz in Figure 5.1b. The outcome for the TPC strategy was determined by selecting the minimum average IRT per distance among all parameter combinations using the respective fixed transmission rate.

As expected, we observe that the outcome at a particular target distance depends on vehicle density. In Figure 5.1a for example, a vehicle with a target distance of 300 m would experience a near-optimal reception performance at its target distance and a vehicle density of 50 vehicles/km, while the difference to the optimal outcome would be 127 ms in the case of 100 vehicles/km and 841 ms in the case of 200 vehicles/km. A comparison between both subfigures yields that a lower fixed transmission rate allows to extend transmit power, which improves reception performance at farther distances. However, reception performance at close distances is worse compared to a higer transmission rate.



(a) Transmission rate 10 Hz



(b) Transmission rate 5 Hz

Figure 5.1: Difference in average IRT between the minimum value for a fixed transmission rate and the global minimum value for each distance

In a nutshell, the TPC approach optimizes reception performance for a distance which is determined by vehicle density. Thus, we conclude that it does not support optimizing reception performance individually for each vehicle at its respective target distance.

5.1.2 Transmission Rate Control (TRC)

In TRC, transmission rate is typically adapted with respect to channel load, while transmit power is fixed to a certain value. That is, TRC determines *a priori* at which distance reception performance will be optimized, i.e., it assumes that all vehicles share the same target distance.

Figure 5.2 on page 107 illustrates the difference between the globally minimum average IRT per distance and the result obtained for a TRC approach, i.e., the minimum

value among all combinations using a certain transmit power. The two subfigures correspond to different fixed transmit power values, i.e., 16 dBm in Figure 5.2a and 20 dBm in Figure 5.2b.

We can see that as expected, the TRC approach optimizes reception performance for a certain distance, independent of vehicle density. However, reception performance degrades significantly beyond this point. In Figure 5.2a for example, a vehicle with a target distance of 300 m would experience a difference of 16 ms between the outcome of TRC and the optimal value at a vehicle density of 50 vehicles/km. The difference increases to 204 ms for 100 vehicles/km and to 406 ms for 200 vehicles/km. The comparison with Figure 5.2b shows that fixing transmit power at a higher value results in a greater optimization distance. However, in comparison the performance at close distances deteriorates.

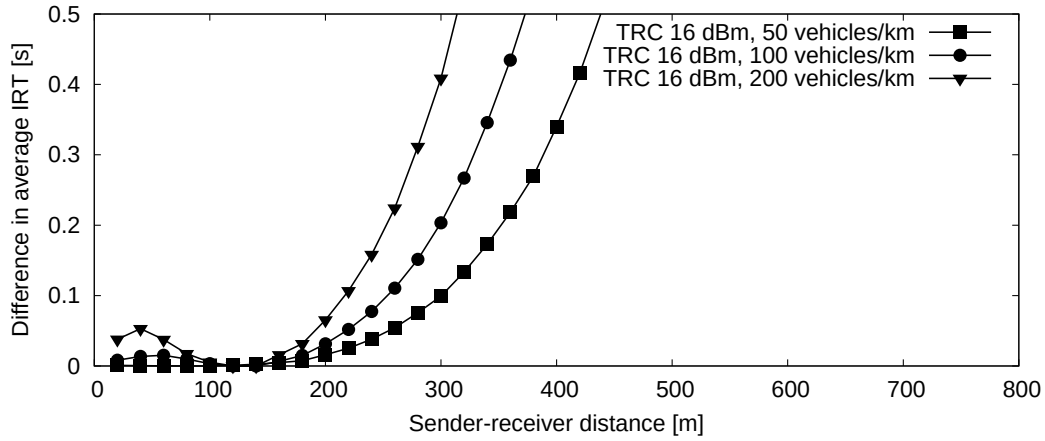
To summarize, while the TRC strategy results in a more predictable reception performance with respect to distance than the TPC strategy, it requires the selection of a fixed transmit power which determines the maximum target distance. Beyond this point, reception performance deteriorates rapidly.

5.1.3 Joint Control of Transmit Power and Transmission Rate

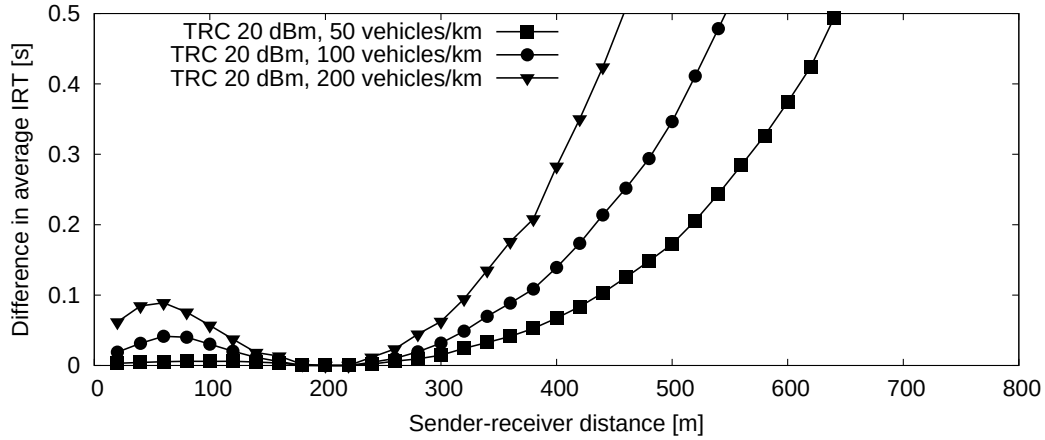
Between TPC and TRC, the latter seems like the better choice if we would like to optimize reception performance for a certain target distance. However, as we have seen there is room for improvement if target distances vary among vehicles. In this case, we suggest the following joint control strategy: *Select transmit power as optimal for the current target distance, i.e., based on the driving context, then adjust transmission rate with respect to channel load.* In a nutshell, the suggested strategy is an extension of TRC which on the one hand reduces unnecessary interference caused by excessive transmit power and on the other hand helps each vehicle to reach its respective target distance. When its target distance changes, a vehicle adapts its transmit power accordingly.

We further assume that the transmission rate is adapted between a minimum and a maximum value. A global minimum rate of, e.g., 1 Hz may be applied to notify newly arriving neighbor vehicles, while a global maximum rate of, e.g., 10 Hz may reflect the idea that the additional safety benefit of receiving beacon information is likely to decrease with update frequency. Optionally, minimum and maximum transmission rate could be determined individually for each vehicle depending on its driving context. We further discuss this idea in Section 5.4 on page 123.

Note that the existence of a minimum transmission rate requires our joint control to have a fallback mechanism in the case that all vehicles have already reached their minimum transmission rate and channel load is still too high. In this case, we suggest to reduce transmit power to ensure communication with close neighbors. When the channel congestion has been resolved, transmit power is increased first to its desired value before transmission rate is increased.



(a) Transmit power 16 dBm



(b) Transmit power 20 dBm

Figure 5.2: Difference in average IRT between the minimum value for a fixed transmit power and the global minimum value for each distance

The basic functionality of the suggested joint control strategy is summarized in Algorithm 1 on the next page, where r, r_{min} and r_{max} denote the current, minimum and maximum transmission rate, respectively, U and U_t denote the current and targeted channel load, respectively, P_{Tx} is the current transmit power and f is a function or look-up table to determine P_{Tx} for a given distance and channel load. Note that the provided pseudo-code is intended as a rough outline rather than a concrete protocol. Especially the way of incrementing and decrementing transmission rate could be realized in different ways.

Figure 5.3 on page 109 illustrates the difference in average IRT between the outcome of the suggested joint control strategy and the globally minimum average IRT per distance. The two subfigures correspond to different minimum transmission rates, i.e., 1 Hz in Figure 5.3a and 3 Hz in Figure 5.3b. To create a look-up table for the transmit power to be used at a particular distance, we used the transmit power value optimizing

Algorithm 1: Joint power/rate control strategy (basic functionality)

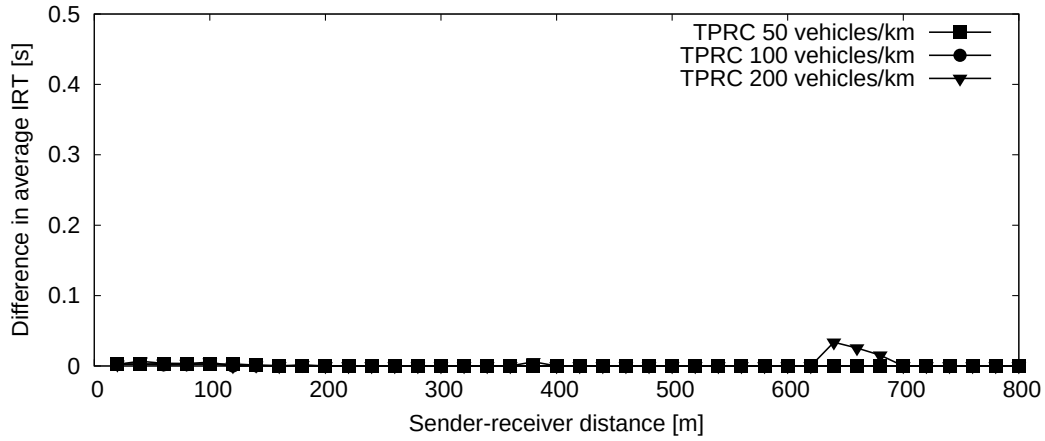
```
Data: r_min, r_max, d_t, U_t, U
r = r_min; P_Tx = f(d_t, U_t);
while congestion control active do
    if U < U_t then
        if P_Tx < f(d_t, U_t) then
            Increase P_Tx;
        else
            if r < r_max then
                Increase r;
            end
        end
    else
        if r > r_min then
            Decrease r;
        else
            Decrease P_Tx;
        end
    end
end
```

average IRT for the majority of vehicle densities in Figure 4.3a on page 92. For the comparison, we then selected the minimum resulting IRT among all combinations using this transmit power and any transmission rate.

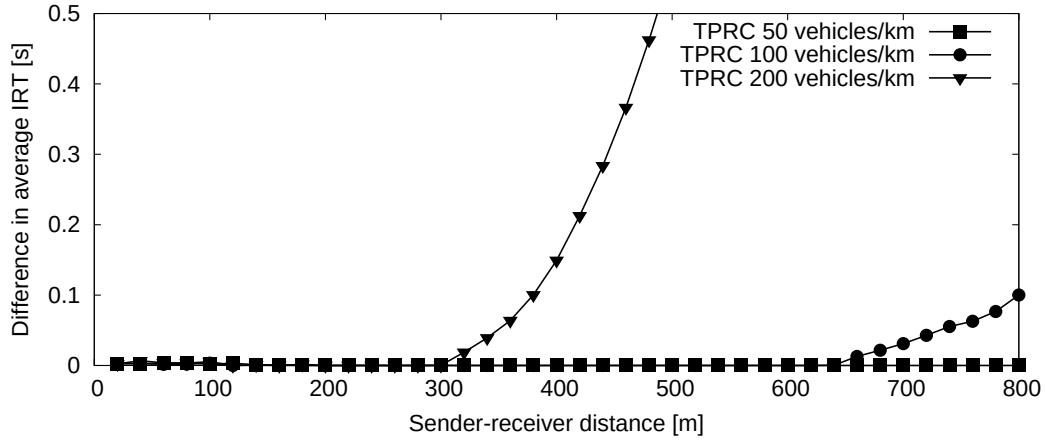
In Figure 5.3a, we can see that the joint control strategy effectively approximates the optimal performance for all evaluated vehicle densities and (almost) all target distances. There are some outliers between 600 m and 700 m which result from the granularity of our parameter space. More specifically, the transmit power optimizing average IRT for a vehicle density of 200 vehicles/km deviates from the optimal transmit power for the majority of vehicle densities, cf. Figure 4.3a on page 92. As a result, the strategy does not select the optimal value.

In Figure 5.3b on the next page, we can see that for a higher minimum transmission rate, the joint control strategy optimizes reception performance up to the distance which the channel permits.

To conclude, what we see as the main advantage of the joint control strategy introduced in this section is that it efficiently approximates the identified optimal transmission parameter configurations for a given channel load target and target distance while reducing complexity due to its independence of vehicle density. Note that if the safety application(s) run by a vehicle require the optimization of reception performance at more than one target distance, the approach could be extended, e.g., by running multiple instances with different target distances simultaneously. A further advan-



(a) Minimum transmission rate 1 Hz



(b) Minimum transmission rate 3 Hz

Figure 5.3: Difference in average IRT between the suggested joint control strategy and the global minimum value for each distance

tage of the suggested joint control strategy is that by applying TPC as a secondary mechanism, it is less dependent on a fine granularity of transmit power than pure TPC approaches. Since a precise control of the realized output power is challenging to implement, practitioners may consider using rather coarse power settings such as “low”, “medium” and “high”.

5.1.4 Discussion

In the last part of this section, we address two potential points of criticism of the presented optimization study. First, we discuss which conclusions we would draw based on a different optimization criterion, i.e., the maximization of the Cumulative Density Function (CDF) of IRT for a certain time window. Second, we address the impact of different channel models, especially regarding non-line-of-sight conditions.

Optimization of T-Window Reliability

In Section 4.1.2 on page 86, we discussed the option of using the maximization of the CDF of IRT for a certain time window t_{Rx} at the target distance d_t as the optimization criterion for our study. The CDF of IRT is commonly referred to as T-window reliability (using $t_{Rx} = T$) and is frequently applied in state-of-the-art awareness metrics.

In the context of our optimization study, we argue that an optimization criterion based on a specific value of t_{Rx} in Equation (4.4) on page 87 would lead to a black and white view of safety, since it assumes a binary transition between "safe" (a message is received within t_{Rx}) and "unsafe" (no message is received within t_{Rx}). Taken literally, this binary transition means that for a given t_{Rx}, any transmission interval $t_{Tx} = t_{Rx} + \varepsilon$, $\varepsilon \to 0$, does not provide any safety benefit at all, while any $t_{Tx} = t_{Rx} - \varepsilon$ would not yield any additional safety benefit compared to t_{Tx}. As a result, only values of t_{Tx} which are multiples of t_{Rx} would increase safety benefit. We argue that while it makes sense to use t_{Rx} as a minimum performance requirement, a further reduction of t_{Rx} could have an additional safety benefit, e.g., by reducing tracking error. In addition, the calculation of the required value of t_{Rx} in a particular traffic situation is typically subject to a certain degree of uncertainty due to factors like the driver's perception-reaction time, road friction, etc. We thus consider a graceful degradation of the safety benefit from a minimum value of t_{Rx} to a maximum value of t_{Rx} to be more realistic.

From an optimization point of view, the CDF value in Equation (4.4) can be increased by either applying a higher multiple of t_{Rx} to the transmission interval t_{Tx} or by increasing the reception probability p_{Rx}. When applying this criterion to our study, the resulting optimal transmission rate values are multiples of t_{Rx} as expected. However, neither the resulting optimal transmission power values nor the optimal transmission rates show any apparent commonalities across vehicle densities. That is, a control strategy to converge toward these optimal points would require knowledge of the current vehicle density as well as of the trade-off in terms of additional channel load between increasing transmission rate by x Hz and increasing transmit power by y dB. However, such a decision would require both an accurate model to predict channel load as well as the potential to tune transmit power in a fine granular way.

Channel Model and Measurements

The presented results of our optimization study are based on simulations using the power-law model described in Appendix A.1 on page 223. We also validated that the independence of the optimal transmit power of vehicle density also holds for other types of fading, e.g., Nakagami with different m-parameters. Still, simulations have their limitations when it comes to the potential influencing factors encountered in a real-world environment.

One of such aspects is the transition between Line of Sight (LOS) and Non Line of sight (NLOS) conditions encountered not only due to static obstacles like buildings

but also due to mobile obstacles such as trucks. In other words, the probability of message reception p_{Rx} might not be constant, as assumed in this study, which would have an impact on the distribution of IRT. In the related work, different authors came to different conclusions as to whether the geometric distribution applies to the IRT observed in a real-world environment. In [97], Martelli et al. conducted a measurement campaign in different environments and found the geometric distribution to be a poor match for the obtained IRT measurements. However, it is not clear if the authors evaluated whether their results matched a geometric distribution for either LOS or NLOS. In [212], Sepulcre et al. found the measured IRT in both urban and highway scenarios with LOS obstructions to be sufficiently well approximated by a geometric distribution to apply the T-window reliability concept. The examples illustrate that further measurement campaigns in different environments are required to better understand the reception performance in terms of IRT.

As for the joint control strategy introduced in this section, an unpredictable change in LOS/NLOS conditions could be countered by choosing transmit power based on NLOS when in doubt. The reception performance under LOS conditions would then be better. The change in LOS conditions can also be interpreted as a further indicator why it could be beneficial to transmit at a higher rate than absolutely necessary, since it increases the chance of transmitting a message during LOS conditions.

5.2 System Model and Convergence Target

In the previous section, we introduced a joint control strategy which selects transmit power with respect to the desired target distance and adapts transmission rate with respect to channel load. In the terminology of control theory, we thus suggested an *open loop* control, i.e., without feedback, for transmit power and a *closed loop* control, i.e., with feedback, for transmission rate.

To design a closed loop controller, we require suitable feedback metric as well as a model of the system to be controlled which describes the relationship between input and output. In [71], Kenney et al. introduce a rate control model which can be used to describe the relationship between a vehicle's transmission rate and the observed channel load in terms of Channel Busy Ratio (CBR). The model forms the basis of the LIMERIC algorithm discussed in Section 7.4.3 on page 211 and was validated against simulations as well as measurements in a controlled environment without fading and hidden terminal effects. In the first part of this section, we introduce the model by Kenney et al. and compare it to simulation results obtained in a scenario where all nodes share the same location. In the second part, we use the concept of communication density [112] to show based on our simulation results of the optimization study presented in Section 4.2.2 on page 89 that the linear relationship between transmission rate and CBR can also be observed in a distributed hidden-

node dominated environment. Parts of the contributions in this section have been previously published in [26].

5.2.1 Rate Control Model According to Kenney et al.

In [71], Kenney et al. model the channel load experienced by the network as the aggregate transmission rate of all participating network nodes. More specifically, they calculate the fraction of network capacity r_C used by all K nodes sharing the channel at time t as

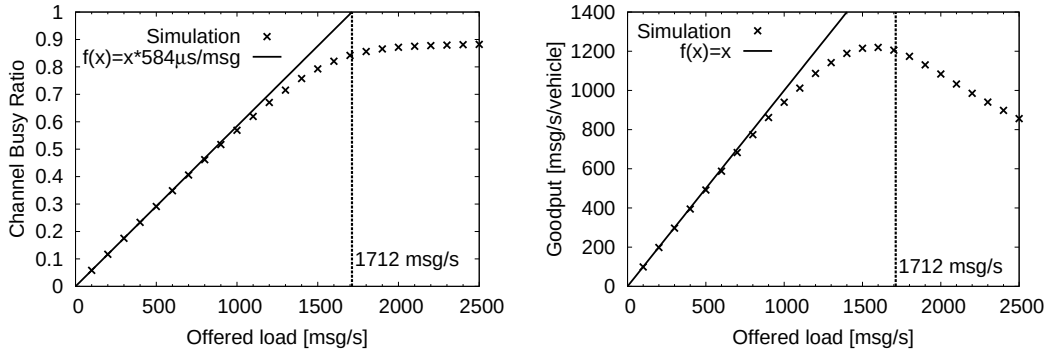
$$r_C(t) = \sum_{j=1}^{K} r_j(t) \tag{5.1}$$

where r_j is the transmission rate of node j expressed as a fraction of the total network capacity. In our setup for example, the transmission of a single message takes 584 μs, resulting in a maximum of 1712 messages/s to be transmitted on a 6 Mbit/s channel. A node transmitting at 10 Hz would thus occupy a fraction of $r_j = 0.00584$ of the total capacity. Kenney et al. further show that $r_C(t)$ can be effectively approximated by means of CBR measurements.

Figure 5.4a on the next page illustrates the relationship between offered load, i.e., the total number of messages generated by all nodes in the network, and the resulting CBR in a scenario where 100 wireless nodes are placed in the same location, which is similar to the setup used by Kenney et al. [71][213]. The scenario allows to evaluate the performance of IEEE 802.11 under idealized conditions, i.e., without hidden nodes and radio-wave propagation effects. The results presented in the figure are based on the simulator configuration described in Appendix A on page 223. Note that in the given scenario, the offered load can be trivially calculated as $K \cdot r$, where r is the uniform transmission rate of all nodes.

As expected according to Equation (5.1), we observe a linear relationship between offered load and measured CBR with a slope of 584 μs/message, i.e., the duration of one message in our setup. However, beyond an offered load of approximately 1200 messages/s, i.e., a CBR near 0.7, the relationship starts becoming increasingly nonlinear as the CBR approaches a saturation point near 0.9. The example illustrates that under high channel load, the observed CBR is most likely not a reliable indicator for the amount of offered load.

Figure 5.4b illustrates the relationship between offered load and goodput, i.e., the number of successfully received messages per node and second, in the same scenario as before. We observe a 1:1 ratio between offered load and goodput up to approximately 800 messages/s, the point at which the relationship between offered load and CBR first begins to deviate visibly from linearity. At this point, packet collisions start to occur due to the limited size of the contention window, which leads to simultaneous transmissions, cf. Section 2.1.2 on page 13. Note that goodput is maximized near an offered load of 1500 messages/s, i.e., before the theoretical maximum of 1712 messages/s



(a) CBR (b) Goodput

Figure 5.4: Observed CBR and goodput in a simulation where 100 nodes share the same location, compared to the theoretically expected outcome according to the model of Kenney et al.

has been reached. At the point of maximum goodput, the Packet Delivery Ratio (PDR) is approximately 0.82 and the observed CBR is approximately 0.79.

5.2.2 Relationship of Communication Density and CBR

While the model by Kenney et al. has been validated in a scenario where all nodes can hear each others' transmissions and the Carrier Sense Multiple Access (CSMA) mechanism works perfectly, the question arises whether the linearity between offered load and CBR still applies in a distributed, hidden-node and fading dominated environment as frequently encountered in VSC. To tackle this question, we use the simulation results obtained in the optimization study of Section 4.2.2 on page 89 and evaluate them based on the concept of Communication Density (CD) [112], cf. Section 2.3.3 on page 40.

CD is a theoretical metric calculated based on global knowledge. It describes the relationship between transmit power, transmission rate and vehicle density and is proportional to the number of carrier sensible events at a location [112]. In this section, we calculate the CD according to Equation (2.11) on page 41. In the following, we interpret CD as a distributed form of the aggregate transmission rate r_C in Equation (5.1). For example, if a total of 856 messages can be carrier sensed at a particular location, we expect a CBR of approximately 50%.

Channel Performance with Respect to Transmit Power and Transmission Rate

Figure 5.5 on page 115 illustrates three different performance metrics in terms of channel state (rows of subfigures) in two different vehicle densities (columns of subfigures) based on the simulation results of the optimization study presented in

Section 4.2.2 on page 89. In addition to CBR and goodput, the figure further shows the observed average Channel Access Time (CAT)[1], i.e., the time it takes a node from generating a message until accessing the wireless medium, for each combination of transmit power and transmission rate. In Figures 5.5a and 5.5b we observe a flat region of a high CBR near 0.9 which resembles the shape of the CBR curve in Figure 5.4a on the preceding page. However, comparing the two figures, we can see that the region moves with respect to vehicle density. A similar observation can be made in Figures 5.5c and 5.5d, where we can see that the set of transmission parameters maximizing goodput differs between vehicle densities. Finally, Figures 5.5e and 5.5f illustrate that within the region resulting in a flat CBR, the CAT increases significantly, i.e., up to an average of 70 ms in the case of 50 vehicles/km and up to an average of 106 ms in the case of 100 vehicles/km. At such conditions, it is a likely event that a vehicle generates a message while its previous message is still in the Medium Access Control (MAC) queue. This condition has been denoted as *local congestion* in the related work [182]. For safety, this situation is particularly critical, since it not only prevents a vehicle from communicating its current status but may also lead to the transmission of outdated information unless the packet in the MAC queue is replaced by the newly arriving one.

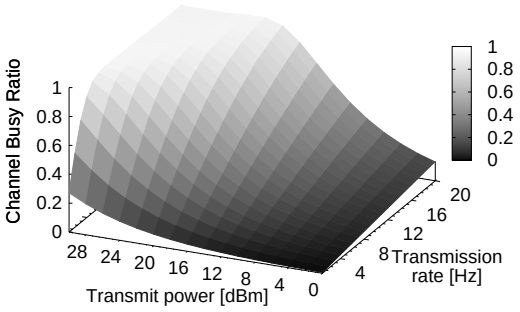
Channel Performance with Respect to Communication Density

Based on the representation of the data in Figure 5.5, it is difficult to draw any conclusions regarding a suitable operation region for congestion control. In particular, we require an a priori perception of the current vehicle density in the scenario to find the set of transmission parameters maximizing goodput while keeping CAT under control. Figure 5.6 on page 116 illustrates the same data, complemented by the results for 200 vehicles/km, using the representation of CD. That is, we calculated the CD according to Equation (2.11) on page 41 for each parameter combination and plotted our simulation results accordingly.

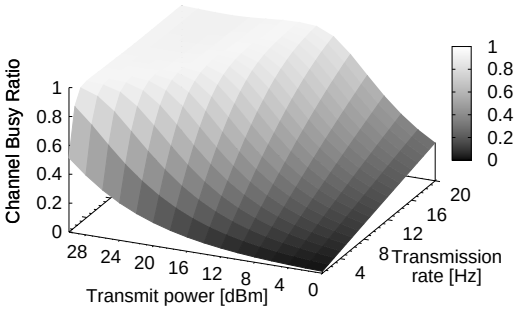
In Figure 5.6a, we observe an unambiguous relationship between CD and CBR up to a certain saturation point. The relationship is independent of vehicle density and, up to a CBR near 0.7, can be approximated using a linear function with a slope of 584 μs/message. That is, our results indicate that a linear rate control based on Equation (5.1) can be applied to a distributed environment if the aggregated transmission rate is calculated according as the CD. Note that for d_{cs} = 1 m and $\rho = N$ vehicles/2 m, Equation (2.11) becomes $r \cdot N$, which is how we calculated the offered load in Section 5.2.1 on page 112. We conclude that in practice, CBR measurements can be used to approximate the CD in a scenario.

Figure 5.6b illustrates the relationship between CD and goodput based on our simulation results. The figure shows the combined goodput in units of 1000 messages/s

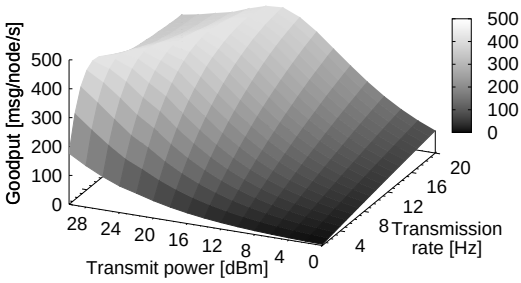
[1]The results presented in this thesis are based on the backoff countdown method of Distributed Coordination Function (DCF). With Enhanced Distributed Channel Access (EDCA), lower CAT values can be observed. We discuss this issue in Section 7.5.2 on page 216



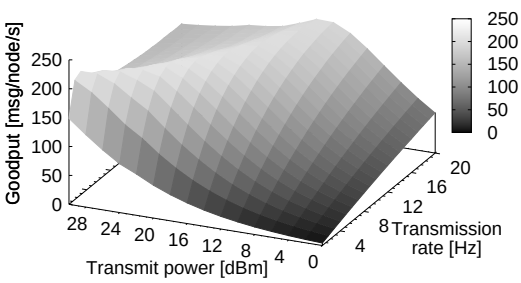
(a) 50 vehicles/km: Channel Busy Ratio (CBR)



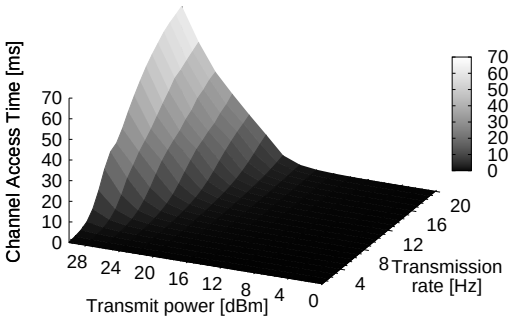
(b) 100 vehicles/km: Channel Busy Ratio (CBR)



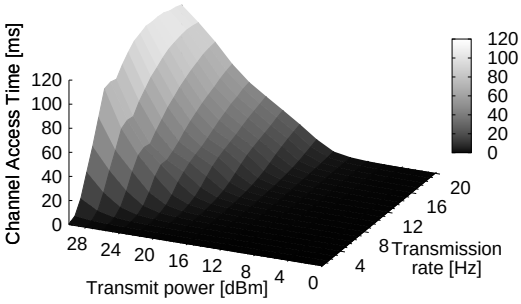
(c) 50 vehicles/km: Goodput



(d) 100 vehicles/km: Goodput



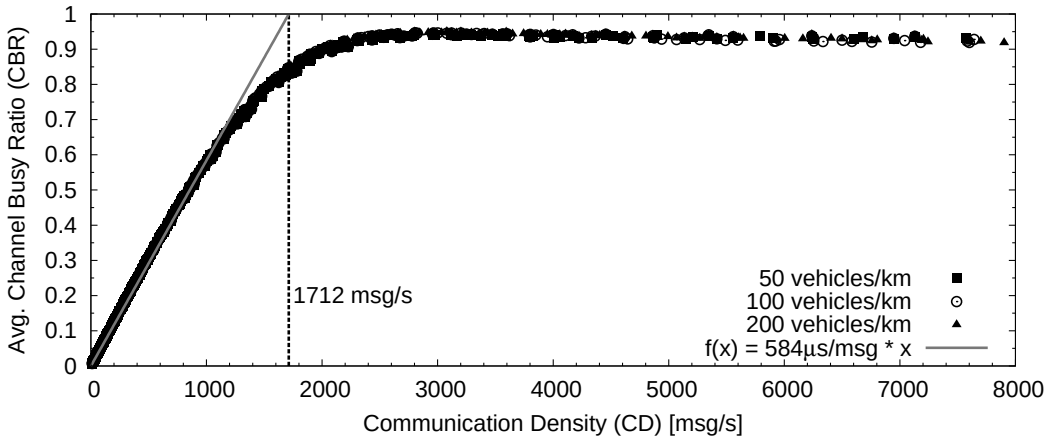
(e) 50 vehicles/km: Channel Access Time (CAT)



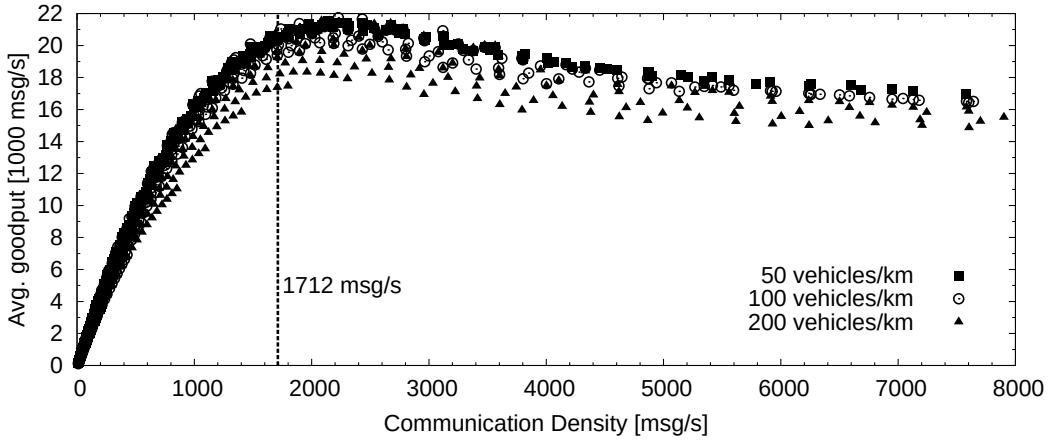
(f) 100 vehicles/km: Channel Access Time (CAT)

Figure 5.5: Channel performance metrics observed in the optimization study with uniform node configuration

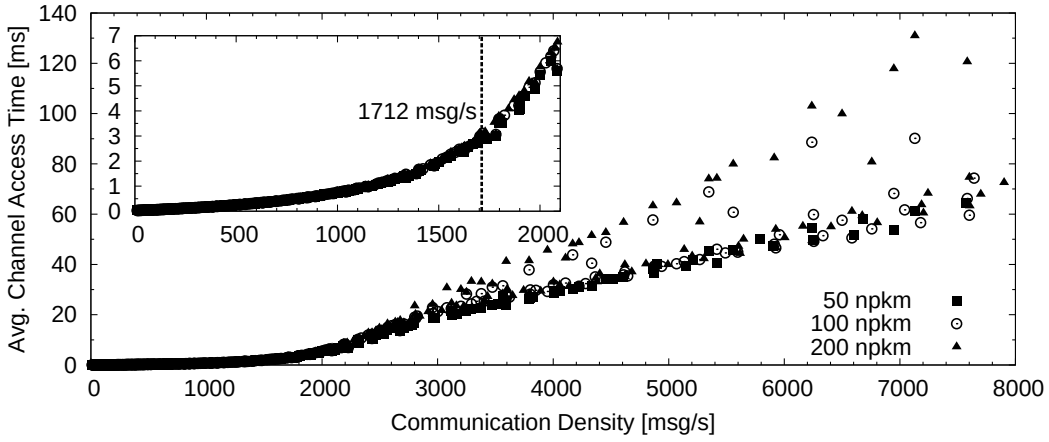
of all nodes in the scenario. We observe that goodput is maximized near a CD of 2100 messages/s, i.e., beyond the theoretical maximum of 1712 messages/s that can be accommodated by the channel. We explain this discrepancy based on the carrier sense threshold, which was set to -95 dBm for the simulations in this thesis. In contrast, at least -91 dBm are required to receive a message, which means that not all carrier



(a) Channel Busy Ratio (CBR) vs. Communication Density (CD)



(b) Goodput vs. Communication Density (CD)



(c) Channel Access Time (CAT) vs. Communication Density (CD)

Figure 5.6: Relationship between Communication Density (CD) and different performance metrics

sensible messages can be received. We further observe an increased variance of the relationship between goodput and CD with increasing vehicle density. The result indicates that scenarios with the same CD experience similar but not necessarily identical channel conditions.

Figure 5.6c depicts the relationship between CD and average CAT according to our simulation results. We observe an unambiguous relationship between both metrics up to a CD of approximately 2000 messages/s. Note that the extreme values of a CAT beyond 100 ms observed in Figure 5.5f correspond to a CD of more than 6000 messages/s, i.e., significantly beyond the channel capacity. That is, by controlling CD, we can avoid channel conditions in which local packet drops are likely to happen.

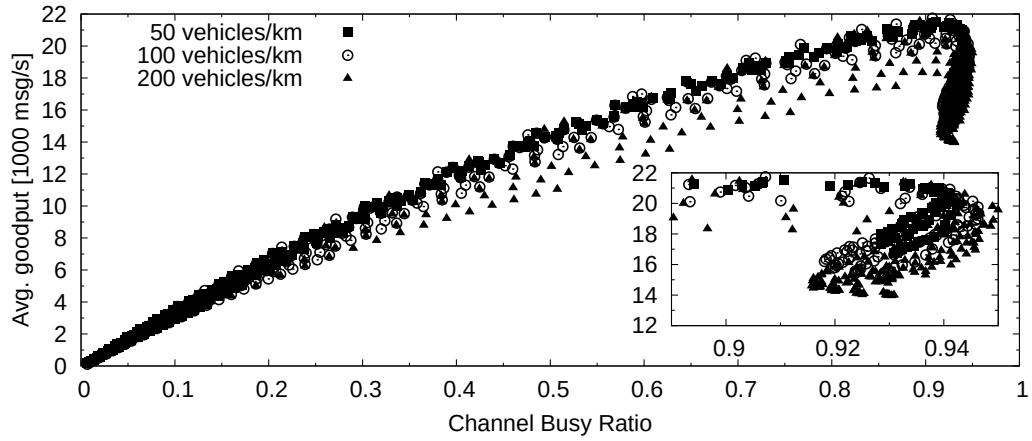
Channel Performance with Respect to CBR

As discussed above, the observed CBR can be used to approximate the current CD in a scenario. Figure 5.7 on the following page illustrates this approach by plotting goodput and CAT against CBR directly[2]. In Figure 5.7a, we observe a monotonically increasing relationship up to a "cliff" near a CBR of 0.91. In Figure 5.7b, we can see that this CBR value is also critical in terms of CAT, since beyond this point, the relationship becomes ambiguous. In the reverse conclusion, the figure shows that CBR measurements can be used as a feedback in order to keep the channel in an operational state with high goodput and limited CAT.

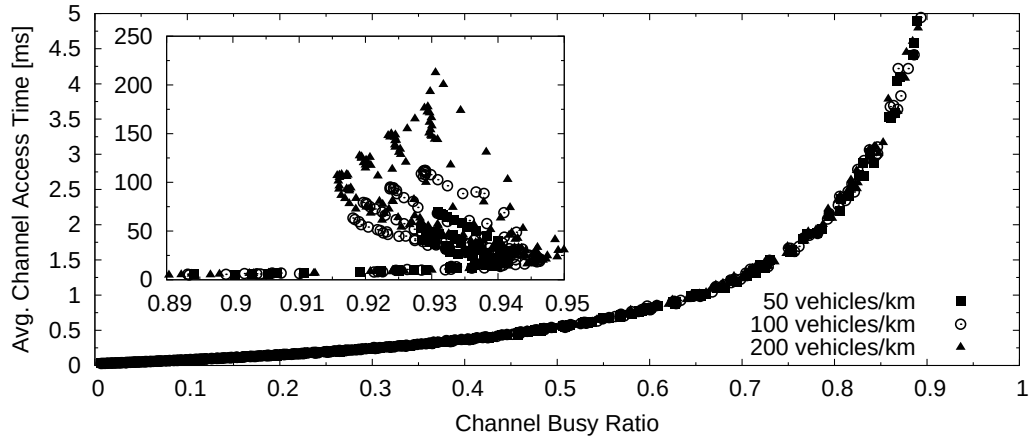
What remains in terms of the design of congestion control is the identification of a channel load target. In Section 3.1 on page 46, we discussed the metric of network power according to Jain et al. [115], which denotes the quotient of throughput and response time, cf. Equation (3.1) on page 48, and can be used to identify the "knee" of the throughput curve at which a congestion avoidance scheme should be operating. While Jain et al. do not define response time explicitly, we interpret the term as the total end-to-end delay between sender and receiver. Using this interpretation, we calculated the network power metric in Figure 5.7c by dividing the goodput as shown in Figure 5.7a by the sum of the CAT as shown in Figure 5.7b and the transmission time of one message[3], i.e., 584 μs. As we can see, the power metric according to this interpretation is maximized at a CBR between 0.4 and 0.5., a point at which the network reaches approximately 2/3 of the maximum goodput.

[2] This result is similar to the findings of Fallah et al. in [193], where the authors study the relationship between CBR and IDR, i.e., the goodput within a certain sender-receiver distance. The authors also study a large parameter space of transmit power and transmission rate and find that the IDR values resulting from different parameter choices fall on a single line. That is, our results presented here basically confirm the findings of Fallah et al. which were obtained using a different simulator and configuration.

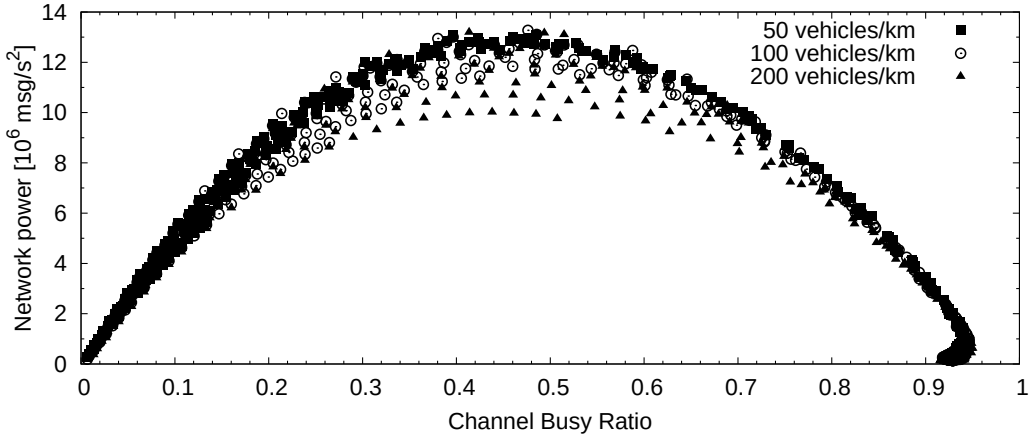
[3] This calculation of the end-to-end delay does not take into account propagation time, which we assume to be negligible for the considered distances. The power metric as calculated here shows a low sensitivity to small changes in the additive delay factor.



(a) Goodput vs. Channel Busy Ratio (CBR)



(b) Channel Access Time (CAT) vs. Channel Busy Ratio (CBR)



(c) Network power vs. Channel Busy Ratio (CBR)

Figure 5.7: Relationship between Channel Busy Ratio (CBR) and different performance metrics

Assuming that our interpretation of the power metric is correct, the result would indicate a CBR target of approximately 0.5 for a congestion avoidance scheme. On the other hand, we saw in Figure 4.7 on page 96 that the parameter combinations optimizing IRT resulted in a CBR near 0.9, i.e., at the point where the goodput curve has its maximum in Figure 5.7a. From a safety perspective, the optimization of IRT includes the optimization of channel access delay as well as of the update delay resulting from packet loss [90]. It is thus not clear at this point whether the power metric as applied above is the right means of defining the "knee" of the goodput curve in the context of VSC. We leave a further exploration of this topic for future work.

In this thesis, we use a CBR target of 0.7, which corresponds to the point beyond which the relationship between offered load and CBR starts to become nonlinear in Figures 5.4a on page 113 and 5.6a on page 116 and can thus be interpreted as a "knee" in the corresponding curves. In addition, the value provides a certain distance from the "cliff" at 0.91 CBR while resulting in approximately 80% of the maximum goodput in Figure 5.7a on the facing page. A CBR target between 0.6 and 0.7 is in line with other works, e.g., [71][178][193]. Note, however, that the approach presented in this thesis is not limited to a specific CBR value.

5.3 Aspects of Fairness

Since congestion control and resource sharing are two sides of the same coin [126], the topic is closely linked with the question of fairness. In computer networks, fairness is typically expressed with respect to throughput, i.e., the share of bandwidth occupied by each user. Two commonly encountered fairness principles are max-min fairness, which maximizes the minimum share in the network, and proportional fairness, which aims at a balance between fairness and the utilization of resources, cf. Section 3.1.2 on page 48.

In the context of congestion control for VSC, fairness has been defined differently by different authors. Some examples are shown in Table 5.1 on the following page. In [25], we proposed three fairness principles to consolidate the views expressed in the related work with our own considerations, i.e., local fairness, global fairness and participation fairness. We describe each of these aspects in the remainder of this section. The terminology has been taken up in different works, e.g., [71][196][204].

5.3.1 Local Fairness

In a wireless channel, vehicles located physically close to each other typically experience similar conditions in terms of channel load and interference. The local fairness principle states that consequently, neighboring vehicles should be regulated by congestion control to the same *degree* [25]. Note that this principle does not necessarily

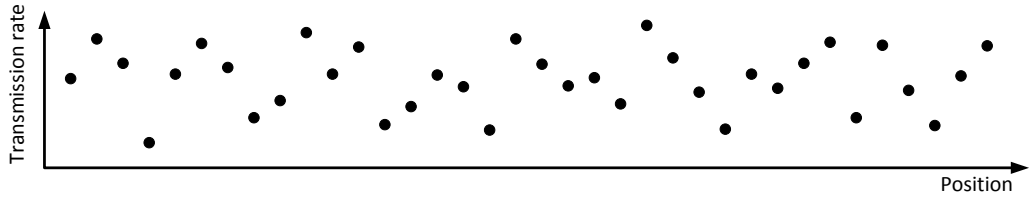
Approach	Fairness definition
D-FPAV [20]	"Maximize the minimum transmit power value over all transmission power levels assigned to nodes that form the vehicular network"
SPAV [189]	Same as D-FPAV [20]
Khorakhun et al. [191]	"[E]qually distributed update rates or transmission powers over all nodes in transmission range of each other"
SOURC [202]	"[U]niform distributed broadcast interval used by all vehicles staying inside the carrier sense range"
Kloiber et al. [196]	"[S]tatistical fairness as all vehicles use the same probability distribution and thus the same average TX power"
GRC [192]	"Space fairness"
Baldessari et al. [207]	"[A]llocate an equal amount of resources to each node by means of direct partitioning"
CTA [182]	"[S]ame chances to access the channel"
LIMERIC [71]	Local fairness (as defined in [25])
EMBARC [204]	Local and global fairness (as defined in [25])

Table 5.1: Examples of fairness concepts in the related work

require nearby vehicles to share the same transmission parameters. For example, the adaptation of the relative transmission rate as introduced in Section 5.4 on page 123 may result in the same *relative* transmission rate but different *absolute* transmission rates for neighboring vehicles[4].

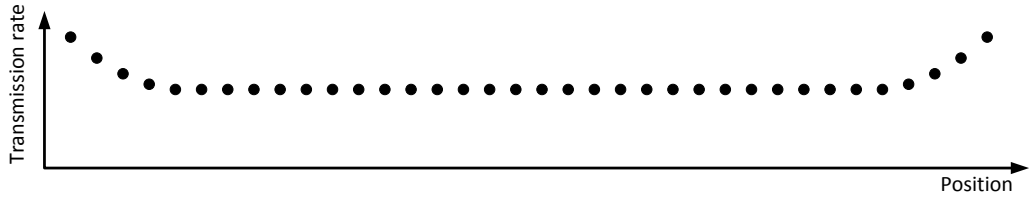
Figure 5.8 on the next page provides a schematic illustration of the local (and global, see below) fairness principle using the example of transmission rate control with a uniform minimum and maximum rate. In the figure, each dot represents the transmission rate of a vehicle with respect to its position on the road. The assumption is that all rate allocations shown in the figure are feasible without violating the channel load threshold. Figure 5.8a shows a random distribution of transmission rates which does not fulfill the local fairness principle. In contrast, Figures 5.8b and 5.8c illustrate allocations which meet the local fairness principle, since neighboring vehicles are assigned similar transmission rates.

[4] Further note that the local fairness principle as stated in [25] has been designed with the adaptation of one control dimension in mind. As we discuss in Section 7.3 on page 191, the principle may require a redefinition for the case of multiple adapted control dimensions, e.g., in the sense that vehicles *contributing to* rather than being *located at* the same region of the wireless channel should be regulated to the same degree.



(a) Neither local nor global fairness

(b) Local but not global fairness



(c) Local and global fairness

Figure 5.8: Schematic illustration of local and global fairness from a rate control perspective (with uniform maximum and minimum rate)

5.3.2 Global Fairness

In [214], Torrent Moreno et al. discuss fairness in the context of congestion control for VSC. They argue that from a safety point of view, it does not make sense to increase the overall throughput in terms of transmitted Cooperative Awareness Messages (CAMs)/Basic Safety Messages (BSMs) while potentially throttling individual vehicles. These vehicles which are not able to make themselves sufficiently heard may become a safety risk to others who are not aware of their presence. The authors thus advocate the application of the max-min fairness principle. However, in a wireless channel, true max-min fairness in its formal sense is difficult to achieve due to the unbounded and probabilistic nature of radio-wave propagation. In [25], we thus used the term *global fairness* to describe a best-effort approach to fulfill the max-min fairness criterion.

Figure 5.8b schematically illustrates a transmission rate allocation which does not meet the global fairness criterion if a rate allocation as shown in Figure 5.8c is also feasible in the scenario without exceeding the channel load limit. The key difference between both rate allocations is that in Figure 5.8c, the minimum transmission rate is higher, which means a higher degree of global fairness. We are going to discuss further

examples of global fairness throughout the remainder of this thesis. For example, the rate allocation in Figure 6.19d on page 163 provides a higher degree of global fairness than the one in Figure 6.16d on page 159.

5.3.3 Participation Fairness

Intuitively speaking, the global fairness principle implies that all vehicles that *contribute* to congestion at a certain location should *participate* in congestion control. Typically, the carrier sense range is used as a boundary for this purpose, i.e., the distance up to which a receiving node can sense the signal of the transmission and thus considers the channel to be busy, e.g., [20][189][207]. A busy channel not only prevents a node from transmitting a message, it may also cause the node to reduce its transmission parameters if CBR measurements are used as a feedback in congestion control.

Assuming a deterministic radio-wave propagation, i.e., based on path loss only, the carrier sense range can be calculated as a fixed distance with a binary transition from carrier sensing to non-carrier sensing. However, if fading is taking into account, radio-wave propagation becomes probabilistic and the theoretically calculated carrier sense range becomes just one distance with a certain carrier sense *probability*. To account for the fact that the theoretical carrier sense range is not the only distance which could be used as the boundary for the participation in congestion control, we use the term *participation range* instead.

In this thesis, we set a node's participation range to the distance at which its transmissions can be carrier sensed with a probability of 10% (using a fixed carrier sense threshold). Based on the power-law model described in Appendix A.1 on page 223, we calculate a node's participation range as:

$$d_{part} = 10^{(P_{Tx}+z_q\sigma-L-P_{csN})/(10\gamma)} \tag{5.2}$$

where P_{csN} is the difference between carrier sense threshold and noise floor in dB according to Equation (2.7) on page 39, γ is the path loss exponent, z_q is z-value of the standard normal distribution corresponding to the desired percentile, e.g., 1.28 for 90%, and σ is the standard deviation of the lognormal noise.

To give an example, Figure 5.9 on the facing page illustrates the theoretical carrier sense probability with respect to distance using the power-law model with deterministic radio-wave propagation as well as with log-normal fading for a transmit power of 20 dBm and a carrier sense threshold of -95 dBm. We can see that under the deterministic model, there is a one-to-zero transition of the carrier sense probability. The theoretical carrier sense range amounts to 1283 m in our setup and can be calculated according to Equation (2.8) on page 39. The figure further shows the resulting participation distance according to Equation (5.2), which amounts to 2138 m in our configuration.

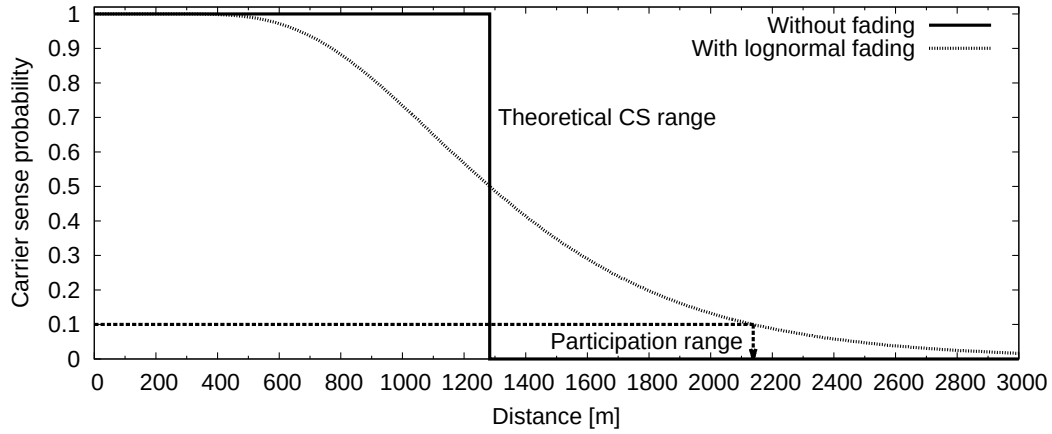


Figure 5.9: Theoretical carrier sense range and participation range for a transmit power of 20 dBm in our setup

According to the participation fairness principle, the transmitting node should participate in congestion control if congestion occurs within its participation range. To do so, the transmitter has to be *aware* of its contribution in the first place. In the reverse conclusion, the participation principle thus implies that an exchange of congestion information among nodes is necessary. We will get back to this topic in Section 6.4 on page 152, when we motivate and describe the information sharing mechanism of PULSAR.

5.4 Adaptation of the Relative Transmission Rate

In Section 4.1 on page 84, we analyzed the spatio-temporal aspects of awareness in order to define an optimization criterion for a congestion control strategy. After a series of considerations, we concluded in Section 5.1.3 on page 106 to select transmit power based on target distance and to adapt transmission rate with respect to channel load. However, different works analyzing the requirements of safety applications have pointed out the existence of a situation-dependent IRT requirement above which the safety application does not function, e.g., [24][210]. That is, if the safety application requires a certain IRT t_{Rx}, it does not make sense to apply any transmission rate below $1/t_{Rx}$, which might be the case if the transmission rate is adapted solely based on channel load.

In this section, we introduce an approach how both worlds, i.e., a situation-dependent transmission rate requirement and an adaptation of transmission rate based on channel load, could be consolidated with each other. In the first part of this section, we describe the general concept of adapting the *relative* transmission rate rather than the absolute transmission rate. The concept is based on a function mapping transmission rate to the safety benefit in a vehicle's current traffic situation and has been previously

described in [25]. In the second part, we give a concrete example of what such a function could look like. We conclude this section by describing how the relative rate control could be implemented using Additive Increase Multiplicative Decrease (AIMD).

It is important to note that the adaptation of the relative transmission rate as well as the proposed metric of Inter-Reception Distance (IRD) is still ongoing work. The findings presented in this thesis are more intended as a snapshot of the current state of development to provide a potential starting point for future studies. In future work, the proposed metric as well as the control method itself and its impact safety applications should be further investigated.

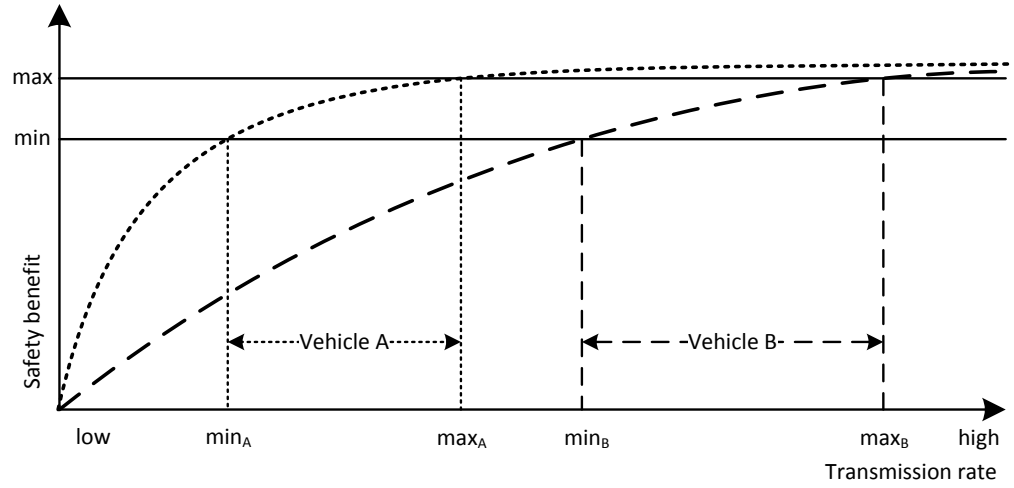
5.4.1 General Concept

As discussed in Section 5.1.4 on page 109, we expect transmission rate control to operate between a minimum and maximum value. These transmission rate boundaries could either be defined globally or be calculated individually for each vehicle based on its driving context. Safety applications are typically expected to have a situation-dependent maximum IRT, i.e., a minimum transmission rate, beyond which the application cannot work [24][165][210]. On the other hand, we do not expect that an infinite increase of transmission rate, even if it were tolerable by the channel, would result in an infinite increase of safety benefit. Rather, we expect a certain saturation point beyond which a further increase of transmission rate would result in a negligible additional safety benefit. For example, a tracking accuracy of 1 mm is most likely not required for the functionality of a safety application.

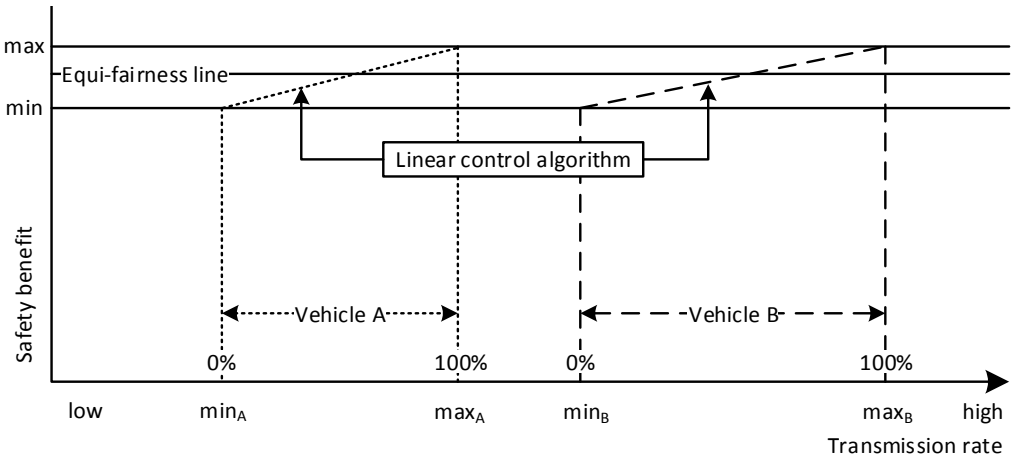
To sum up these considerations, Figure 5.10a on the facing page illustrates a sketch of a function mapping a vehicle's transmission rate to the achieved safety benefit. Note that "safety benefit" could be expressed in different metrics whose exploration we leave for future work. We provide a concrete example in Section 5.4.2 on page 126. The figure shows the expected shape of the mapping function for two vehicles in different driving contexts. For example, vehicle A could be a slow vehicle stuck in a traffic jam, while vehicle B could be passing on the other side of the highway at high speed under free-flow conditions. Assuming that a minimum and maximum safety benefit exist as discussed above, we can use these values to derive a minimum and maximum transmission rate for each vehicle.

Figure 5.10b on the facing page illustrates the application of a linear control algorithm between the transmission rate boundaries of each vehicle. While the safety benefit function may be nonlinear between the respective rate boundaries, we assume a linear control algorithm based on CBR measurements according to the findings presented in Section 5.2 on page 111. We leave the exploration of other options for future work. The linear control algorithm operates on the *relative* transmission rate which can be calculated as

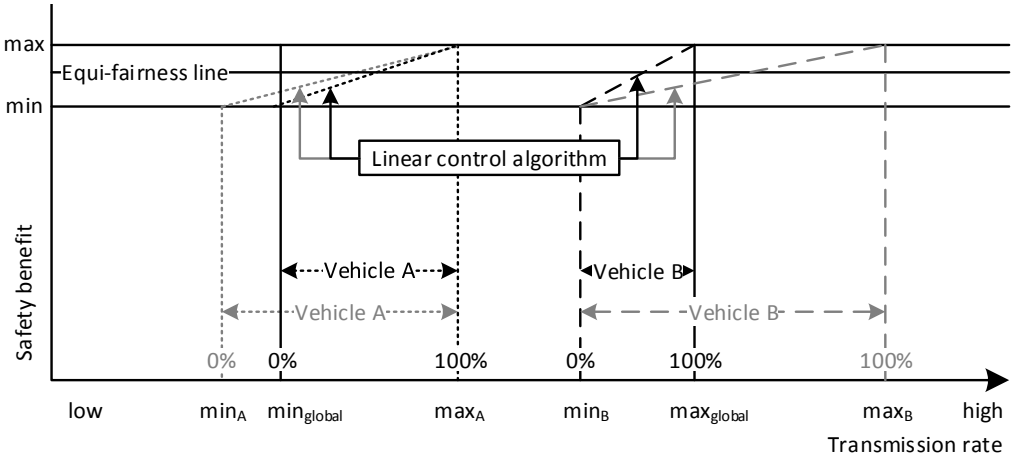
$$r_{rel} = \frac{r_{max} - r}{r_{max} - r_{min}} \in [0, 1] \tag{5.3}$$



(a) Derivation of transmission rate boundaries



(b) Adaptation of the relative transmission rate



(c) Alternatives if a global minimum and/or maximum rate exists

Figure 5.10: Illustration of the concept of relative transmission rate adaptation

where r, r_{min} and r_{max} denote the vehicle's absolute, minimum and maximum transmission rate, respectively. In other words, r_{rel} corresponds to a *fraction* of the interval $[r_{min}, r_{max}]$. Consequently, a rate allocation is considered to be *fair* if it assigns the same r_{rel}, i.e., the same safety benefit, to all participating vehicles.

Note that depending on the applied function mapping transmission rate to safety benefit, it is possible that the maximum transmission rate calculated for a particular vehicle exceeds a globally defined maximum rate, e.g., 10 Hz. Similarly, a global minimum rate of, e.g., 1 Hz could be desired. If this is the case, the question arises within which boundaries the relative transmission rate should be calculated. Figure 5.10c on the preceding page illustrates two options to integrate global transmission rate boundaries into the concept. In both cases, vehicle j's absolute transmission rate stays within the interval $[max(r_{min,j}, r_{min,global}), min(r_{max,j}, r_{max,global})]$. The difference, however, is the calculation of the relative transmission rate.

In the first option, indicated in black lines, $r_{rel,j}$ is calculated based on the interval $[max(r_{min,j}, r_{min,global}), min(r_{max,j}, r_{max,global})]$. The advantage of this approach is that if the channel is congested, all vehicles react immediately by reducing their absolute transmission rates. However, the approach may have consequences for fairness. In the figure for example, the new 50% value of vehicle B's r_{rel} corresponds to less than 25% of its actual interval $[r_{min,B}, r_{max,B}]$.

To retain fairness, the second option indicated using gray lines calculates vehicle j's transmission rate based on the interval $[r_{min,j}, r_{max,j}]$ and applies the global minimum and maximum rates as cut-off values. That is, while the relative transmission rate is further adapted, the absolute transmission rate of vehicle j does not follow. As a consequence however, this approach can result in situations where it takes a longer time to resolve channel congestion. In our example, vehicle B would not decrease its absolute transmission rate until a relative transmission rate of about 30% has been reached.

In this thesis, we decided to use the second option for two reasons. First, it retains fairness between different vehicles. Second, in the scenarios evaluated in this thesis the AIMD mechanism applied in the PULSAR protocol showed a sufficiently fast convergence to efficiency from a channel overload situation. We leave a further exploration of this topic for future work.

5.4.2 Concrete Example Based on Inter-Reception Distance

So far, we assumed that a function exists to calculate the "safety benefit" associated with a certain transmission rate. In the following, we introduce an example of what such a function could look like. Note that the following line of reasoning is similar to [24], where the An et al. present a detailed analysis of the spatio-temporal reception requirements of an Forward Collision Warning (FCW) application in order to avoid false negative as well as false positive warnings. The metric presented in this subsection can be considered a simplification and abstraction of the analysis by An et al. with

the primary objective of creating a simple yet reasonably realistic mapping function between transmission rate and safety benefit in order to evaluate the performance of the adaptation of the relative transmission rate in Section 7.3 on page 191.

We thereby continue the line of reasoning behind the target distance concept which focuses on the maximum relative speed between a vehicle and its neighbors, cf. Section 4.1.1 on page 84. For simplicity, we assume that the transmitting vehicle would like to optimize reception performance for a particular neighbor who is the most safety critical in the current situation. Further note that the metric has some similarities as well with the concept of tracking error [23]. In a sense, the proposed metric can also be understood as a *relative* tracking error between two vehicles.

As discussed in Section 4.1.1 on page 84, a safety application needs to warn the driver sufficiently ahead of time in order to avoid a potential crash. Further, the warning should stop if the danger no longer exists, cf. [24]. From a transmitting vehicle's point of view, this means that it has to make sure that its messages are received by the other vehicle at a sufficient distance and with a certain regularity in order to track changes in their relative distance.

Assuming a constant acceleration for both vehicles[5] as well as a constant heading, the change in the relative distance d_{rel} between two vehicles A and B within a time period $T = t_2 - t_1$ can be approximated as

$$d_{rel}(t_2) - d_{rel}(t_1) = (t_2 - t_1)(v_A - v_B) = T v_{rel} \tag{5.4}$$

If we interpret T as an IRT value, we can calculate the relative change in inter-vehicle distance between successive message receptions under the given assumptions. We denote this distance as Inter-Reception Distance (IRD) d_{Rx}.

Using Equation (4.5) on page 87, we can calculate the kth percentile of d_{Rx} based on the transmission interval t_{Tx}, the transmission rate r and the reception probability p_{Rx}:

$$P_k = t_{Tx} v_{rel} \left\lceil \frac{log(1-k)}{log(1-p_{Rx})} \right\rceil = \frac{v_{rel}}{r} \left\lceil \frac{log(1-k)}{log(1-p_{Rx})} \right\rceil \tag{5.5}$$

Note that analogously, we can derive the expected value of IRD by applying Equation (4.6) on page 87.

By solving for r, we can further calculate

$$r(d) = \frac{v_{rel}}{d} \left\lceil \frac{log(1-k)}{log(1-p_{Rx})} \right\rceil \tag{5.6}$$

[5] This assumption is likely the most limiting aspect of the proposed metric and is a significant simplification compared to the analysis of An et al. [24] who took into account potential changes in the relative speed as well. We make this assumption to keep our metric simple and leave a further exploration of the implications for a safety application for future work.

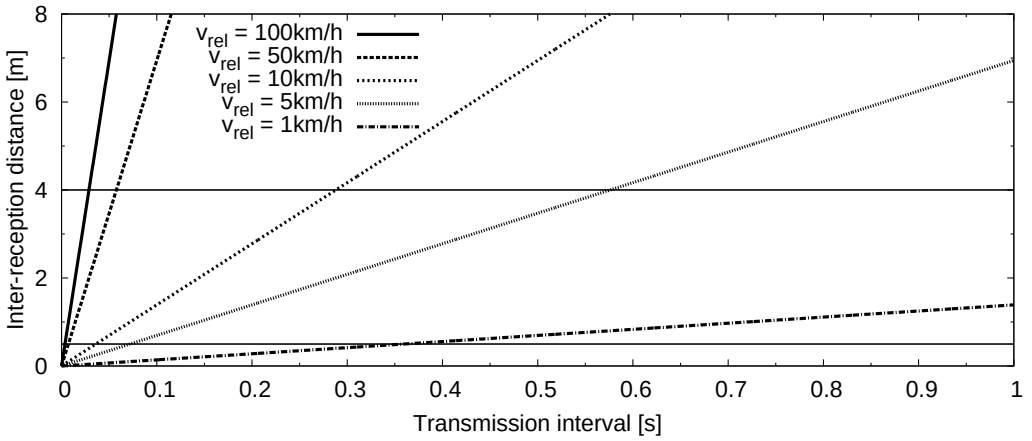
as the transmission rate for which the IRD d_{Rx} is smaller than d with probability k. We can use this equation to calculate a vehicle's minimum and maximum transmission rate based on its driving context as we discuss in the following.

Figure 5.11 on the facing page illustrates the resulting IRD according to Equation (5.5) with respect to transmission interval in Figure 5.11a and with respect to transmission rate in Figure 5.11b. We observe a linear relationship in the former case and a nonlinear relationship in the latter case. We will come back to this aspect later. Each subfigure further shows two horizontal lines at 0.5 m and 4 m, respectively. These values represent hypothetical values for a maximum and minimum performance requirement in the terms of a "safety benefit" as in Figure 5.10 on page 125. We thereby assumed that the maximum tolerable change in inter-vehicle distance to be one car length, i.e., approximately 4 m. On the other hand, we assumed changes in inter-vehicle distance below 50 cm to be negligible for the functionality of the safety application. Note that these values are intended only as examples.

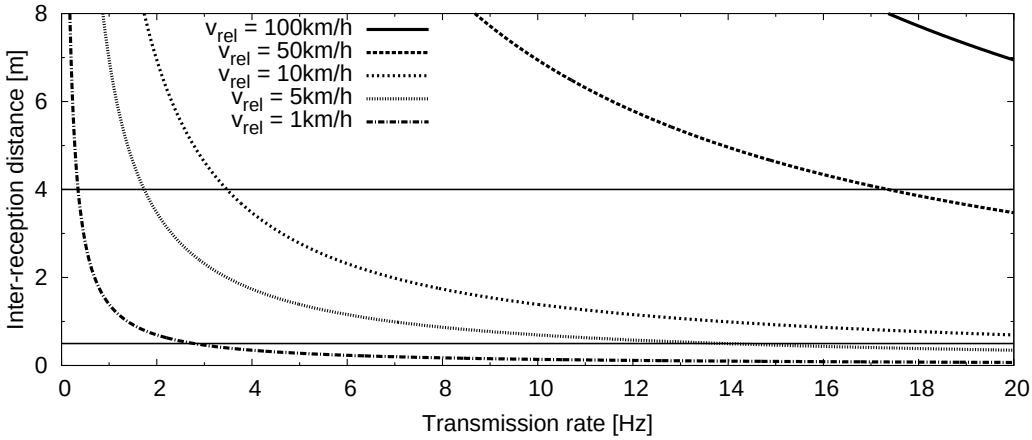
Figure 5.12 on page 130 provides a schematic illustration of how the relative transmission rate concept could be combined with IRD to adapt the resulting "safety benefit" with respect to channel load. In the figure, the left car approaches the one on the right. The crosses along the way represent message receptions by the approaching car. Note that while these points are equidistant in the figure for presentation clarity, they are intended to represent worst case IRDs based on when the last message was received. The actual IRDs may thus vary. The percentage value on the right corresponds to the relative transmission rate determined based on channel load. At 100% relative transmission rate, an IRD according to the maximum requirement is achieved, e.g., 50 cm in our example. At 0% relative transmission rate, on the other hand, the minimum performance requirement, e.g., 4 m, applies. The basic idea is that the minimum performance requirement should ensure a basic functionality of the application, while the maximum performance requirement could help to, e.g., reduce "false alarms" to a minimum[6]. The adaptation of the relative transmission rate can help to provide similar Quality of Service (QoS) levels in terms of safety benefit to all vehicles inside a certain area.

To conclude this introduction of the IRD concept, we would like to come back to the relationship between transmission interval and transmission rate. So far in this thesis, we focused on the adaptation of the latter due to its proportional relationship with channel load, cf. Section 5.2 on page 111. In contrast, there is an inversely proportional relationship between transmission interval and channel load as illustrated by Figure 5.13 on page 131. The figure shows the relationship between transmission rate, transmission interval and observed CBR in a scenario where 100 nodes share the same location. Note that the underlying data is the same as in Figure 5.4 on page 113.

[6] The suggested approach is similar to the introduction of a *tolerance region* for false positive and false negative warnings in [24]. However, An et al. did not discuss if and how the tolerance region could be adapted with respect to channel load. In future work, it should be evaluated how both approaches relate to each other.



(a) With respect to transmission interval



(b) With respect to transmission rate

Figure 5.11: Inter-Reception Distance (IRD) functions based on the 99.9th percentile of IRT (using $p_{Rx} = 0.8$)

Based on this observation, we conclude that the transmission rate is more suitable for a linear control than the transmission interval. We will come back to this topic in Section 7.4.2 on page 206, we evaluate the behavior of the protocol SOURC [202] which applies linear control to the transmission interval.

On the other hand, Figure 5.11 illustrates that IRD is proportional to the transmission interval, while it is inversely proportional to the transmission rate. That is, from a safety point of view the control of the transmission interval may actually be preferable given a metric based on IRT. In this thesis, we approach this conflict by using a linear interpolation between minimum and maximum transmission rate as illustrated by Figure 5.10 on page 125. Thereby, we prioritize the linearity regarding channel load over the linearity regarding IRD. In future work, other ways to approach this issue could be explored.

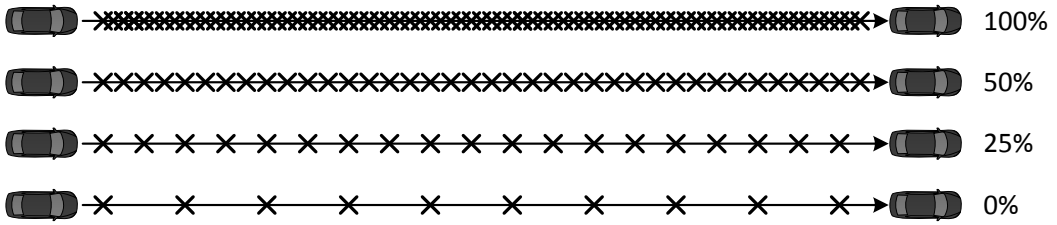


Figure 5.12: Schematic illustration of the relationship between maximum Inter-Reception Distance (IRD) and relative transmission rate

5.4.3 Implementation using AIMD

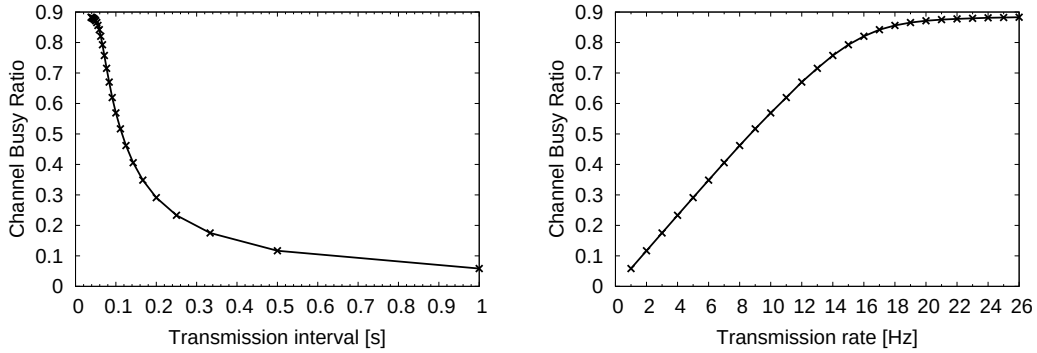
In the last part of this section, we illustrate how AIMD can be used to implement the adaptation of the relative transmission rate. The analysis by Jain et al. in [115] has shown that AIMD converges to fairness and efficiency if executed synchronously, cf. Section 3.1.3 on page 49. In the following, we use the vector representation of Jain et al. according to Figure 3.3 on page 52 to illustrate the convergence of AIMD if the relative allocation of resources is adapted instead of the absolute one.

To apply the notation of Jain et al., we use the example of a system with two users A and B and a total capacity of $c = 10$ units. If AIMD is applied in absolute terms, each user i adds α_I units to his allocation x_i if $x_1 + x_2 < c$ units. Otherwise, he calculates his new allocation as $(1 - \beta_D)x_i$ units.

If AIMD is applied in *relative* terms, each user adapts the *fraction* $x_{i,rel}$ of his minimum and maximum allocation interval $[x_{i,min}, x_{i,max}]$. He adds $\alpha_{I,rel}$ percentage points to his allocation $x_{i,rel}$ if $x_1 + x_2 < c$ units. Note that x_i represents each user's *absolute* allocation of resources. Otherwise, he calculates his new relative allocation as $(1 - \beta_{D,rel})x_{i,rel}$. When a new relative allocation has been determined, the absolute allocation is calculated as $x_i = x_{i,rel}(x_{i,max} - x_{i,min}) + x_{i,min}$.

Figure 5.14 on page 132 provides two examples of a relative AIMD adaptation using the vector representation of Jain et al. In Figure 5.14a, both users have identical minimum and maximum allocations, i.e., $x_{i,min} = 1$ units and $x_{i,max} = 10$ units. The left subplot shows the adaptation of the relative allocation $x_{i,rel}$, while the right subplot shows the resulting absolute allocation x_i. The area spanned by the minimum and maximum allocations of both users is indicated in gray shade. The figure contains AIMD's convergence from four different starting points which are indicated by circles. The starting points were chosen to cover each of the four possible combinations of feasibility and fairness in terms of overload/underload and preference of user 1/user 2. Notice that the convergence to fairness takes place in terms of relative allocation, while the convergence to efficiency takes place in terms of absolute allocation.

Figure 5.14b illustrates the convergence from the same relative starting points using a different set of allocation boundaries, i.e., $x_{1,min} = 1$, $x_{1,max} = 4$, $x_{2,min} = 4$ and $x_{2,max} = 10$. In contrast to the previous example, the allocation boundaries of the two users are



(a) CBR vs. transmission interval (b) CBR vs. transmission rate

Figure 5.13: Relationship of transmission rate, transmission interval and observed CBR in a scenario where 100 nodes share the same location

now asymmetrical. We observe that AIMD based on the relative allocation converges to fairness and efficiency in this configuration as well. Speaking in the terms of the vector representation, the adaptation of the relative allocation basically creates a fairness line within the gray rectangle. Notice that the figure contains an additional adaptation starting from the relative allocation 0.95, 0.1. From this starting point, relative AIMD would normally calculate an allocation of more than 100% for user 1, since the system is below the efficiency line. It is thus necessary to enforce the boundaries of the relative rate adaptation as shown in the example.

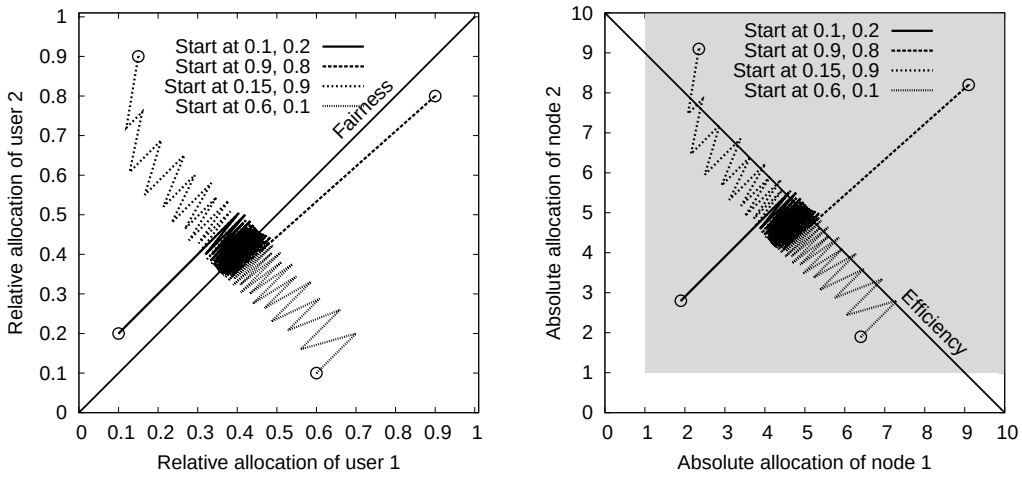
To better illustrate the behavior of relative AIMD near the equilibrium point, Figure 5.15 on page 133 plots the chronological sequence of the adaptation shown in Figure 5.14b. We can see that the typical saw tooth shape of the adaptation curve is retained, while the two users converge to different (absolute) allocations.

To conclude, the results presented in this section indicate that an implementation of the relative transmission rate concept based on AIMD is feasible. However, a formal proof is still required, which we leave for future work.

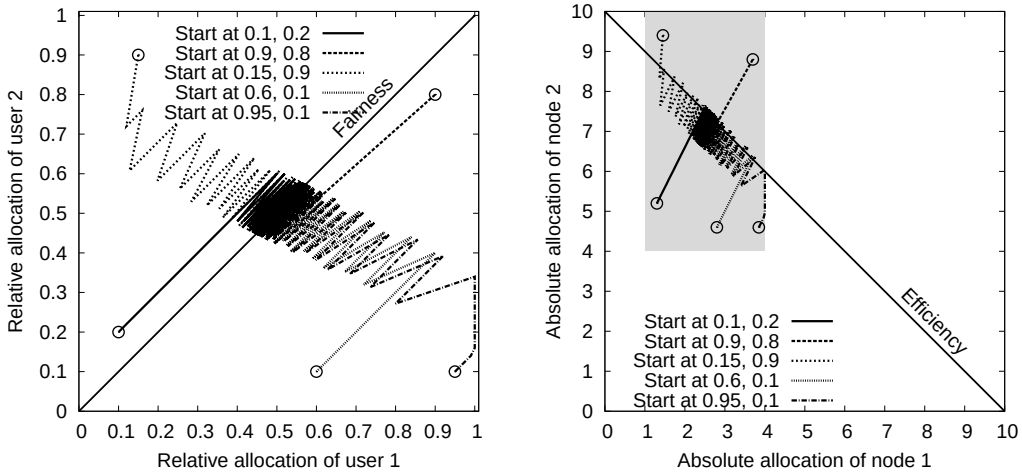
5.5 Summary of Design Principles

Over the course of this chapter, we have identified a number of desired properties of a congestion control protocol for VSC. We summarize these findings in the following five design principles [25]:

1. *Decentralization*. Since VSC based on IEEE 802.11p is meant to operate in ad-hoc mode, a congestion control protocol should be distributed.



(a) Using $x_{1,min} = x_{2,min} = 1$ and $x_{1,max} = x_{2,max} = 10$



(b) Using $x_{1,min} = 1$, $x_{1,max} = 4$, $x_{2,min} = 4$ and $x_{2,max} = 10$

Figure 5.14: Examples of relative AIMD adaptations using $\alpha_{I,rel} = 0.02$ and $\beta_{D,rel} = 0.2$ in the vector representation of Jain et al. [115]

2. *Local fairness*. Vehicles contributing to congestion in the same location should be regulated to a similar extent. Typically, such vehicles are located physically close to each other, cf. Section 5.3.1 on page 119.

3. *Global fairness*. If individual vehicles are throttled to increase overall system throughput, they may become a safety risk to others. Congestion control should thus follow a best-effort approach to converge to max-min fairness, cf. Section 5.3.2 on page 121.

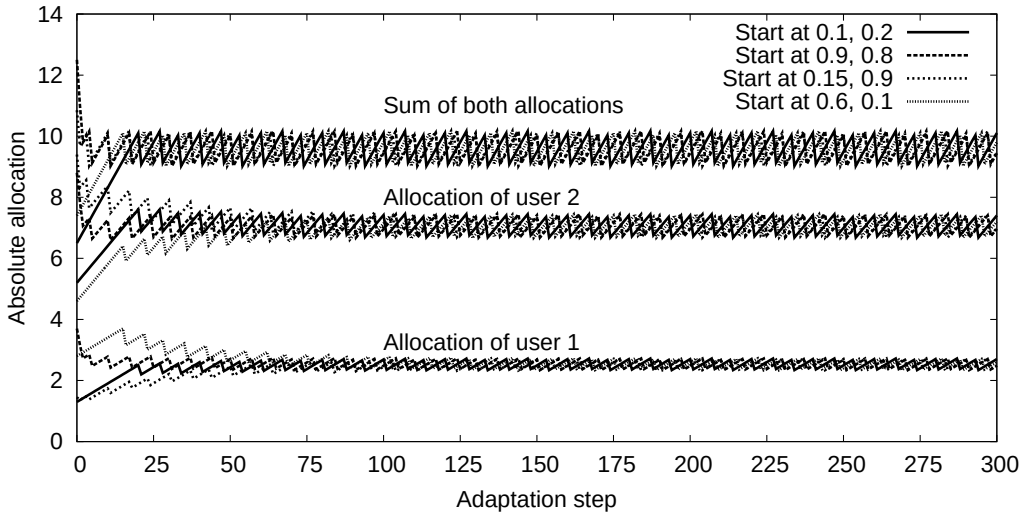


Figure 5.15: Convergence of relative AIMD over time (continued example)

4. *Participation fairness.* All vehicles contributing to congestion at a certain location to a non-negligible extent should participate in congestion control. For this, they need to receive feedback from their peers, cf. Section 5.3.3 on page 122.

5. *Deference to safety applications.* In order not to compromise safety, congestion control should work under the guidance of and operate within the boundaries set by safety applications, cf. Section 5.4 on page 123.

6

Protocol Design and Description

In this chapter, we flesh out the previously discussed design methodology and design principles with concrete protocol mechanisms and introduce Periodically Updated Load Sensitive Adaptive Rate control (PULSAR) as a resulting congestion control protocol for Vehicle Safety Communications (VSC) [25]. Since many design decisions are based on performance issues, this chapter deviates to a certain degree from a classical descriptive style. For each design decision, we provide examples why or why not we chose which alternative. In the following, we first provide an overview of the protocol's general structure. Then, we introduce each of its three building blocks separately and discuss the underlying design decisions and considered alternatives.

6.1 General Structure

The findings in Chapter 5 on page 103 suggest that for each target distance, there is a certain transmit power which optimizes reception performance. PULSAR thus uses an open loop controller[1] for transmit power based on the current target distance. Transmission rate, on the other hand, is controlled in closed loop using Channel Busy Ratio (CBR) measurements as a feedback, cf. Section 2.3.3 on page 40. The adaptation is triggered periodically and bounded by the minimum and maximum transmission rate set by the safety application(s). To meet the fairness principles

[1] A potential drawback of an open loop controller is the dependency on a correct look-up table. In future work, the protocol could be extended to use a closed loop controller for transmit power as well. The approach of Guan et al. [186] could be used as a starting point for this purpose.

derived in Section 5.3 on page 119, channel state information is shared among nodes and aggregated before adaptation. The protocol's general structure is summarized in Figure 6.1 on the facing page.

Note that PULSAR supports the dynamic change of the input parameters from safety applications. We expect target distance as well as minimum and maximum transmission rate to be adapted based on the driving context of the vehicle running the protocol. However, we expect these adaptations to take place at a different time scale, i.e., in the order of seconds rather than milliseconds.

In contrast to a proactive approach which tries to predict future channel load based on channel models, PULSAR is a reactive protocol which relies on feedback. We chose a reactive approach over a proactive one, since our objective is to operate near the "knee" of the channel load curve and thus a slight overload situation does not have detrimental effects on reception performance. In addition, the reactive approach depends less on channel models and is thus less vulnerable if the model does not match the real-world situation.

Another general design decision concerns the frequency of rate adaptations. While other approaches, e.g., [191], adapt transmission rate after each transmission event, we chose a periodic trigger for transmission rate adaptation with a fixed *adaptation interval* t_{adp}. Rate adaptation after constant time intervals leads to more fairness, since it provides the same number of adaptation opportunities to all nodes. For example, if adaptation and transmission events are coupled, nodes with lower transmission rate get less opportunities to increase their share of the channel, while those with already higher transmission rates get more opportunities.

PULSAR is structured in three logical building blocks, i.e., local channel load assessment, transmission rate adaptation and information sharing. The protocol's modular structure is intended to facilitate an exchange of mechanisms in the future. For example, the information sharing logic of PULSAR was used by Bansal et al. in [204] to aggregate feedback for the EMBARC rate control algorithm. In the following, we describe PULSAR's three building blocks in more detail.

6.2 Local Channel Load Assessment

In Section 5.2 on page 111, we saw that below the "knee" of the curve, the CBR follows approximately a linear relationship with the Communication Density (CD). Thus, we conclude that the CBR can be used to approximate the current CD in a scenario. In addition, IEEE 802.11 hardware supports CBR measurements due to standard specifications, cf. Section 2.3.3 on page 40. For these reasons, PULSAR uses CBR measurements as a control feedback.

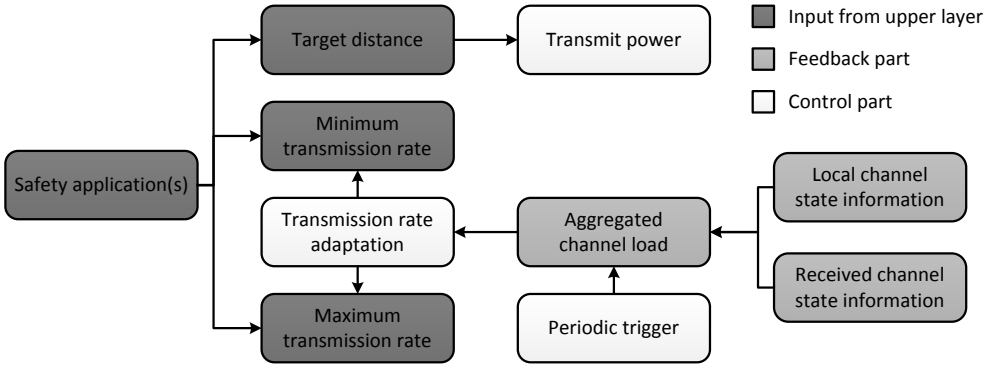


Figure 6.1: Overview and building blocks of the PULSAR protocol

6.2.1 CBR Filtering Techniques

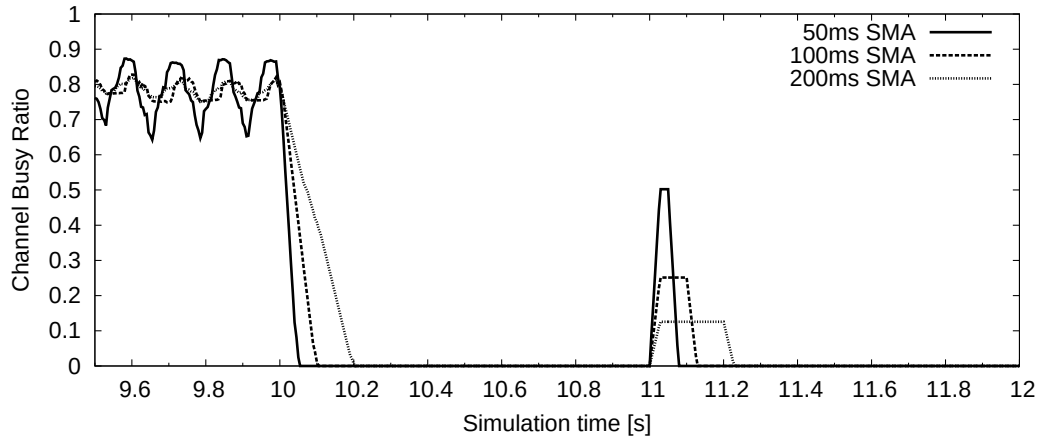
To calculate the CBR, the state of the channel is typically monitored for a certain amount of time, which we refer to as *channel monitoring time* t_{mon}. The CBR then denotes the fraction of time the channel was perceived as busy by the Medium Access Control (MAC) layer during this period. In essence, this form of CBR calculation corresponds to a Simple Moving Average (SMA) of true or false channel samples within a sliding window of length t_{mon}, cf. *channel probing* in Section 2.3.3 on page 40. An SMA is a simple low-pass filter which allows the controller to react to long-term changes in the system by filtering out high frequencies.

Simple Moving Average (SMA)

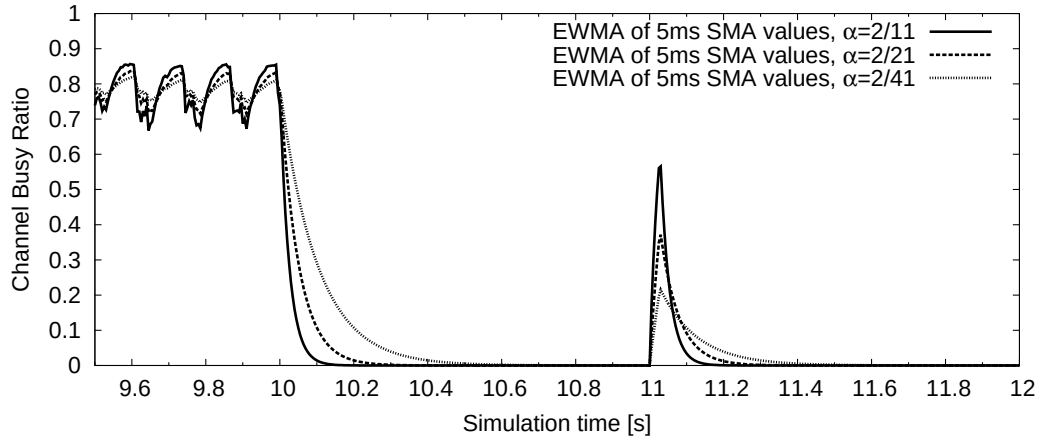
Figure 6.2a on the following page illustrates the resulting CBR measurements using an SMA of different t_{mon} triggered every 5 ms in a scenario with 200 nodes sitting in the same location. We use this scenario in the following to study the behavior of different mechanisms without the influence of fading[2]. In the simulation, all nodes transmit at 8 Hz until they stop transmitting abruptly at simulation time 10 s. There is no transmission rate adaptation except for *on* or *off* in this scenario. At simulation time 11 s, all nodes transmit exactly one packet before going silent again, emulating an "impulse response". We observe that before simulation time 10 s, the CBR shows oscillations near a value of 0.8. These result from a random initial time offset at which each node starts its transmission as well as from a random jitter for each transmission. If the initial transmissions had been distributed evenly within the transmission interval and no jitter had been applied, the CBR would remain nearly constant at a value of approximately 0.8.

Comparing the different CBR curves, we observe that a longer t_{mon} effectively smoothes out oscillations, but introduces a feedback delay. Its impulse response is finite and box-shaped as we can see near simulation time 11 s. According to

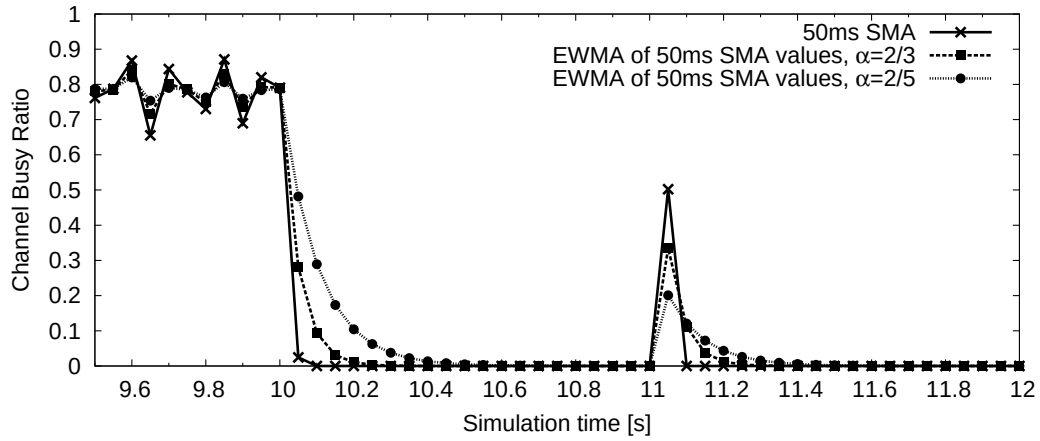
[2] This setup was also used by Kenney et al. in [71].



(a) Simple Moving Average (SMA), triggered every 5 ms



(b) Exponentially Weighted Moving Average (EWMA) of 5ms SMA values, triggered every 5 ms



(c) EWMA of 50ms SMA values, triggered every 50 ms

Figure 6.2: Impact of different filtering techniques on the observed CBR

[215], the SMA is a very good smoothing filter, i.e., in the time domain, but it has a poor performance in the frequency domain, since its frequency response is a sinc function whose "ripples" let through some high frequencies. This may cause "spikes" in the adaptation process.

Exponentially Weighted Moving Average (EWMA)

As an alternative to the SMA, we consider an Exponentially Weighted Moving Average (EWMA) with averaging parameter α in the form

$$U_{avg} = (1 - \alpha) \cdot U_{avg} + \alpha \cdot U_{new} \tag{6.1}$$

where U_{new} is the newly obtained CBR sample.

In this thesis, we assume that the radio offers a function to retrieve an SMA-based CBR value at any point in time with a configurable interval t_{mon}. This assumption is based on unofficial information from hardware suppliers. We thus use SMA-based CBR values as an input for Equation (6.1) rather than true or false channel samples. We approximate the latter behavior by using very small values of t_{mon}, e.g., 5 ms.

The EWMA is a low-pass filter like the SMA, but it assigns a higher weight to newer samples and theoretically considers an infinite history of previous input. It has an infinite frequency response without the "ripples" of the sinc function. Thus, it is less likely to prodcuce "spikes" in the adaptation.

To approximate the behavior of a corresponding SMA, the averaging factor α which assigns approximately 86% of the weight to the first N data points can be calculated as[3]

$$\alpha = \frac{2}{N + 1} \tag{6.2}$$

For example, if we would like to calculate α for a decay over 100 ms and we use 5 ms SMA values as an input, we get $N = 20$ and $\alpha = 2/21 \approx 0.095$.

Figure 6.2b on the facing page illustrates the CBR calculated as an EWMA of 5 ms SMA values in the scenario considered above. We observe that while lower values of α smooth out the oscillations resulting from the random spacing of transmissions as well, the EWMA is generally more sensitive to these oscillations than the SMA.

Self-Averaging: A Combination of SMA and EWMA

As a third alternative, we consider a combination of SMA and EWMA which we denote as *self averaging* [25]. The CBR is calculated according to Equation (6.1) and uses SMA samples with $t_{mon} = t_{adp}$ as an input. The CBR over time resulting from the suggested self-averaging approach in the scenario considered above is depicted

[3] A derivation of this relationship can be found in textbooks on statistics or at *http://en.wikipedia.org/wiki/Moving_average*.

in Figure 6.2c on page 138. The intention behind the self-averaging approach is to combine the low sensitivity to small-scale oscillations of the SMA with the graceful degradation behavior of the EWMA while introducing a comparable feedback delay to either of the two approaches.

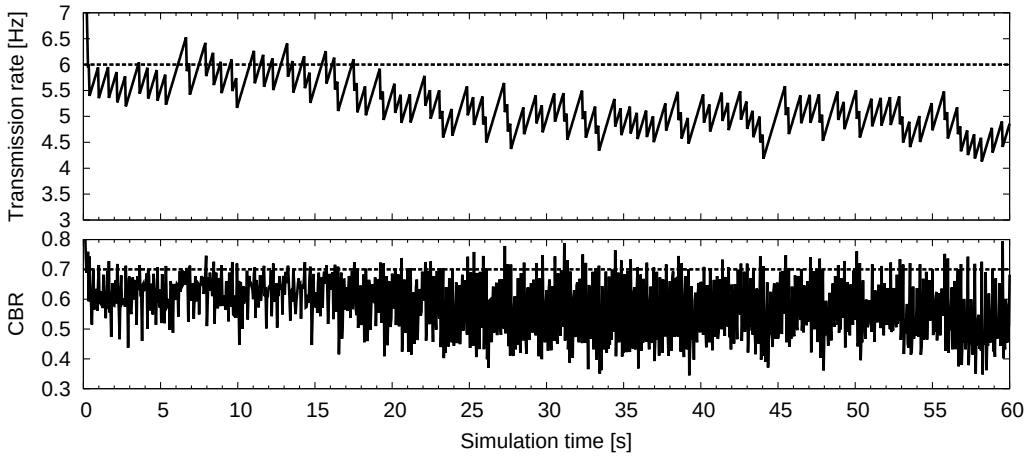
6.2.2 Impact of the Nyquist-Shannon Sampling Theorem

To achieve a stable control, the *Nyquist-Shannon sampling theorem* has to be fulfilled. The theorem "indicates that a continuous signal can be properly sampled, only if it does not contain frequency components above one-half of the sampling rate" [215]. Applied to our case, this means if we adapt in intervals of t_{adp}, i.e., at a rate of $1/t_{adp}$, our CBR filter needs to filter out all frequencies above $1/(2t_{adp})$. For example if $t_{adp} = 100\,ms$, an SMA filter has to be based at least on an interval of 200 ms. Note, however, that the sampling theorem assumes a perfect low-pass filter. In practice, a factor of approximately 2.2 is typically applied [215].

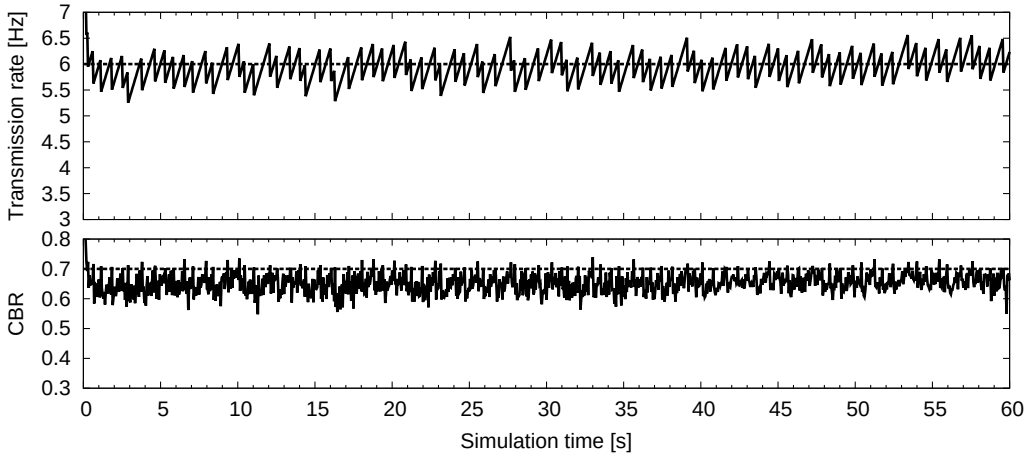
For the design of PULSAR, we evaluated the impact of the three CBR averaging techniques described in the previous subsection. The obtained results indicate that Additive Increase Multiplicative Decrease (AIMD), i.e., the increase/decrease mechanism applied in PULSAR, is robust against the choice of a low-pass filter, as long as the sampling theorem is fulfilled. Figure 6.3 on the next page illustrates an example of the development of transmission rate and measured CBR over time based on the same initial setup but different CBR feedback. The figure is based on the same scenario as above where 200 nodes are located at the same position. All nodes adapt their transmission rates simultaneously every 50 ms using AIMD with parameters $\alpha_I = 0.05$ and $\beta_D = 0.1$ and the rescheduling mechanism described in Section 6.3.1 on page 142. They initially pick a random time offset for their first transmission and apply a 5 ms random jitter to each transmission.

Figure 6.3a on the next page shows the behavior of the system based on an SMA with $t_{mon} = t_{adp} = 50$ ms. We observe large oscillations between 0.3 and 0.8 CBR, i.e., the system is not stable. In this case, the sampling theorem is not fulfilled. In contrast, Figure 6.3b on the facing page shows the adaptation process based on the self-averaging approach. The 50 ms SMA samples are now averaged in the EWMA to provide a comparable decay to a 100 ms SMA, i.e., using $\alpha = 2/3$. We can see that the oscillations have been significantly reduced and are now centered around the equilibrium point near 6 Hz.

For the remainder of this thesis, we use the self-averaging approach as described above using $\alpha = 10/7$, which corresponds to an SMA using an interval of $t_{mon} = 2.5t_{adp}$. We chose this value to be on the safe side regarding the sampling theorem. Note that while AIMD appears to be robust against the choice of CBR filter, it may be preferable to apply an EWMA filter based on smaller time samples for other



(a) 50 ms SMA (sampling theorem not fulfilled)



(b) EWMA of 50 ms SMA values, $\alpha = 2/3$ (sampling theorem fulfilled)

Figure 6.3: Example of an AIMD adaptation based on CBR filtering with and without the fulfillment of the sampling theorem

increase/decrease algorithms, e.g., LIMERIC [71]. We leave a further exploration of this topic for future work.

6.3 Transmission Rate Adaptation

PULSAR is based on binary rate adaptation using the AIMD mechanism, cf. Section 3.1 on page 46. That is, at each periodic rate adaptation event k, the new transmission rate r_{k+1} is calculated as

$$r_{k+1} = \begin{cases} r_k + \alpha_I & \text{if } U_k \leq U_t \\ (1 - \beta_D) r_k & \text{if } U_k > U_t \end{cases} \tag{6.3}$$

where U_k is the current CBR value and U_t is the CBR target. The parameters α_I and β_D determine the convergence behavior of AIMD, as we discuss in detail in Section 7.1.1 on page 168.

6.3.1 Message Rescheduling

When a new transmission rate r_{k+1} has been calculated, it needs to be applied to the periodic transmission regime. At this point in time, the last transmission was typically less than $1/r_k$ seconds ago, where r_k denotes the previous transmission rate. That is, t_{rem} seconds would still be remaining until the next transmission event.

We consider the following options for the transition from r_k to r_{k+1}:

(a) The next transmission takes place as planned, i.e., after t_{rem}. Then, the new transmission interval is applied[4] as $t_{Tx} = 1/r_{k+1}$.

(b) The next transmission is *rescheduled* [25] to take place after $t_{resc} = t_{rem} \cdot r_k/r_{k+1}$ seconds. Then, $t_{Tx} = 1/r_{k+1}$.

(c) The next transmission takes place after a random interval from within $[0, 1/r_{k+1}]$. Then, $t_{Tx} = 1/r_{k+1}$.

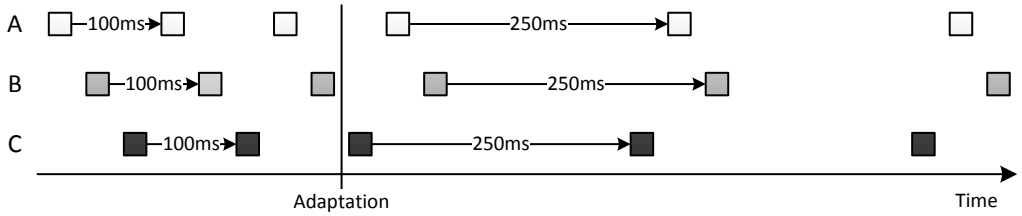
Note that option (c) may result in the starvation of individual nodes, e.g., if $t_{Tx} > t_{adp}$, when the next transmission may be postponed multiple times in a row. In the following, we thus focus on the first two alternatives which take into account the previous transmission.

Impact on Transmission Distribution

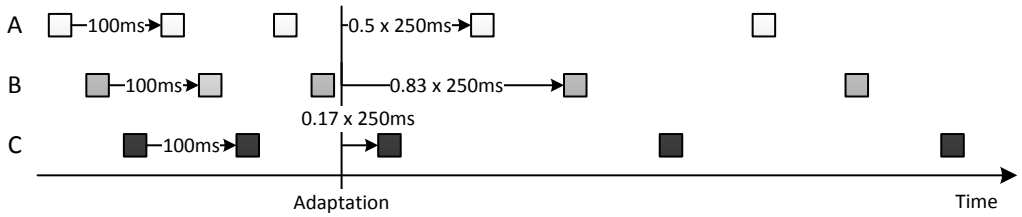
The two subfigures of Figure 6.4 on the next page illustrate schematically the impact of options (a) and (b), respectively, when decreasing transmission rate. The boxes in the figure represent the transmissions of three nodes A, B and C over time, starting at a transmission rate of 10 Hz. At the vertical line, the transmission rate of all nodes is set to 4 Hz. In Figure 6.4a, we observe that option (a) leads to a clustering of transmissions, while large gaps appear between the clusters. In contrast, the *rescheduling* mechanism [25], i.e., option (b), retains the original distribution of transmissions, as depicted in Figure 6.4b. It does so by calculating the remaining fraction of the previous transmission interval and applying it to the new transmission interval. This way, transmissions are more spread out over time compared to option (a), which reduces the risk of packet collisions.

Analogously to Figure 6.4, Figure 6.5 on page 144 illustrates the impact of the rescheduling mechanism when increasing transmission rate, i.e., from 4 Hz to 10 Hz in the example. In Figure 6.5a, we observe that using option (a) leads to a delayed reaction

[4] We abstract from the usage of a random jitter for presentation clarity.



(a) Without message rescheduling: Bursty behavior



(b) With message rescheduling: Transmissions more spread out over time

Figure 6.4: Impact of the rescheduling mechanism when decreasing transmission rate; boxes represent transmissions over time by exemplary nodes A, B and C.

to rate adaptation as well as a slight clustering effect. Note that it is possible to adapt transmission rate multiple times until the adaptation takes effect on the system. In contrast, the rescheduling mechanism in Figure 6.5b causes the adaptation to take immediate effect and retains the original distribution of transmissions.

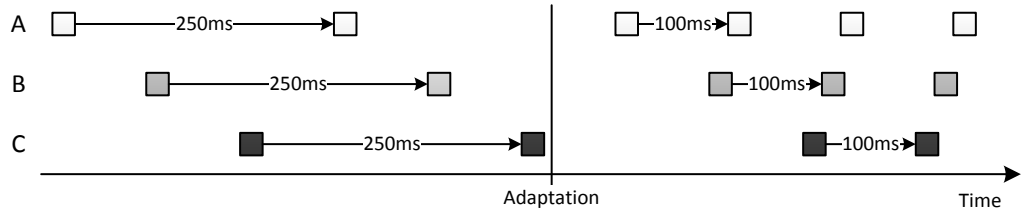
Impact on the Observed CBR

Figure 6.6 on page 145 illustrates the impact of the rescheduling mechanism on the observed CBR in the previously applied simulation scenario with 200 nodes placed in the same location. Initially, all nodes start at 10 Hz, choosing a random offset for their first transmission and applying a jitter of 5 ms. Then, transmission rate is adapted as shown in Figure 6.6a.

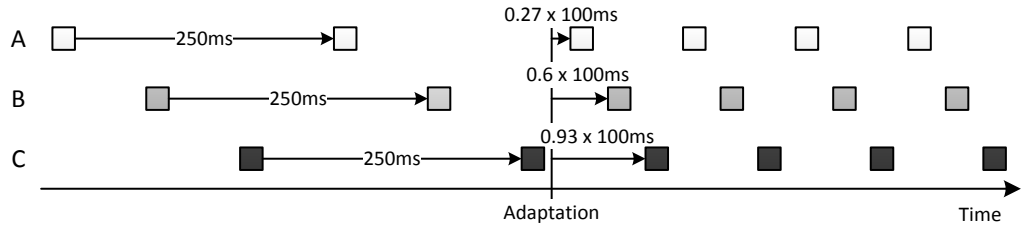
Figures 6.6b and 6.6c illustrate the observed CBR over time without and with the rescheduling mechanism, respectively. The CBR is measured every 100 ms as an SMA of the same length. We observe that without the rescheduling mechanism, i.e., using option (a), the CBR shows large oscillations resulting from the clustering effect described above. In contrast, the oscillations with the rescheduling mechanism are much smaller and result from the randomly chosen initial transmission offset.

Impact on Transmission Rate Adaptation

Figure 6.7 on page 146 illustrates how the rescheduling mechanism affects the AIMD-based rate adaptation in PULSAR. In the example, the initial transmissions of the 200 nodes in the scenario were spread out evenly throughout the transmission interval and no jitter was applied. With message rescheduling, i.e., in Figure 6.7b, we observe a



(a) Without message rescheduling: Adaptation is delayed



(b) With message rescheduling: Adaptation takes immediate effect

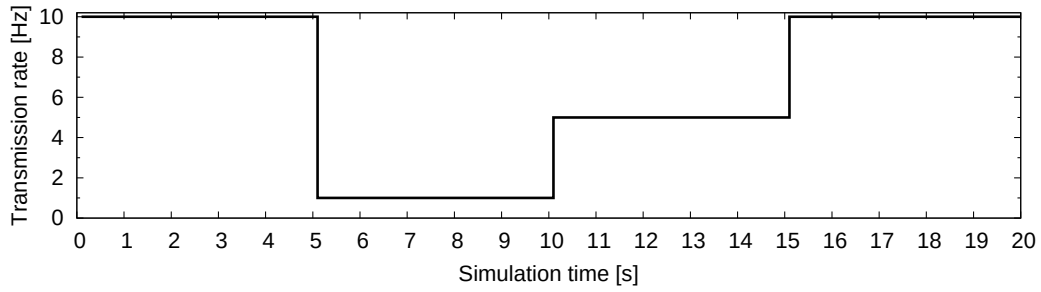
Figure 6.5: Impact of the rescheduling mechanism when increasing transmission rate; boxes represent transmissions over time by exemplary nodes A, B and C.

perfectly saw-tooth shaped adaptation curve. Without message rescheduling, however, we can see that after some initial oscillations, the system converges to a state in which all nodes send at 10 Hz and the CBR remains near 0.63. This result conflicts with the initially measured CBR of 0.88 at simulation time 0 s when all nodes were sending at 10 Hz as well. What happens is that the aforementioned clustering effect causes their transmissions to overlap. The collided transmissions require less time on the channel and thus the CBR is lowered. Note that the synchronization of transmissions also occurs if a random initialization and jitter are applied.

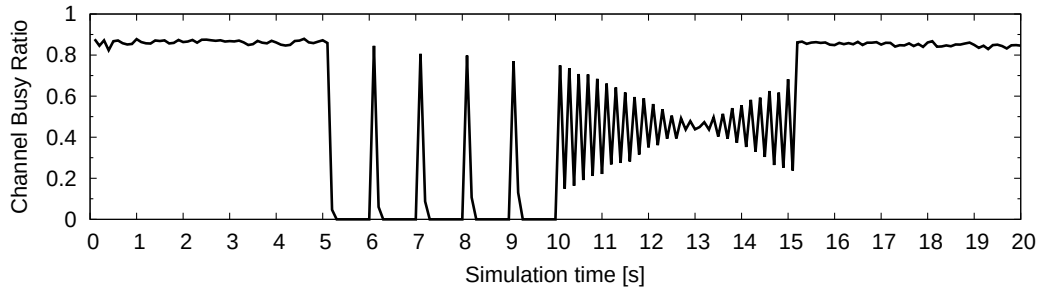
In summary, message rescheduling is a subtle but important mechanism in the PULSAR protocol to ensure that rate adaptation takes immediate effect and no transmission bursts are induced.

6.3.2 Synchronization

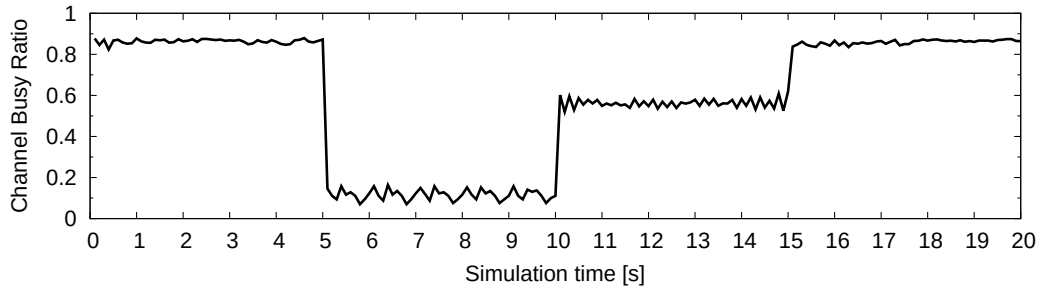
In TCP, the most prominent protocol employing the AIMD mechanism, nodes execute the rate adaptation without explicit synchronization. In fact, a synchronization is neither feasible nor desired, since AIMD is typically assumed to converge to fairness. However, as pointed out by Gorinsky and Vin in [121] and [122], this is not necessarily the case in an asynchronous execution.



(a) Transmission rate



(b) Observed CBR without rescheduling

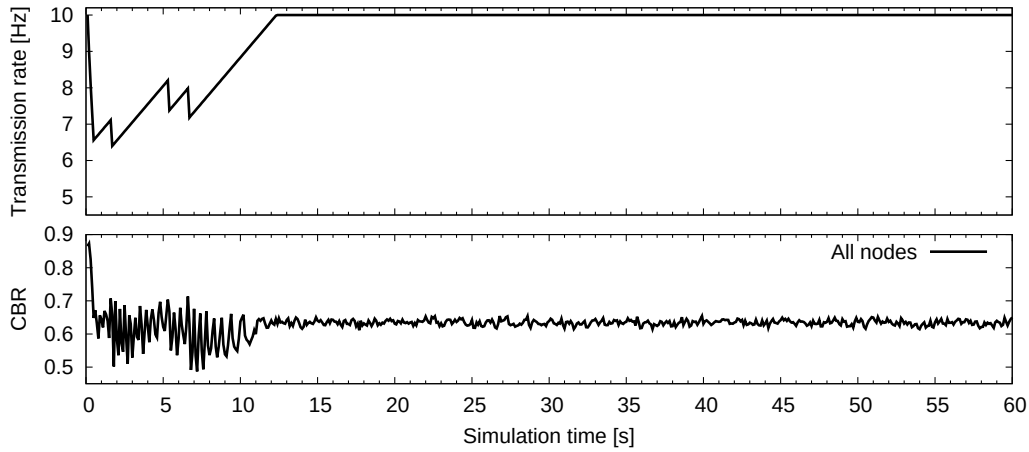


(c) Observed CBR with rescheduling

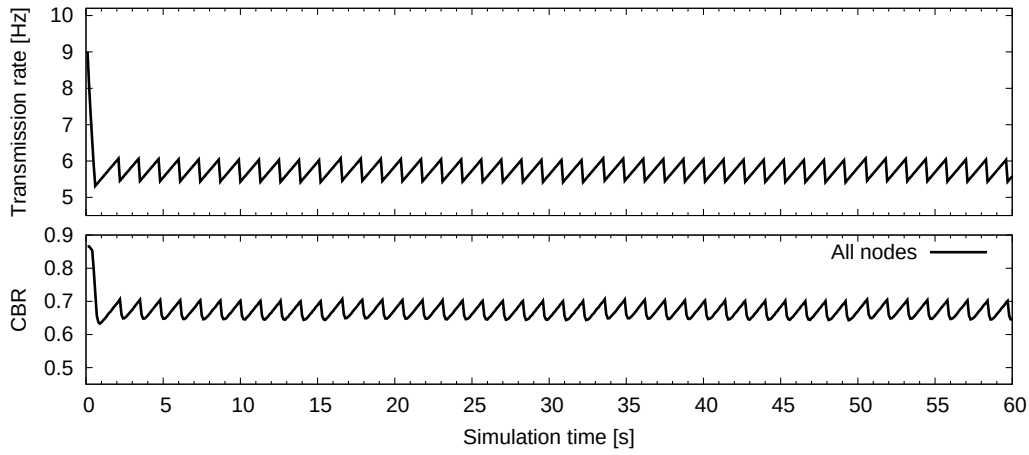
Figure 6.6: Impact of the rescheduling mechanism on the observed CBR

A Simple Example

Figure 6.8 on page 147 illustrates a simple example where an asynchronously applied AIMD converges to fairness in one case and does not converge to fairness in another case. The setup consists of three nodes A, B and C which share a resource with a maximum allocation of 15 units. The three nodes adapt their respective share of the resource in the order A, B, C, A, etc. using AIMD with $\alpha_I = 1$ and $\beta_D = 0.5$. In Figure 6.8a, all nodes start with an initial allocation of 5 units each. We observe that each node performs a similar adaptation in the typical saw-tooth shaped curve of AIMD. In steady state, all nodes receive the same amount of resources, albeit time shifted. The system has converged to fairness.



(a) Adaptation without rescheduling

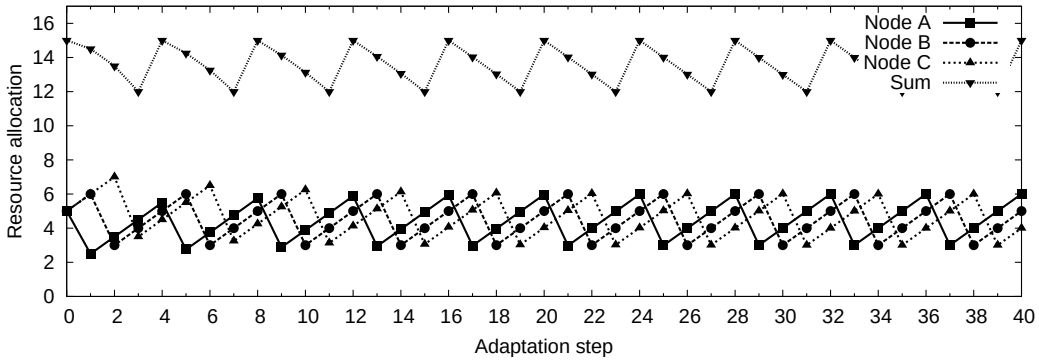


(b) Adaptation with rescheduling

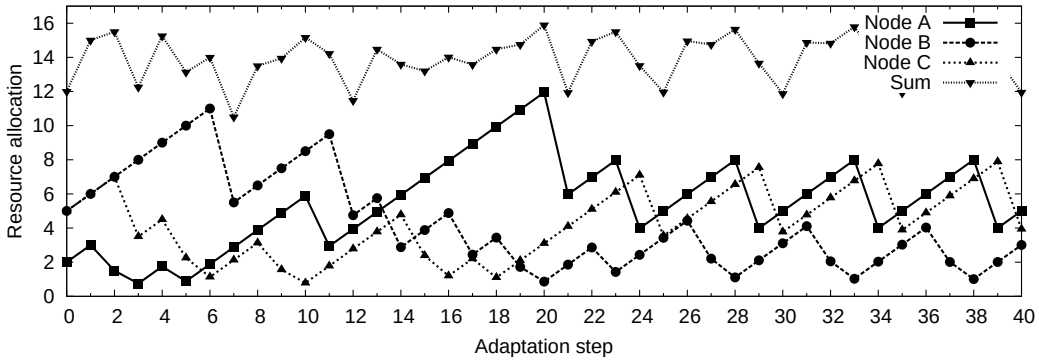
Figure 6.7: Impact of the rescheduling mechanism on AIMD-based rate adaptation in a scenario excluding radio-wave propagation effects (200 nodes, α_I = 0.05 Hz, β_D = 0.1); synchronized adaptation with deterministic initialization and no jitter

Figure 6.8b shows the convergence behavior of AIMD in the same setup with an initial allocation of 2, 5 and 5 for nodes A, B and C, respectively. We observe that the system does not reach steady state until adaptation step 26. Even then, node B gets a significantly lower share of the resource than nodes A and C. Note, however, that the system converges quickly to efficiency, as we can tell from the sum of the share of the three nodes.

In the considered setup, whether or not the system converges to fairness depends on the combination of AIMD parameters, maximum resource allocation and initial setup. If more than three nodes are part of the system, the outcome becomes increasingly difficult to predict.



(a) Convergence to fairness with initial allocation 5-5-5



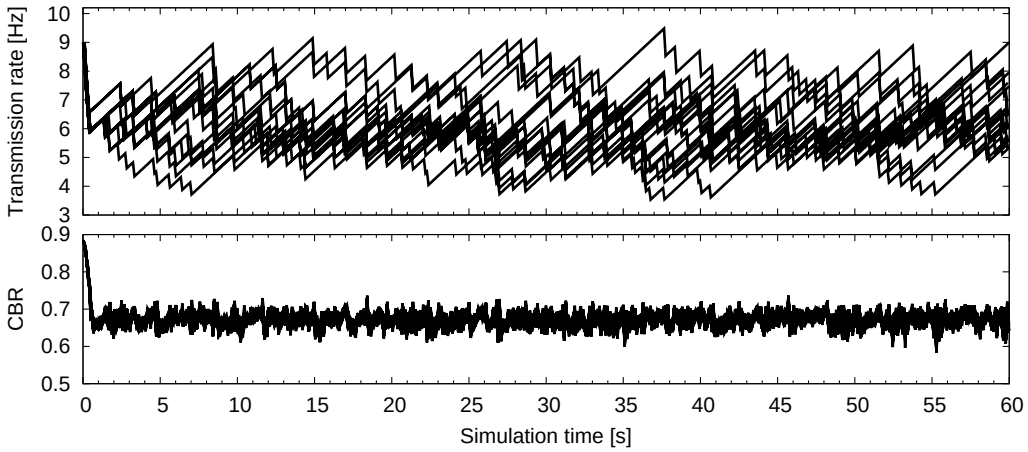
(b) Convergence to unfairness with initial allocation 2-5-5

Figure 6.8: Example of an asynchronous AIMD adaptation with three nodes and a maximum allocation of 15 units

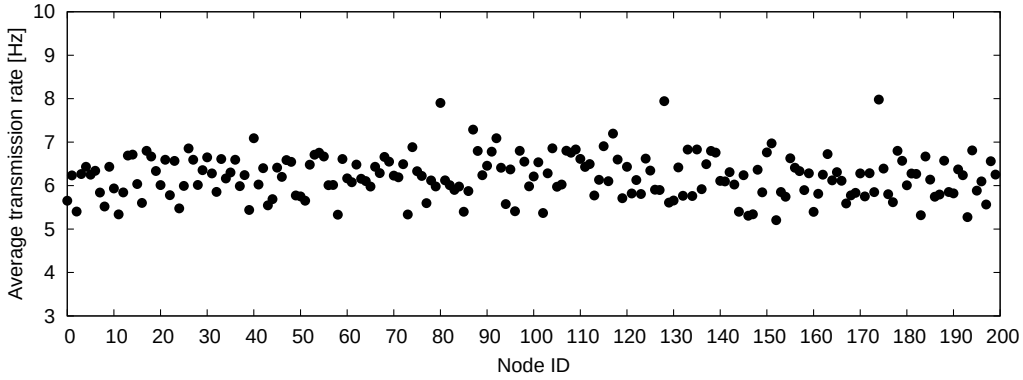
Synchronization in PULSAR

In PULSAR, the asynchronous execution of AIMD typically leads to a convergence to efficiency, but not to fairness. Figure 6.9a on the following page illustrates an asynchronous AIMD-based rate adaptation in the scenario considered above. Initially, each node picks a random time offset to begin its periodic rate adjustments. The figure shows the transmission rates and measured CBR values of 20 exemplary nodes. We observe that while the CBR stays between 0.6 and 0.7, the transmission rates of the different nodes show large oscillations between 3.5 Hz and 10 Hz. In other words, if we assume that all nodes have the same communication demand from a safety perspective, the fair rate allocation between the participating nodes at a particular point in time is compromised. The 60 s time average of the transmission rate of each node is presented in Figure 6.9b. We can see that a certain degree of long-time fairness among nodes is given.

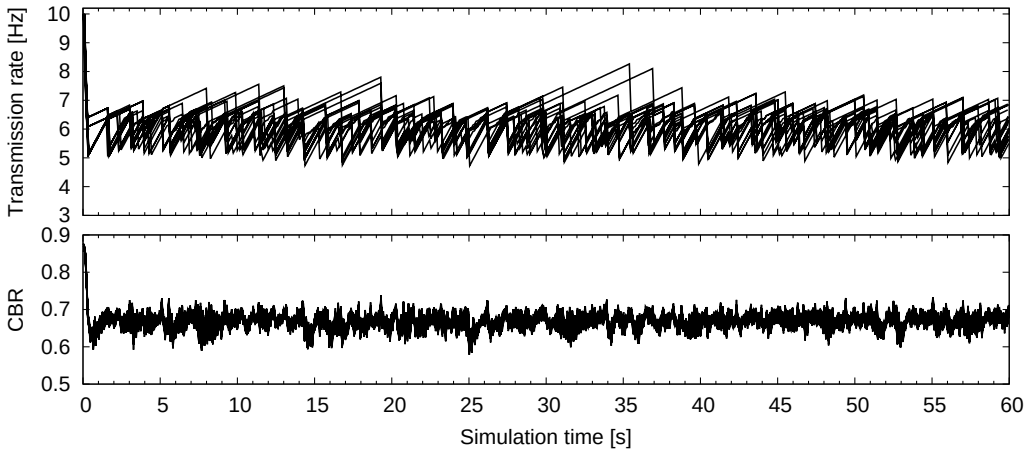
However, in the context of VSC, the instantaneous fairness is at least as important as the long-term fairness, since vehicles which are not able to communicate their current state sufficiently may become a safety risk to others. In PULSAR, we thus



(a) Asynchronous adaptation



(b) 60s average of transmission rate per node with asynchronous adaptation



(c) Asynchronous adaptation with target rate

Figure 6.9: Asynchronous AIMD-based rate adaptation in a scenario excluding radio-wave propagation effects (200 nodes, $\alpha_I = 0.05\,\text{Hz}$, $\beta_D = 0.1$); with rescheduling, random initialization and jitter

assume that nodes are synchronized in their AIMD adjustments. A synchronization could be realized by means of the Global Positioning System (GPS) clock, which has an accuracy in the order of nanoseconds [216], i.e., multiple orders of magnitude higher than the transmission rate adaptations.

Note that in case nodes start drifting apart in spite of the synchronization via GPS, the target rate mechanism introduced in the following subsection can serve as a backup mechanism. As illustrated by Figure 6.9c on the preceding page, the target rate mechanism does not prevent the nodes from drifting apart in an asynchronous adaptation, but it limits the resulting degree of unfairness.

6.3.3 Target Rate Mechanism

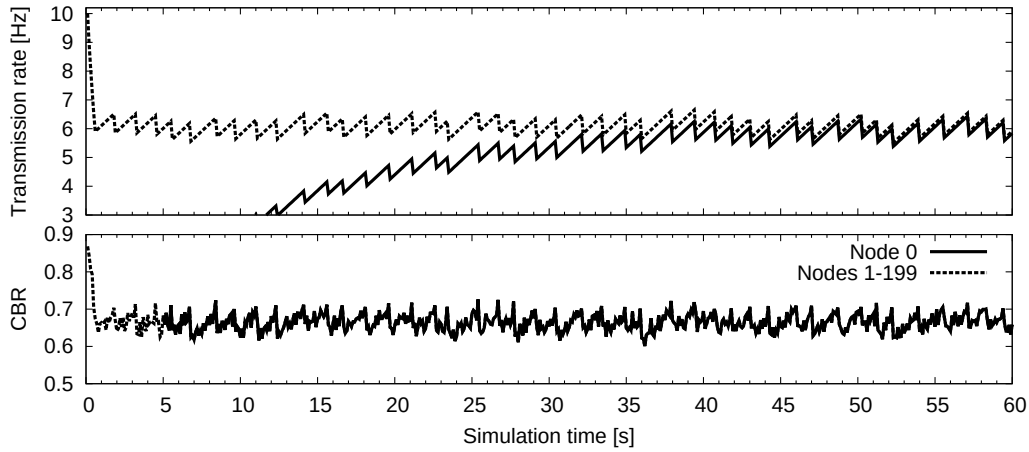
A well-known issue of AIMD is its relatively slow convergence to fairness, cf. e.g. [217]. Unless the additive increase parameter is chosen aggressively, a new node entering the system from a low starting point may require a long time to reach its fair share of the resource. An aggressive parameter choice, however, leads to increased oscillations, cf. Section 3.1 on page 46. A major factor in the slow convergence to fairness is that the new node has to decrease its share with the others every time the capacity limit is reached. Figure 6.10a on the next page illustrates this phenomenon in the previously considered scenario with 200 nodes in the same location. One of the nodes enters the system with a 5 s delay, starting at 1 Hz. We can see that even after 30 s simulation time, the system has not yet converged to fairness.

The slow convergence of AIMD is undesirable in a vehicular network where nodes are moving fast, since it may lead to a violation of the *local fairness* principle, cf. Section 5.3 on page 119. On the other hand, aggressive AIMD parameters are also undesirable, since they lead to increased oscillations. In PULSAR, we thus introduce an additional mechanism to accelerate the convergence speed of AIMD while using non-aggressive parameters. We thereby make use of the unique situation that we can obtain direct feedback on the current usage of channel resources by neighboring vehicles. By piggybacking transmission rates in Cooperative Awareness Messages (CAMs)/Basic Safety Messages (BSMs), vehicles can quickly assess the average transmission rate of their neighbors. We denote this average rate as *target rate* [25].

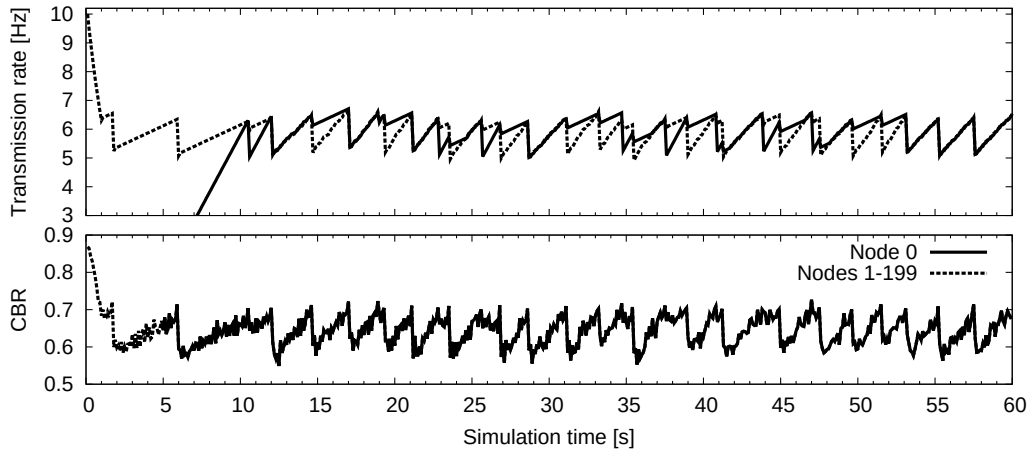
When a new transmission rate information r_{new} has been received, a node can update its target rate as

$$r_t = (1 - \alpha) \cdot r_t + \alpha \cdot r_{new} \tag{6.4}$$

The implementation as an EWMA has some subtle differences compared to a neighbor average obtained from, e.g., a neighbor table. The EWMA favors neighbors with higher transmission rate over those with lower transmission rate and closer neighbors over more distanced ones [25]. We leave a comparison of different techniques for calculating the target rate for future work.



(a) Without target rate mechanism



(b) With target rate mechanism

Figure 6.10: Impact of the target rate mechanism on the convergence of AIMD

The target rate approximates the (locally) fair transmission rate in a node's surroundings. In PULSAR, we therefore use it as a "gravitation pull" for AIMD adjustments. The basic idea is to prevent nodes from moving away too far from the neighbor average. The target rate mechanism [25] doubles or halves a node's rate increment or decrement depending on whether the node's current transmission rate is above or below neighbor average. For example, when increasing transmission rate, all nodes below neighbor average double their increment, while those above average halve their increment. The target rate mechanism is summarized in Algorithm 2 on the next page, where U_k and U_t denote the current CBR and CBR target, respectively, r_k and r_t denote the node's current and target transmission rates, respectively, and α_I and β_D correspond to the AIMD parameters.

Figure 6.10b illustrates the impact of the target rate mechanism on the convergence to fairness in the scenario considered before. We can see that with the target rate mechanism, the delayed node catches up faster with the others, i.e., the convergence

Algorithm 2: AIMD with target rate mechanism

Data: $U_k, U_t, r_k, r_t, \alpha_I, \beta_D$
Result: r_{k+1}
if $U_k < U_t$ **then**
  if $r_k < r_t$ **then**
    $r_{k+1} \leftarrow r_k + \alpha_I \cdot 2$
  else
    $r_{k+1} \leftarrow r_k + \alpha_I \cdot 0.5$
  end
else
  if $r_k < r_t$ **then**
    $r_{k+1} \leftarrow r_k - r_k \cdot \beta_D \cdot 0.5$
  else
    $r_{k+1} \leftarrow r_k - r_k \cdot \beta_D \cdot 2$
  end
end

to fairness has been accelerated. However, we also observe a decreased smoothness of the AIMD adaptation which results from the step transition of the rate adaptation at the target rate. The example illustrates that the target rate mechanism is still crude and requires further development. In future work, the impact of a different function to converge toward the target rate could be explored. Such a function could, e.g., have a sigmoidal shape whose center is at the target rate. We leave a further exploration of this topic for future work.

6.3.4 Adaptation of the Relative Transmission Rate

In Section 5.4 on page 123, we introduced a concept to translate a vehicle's transmission rate into a safety benefit [25]. In essence, the concept adapts a vehicle's minimum and maximum transmission rate based on its driving context. These adaptations are expected to occur in the order of seconds to minutes. Two vehicles are said to have the same safety benefit if they transmit at the same *relative transmission rate*, i.e., at the same fraction of their minimum-maximum transmission rate interval which can be calculated according to Equation (5.3) on page 124.

As further discussed in Section 5.4.3 on page 130, the AIMD mechanism provides a straightforward way to support the adaptation of the relative transmission rate. That is, instead of entering absolute values in the unit of Hertz, the same mechanisms can be applied to percentage values. For example, assume that a node has a minimum and maximum transmission rate of 1 and 10 Hz, respectively and that it is currently transmitting at 5.5 Hz, which corresponds to a relative transmission rate of 0.5. Assume further that the next step is an increment and that the additive increase parameter is $\alpha_{I,rel} = 0.05$. Instead of adding 0.05 Hz, however, we now add 5 *percent*, resulting

in a new relative rate of $r_{rel} = 0.55$. The new absolute transmission rate is thus $r = r_{rel} \cdot (r_{max} - r_{min}) + r_{min} = 5.95\,\text{Hz}$.

In the event that a node's transmission rate boundaries change, it adapts its absolute transmission rate based on the current relative transmission rate. From a safety point of view, this approach has the advantage that a higher transmission rate can be made available instantaneously if required by the safety application. From a channel point of view, it may result in a temporary channel congestion if more than one vehicle suddenly increases their maximum transmission rate. However, such situations can be quickly resolved by the AIMD mechanism which due to its multiplicative decrease part quickly recovers from a channel overload situation. Note, however, that if the combination of the minimum transmission rates of all nodes exceeds the capacity of the channel, additional mechanisms will be required for congestion control, e.g., transmit power control.

Figure 6.11 on the facing page illustrates an exemplary adaptation of the relative transmission rate in the previously considered scenario where all nodes share the same location. The available 200 nodes were split into two groups of 100 nodes each which experience different changes to their minimum and maximum transmission rates every 5 s. From bottom to top, the figure shows the measured CBR by both groups of vehicles, the absolute transmission rate including minimum and maximum rate boundaries and the relative transmission rate, which is the basis of the adaptation and thus equal for both groups. Both groups start with a minimum and maximum transmission rate of 1 Hz and 10 Hz, respectively. The system converges to a CBR of ~0.7, a relative transmission rate of ~0.65 and a resulting absolute transmission rate of ~6.5 Hz. At simulation time 5 s, group 2 increases its minimum transmission rate to 8 Hz. As a result, its absolute transmission rate is immediately set to ~9.1 Hz, which corresponds to 65% of its new transmission rate interval. The resulting channel overload is compensated by PULSAR by converging to a relative transmission rate of ~0.4. At simulation time 10 s, group 1 sets its maximum rate to 3 Hz, allowing the relative transmission rate to go up toward 100%. Then, at simulation time 15 s, group 1 sets its transmission rate limits to 8 Hz and 10 Hz, i.e., to the same level as group 2. The result is that even though the relative transmission rate drops to 0, the channel is still congested. The example illustrates that a careful choice of the transmission rate boundaries is important when applying the relative transmission rate adaptation mechanism. Finally, at simulation time 20 s, both groups reset their minimum and maximum transmission rates to 1 Hz and 10 Hz, respectively. We observe that the system returns to its original state.

6.4 Exchange and Aggregation of Channel Information

So far, we focused on the temporal aspects of transmission rate adaptation only. In this section, we add a spatial dimension to the problem and introduce PULSAR's

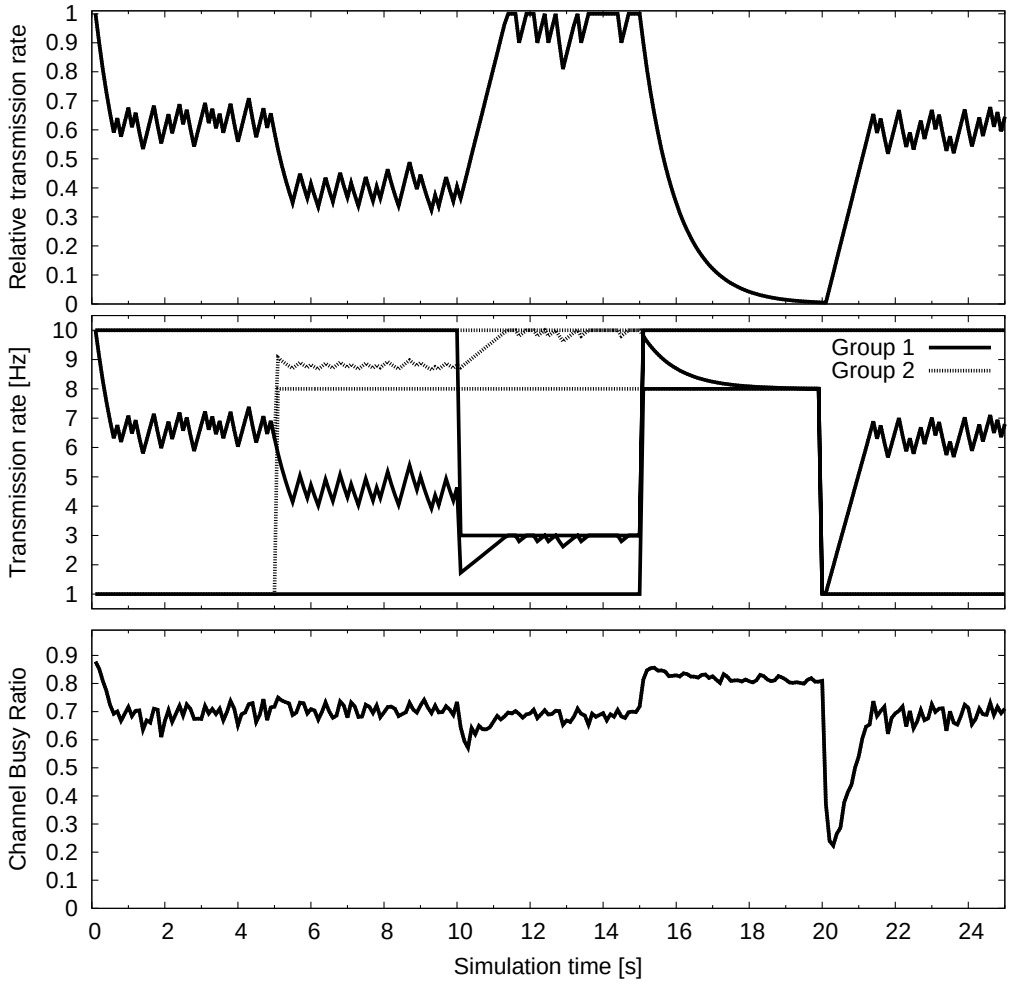


Figure 6.11: Exemplary adaptation of the relative transmission rate using AIMD with $\alpha_{I,rel}$ = 0.05 and $\beta_{D,rel}$ = 0.1

third building block, i.e., a mechanism to share CBR measurements among nodes. As we show in the following, the exchange of channel state information is closely related to fairness.

6.4.1 Motivation

To demonstrate the necessity of exchanging channel state information, we first study the convergence behavior of AIMD on a static linear road. The road has a length of 6 km, a fixed inter-vehicle distance of 10 m and a fixed transmit power of 20 dBm. We set the initial transmission rate to 10 Hz. Radio-wave propagation is calculated according to the power-law model with log-normal fading, cf. Appendix A.1 on page 223. In this setup, the maximum uniform transmission rate which does not violate the CBR threshold of 0.7, i.e., the globally fair transmission rate in the middle of the road, is approximately 5.3 Hz.

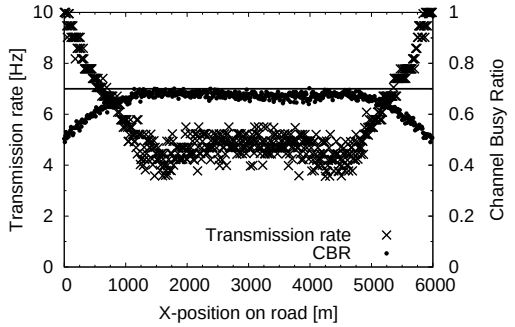
Figure 6.12 on the facing page illustrates four snapshots of the simulation at simulation times 4 s, 13 s, 30 s and 179 s, respectively. Each sub-figure shows the current transmission rate and measured CBR, quantified by the left and right y-axis, respectively, of each vehicle with respect to its position on the road. In Figure 6.12a, we can see that after 4 s, nodes between 1000 m and 5000 m have decreased their rates to between 4 Hz and 5 Hz, while transmission rates gradually increase towards both ends of the road where a lower channel load is measured.

After 13 s, nodes near both ends of the road have further increased their transmission rates, as shown in Figure 6.12b. As a result, two "valleys" have formed in the transmission rate allocation, i.e., near 1300 m and 4700 m. After 30 s, a third "valley" has formed near 3000 m, cf. Figure 6.12c. The steady-state transmission rate allocation is shown in Figure 6.12d, where we observe a spatial wave in both transmission rate allocation and CBR measurements. Note that the formation of the spatial wave is not caused by the boundary effects in the scenario alone, since it can also be observed on a circular road which eliminates boundary effects. On the circular road, the wave is static as well, i.e., it does not move over time. Its wavelength depends on the chosen transmit power, while its phase, i.e., the location of its crests and troughs, appears to be random.

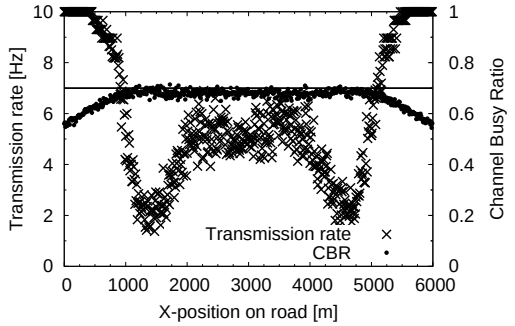
Similar outcomes, i.e., a spatial drift of the transmission rates of neighboring nodes, can be observed in other scenarios as well. From a safety point of view, such a behavior is undesirable, since it may cause some vehicles to not be able to communicate their current status sufficiently. It should be noted that the target rate mechanism introduced above is not sufficient to prevent the spatial drift of nodes' transmission rates. Analogously to Figure 6.12, Figure 6.13 on the next page illustrates the convergence behavior of AIMD with the target rate mechanism. We can see that the target rate mechanism significantly improves local fairness and mitigates global unfairness. However, it does not influence nodes' decisions of increasing or decreasing their transmission rates and thus cannot prevent the formation of the two "valleys" near 1300 m and 4700 m. To influence the *direction* of nodes' transmission rate adjustments and thus to facilitate global fairness, an exchange of CBR information is required, which we study in the following.

6.4.2 One-Hop Piggybacking

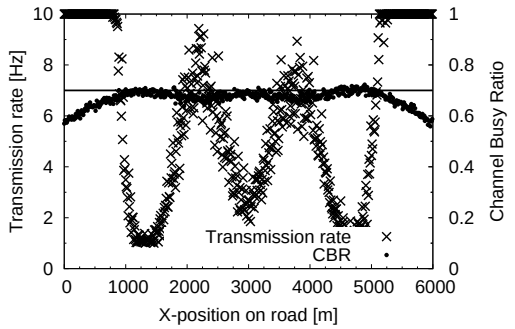
In wireless communications, nodes may cause interference beyond their own detection range. A node experiencing a congested channel thus needs to inform its neighbors of their contribution to the overload situation. In this thesis, we consider two alternatives to disseminate congestion information. First, a congested node could transmit a dedicated congestion message. Second, it could piggyback congestion information in the (periodic) messages it transmits anyway.



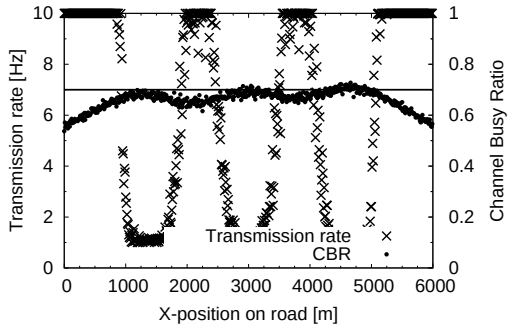
(a) Simulation time 4 s



(b) Simulation time 13 s

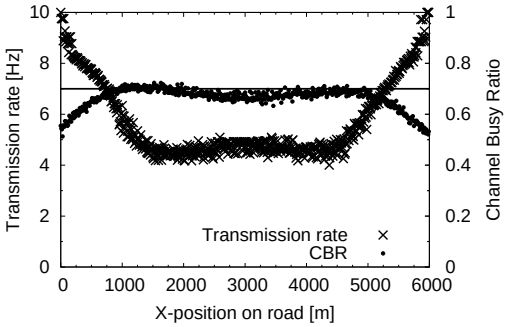


(c) Simulation time 30 s

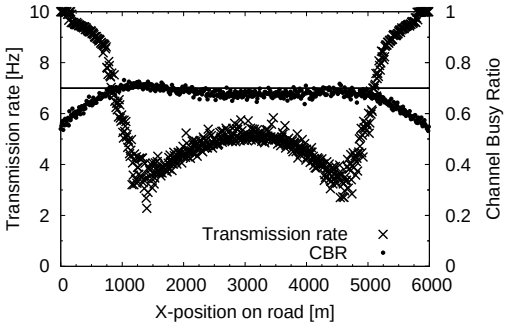


(d) Simulation time 179 s

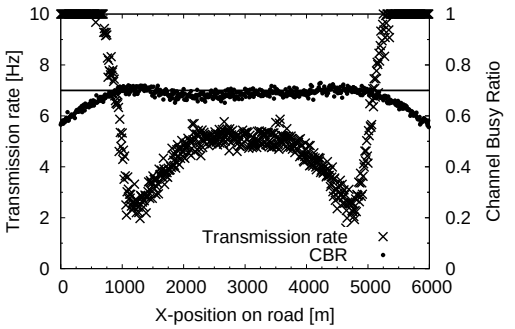
Figure 6.12: AIMD-based adaptation on a linear road (local information, without target rate)



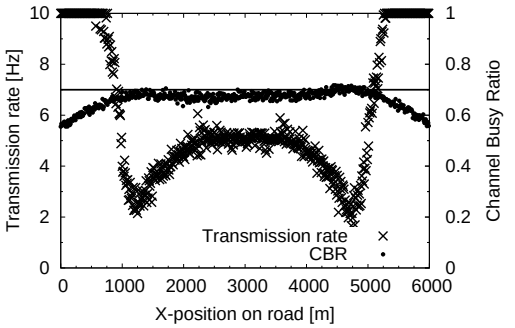
(a) Simulation time 4 s



(b) Simulation time 13 s



(c) Simulation time 30 s



(d) Simulation time 179 s

Figure 6.13: AIMD-based adaptation on a linear road (local information, with target rate)

Piggybacking Versus Dedicated Congestion Messages

A dedicated congestion message has the advantage of not having to wait until a new regular message has been generated. In addition, it could be transmitted at a different transmit power than regular messages and could thus reach beyond the node's typical communication range. However, each additional message comes at the cost of increased overhead, which would be counterproductive especially if the channel is already in an overload situation. The message itself would likely be very small, i.e., in the order of a few bytes, but would require the inclusion of all headers of the protocol stack, cf. Figure 2.5 on page 23. In addition, the message would have to go through the entire Carrier Sense Multiple Access (CSMA) backoff procedure, cf. Section 2.1.2 on page 13. To keep overhead low, we thus focus on piggybacking in this thesis.

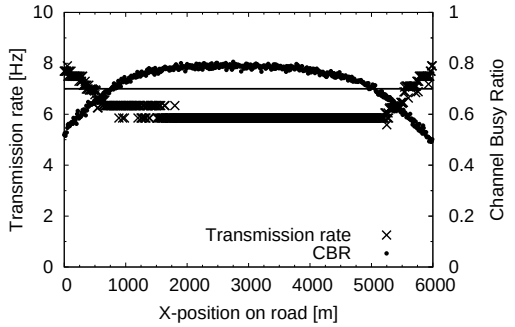
Aggregation of CBR Measurements

When a node receives piggybacked CBR measurements from its neighbors, it needs to aggregate this information before taking a rate control decision. In the following, we consider two alternatives for aggregation, i.e., the 1-norm and the maximum norm. The former corresponds to the arithmetic mean of the CBR values received from each neighbor, while the latter corresponds to the maximum of the received CBR values. The maximum norm is the more conservative approach, since it causes a node to reduce its transmission rate if any one of its neighbors reports a congested channel. By favoring the most congested nodes in the scenario, it helps to maximize the minimum transmission rate, i.e., to achieve global fairness.

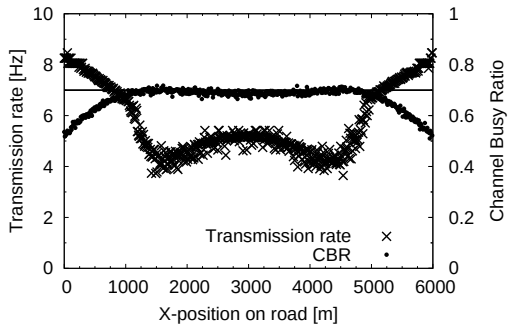
Figures 6.14 and 6.15 on the facing page illustrate the impact of using the 1-norm and maximum norm, respectively, in the aggregation of piggybacked CBR measurements from one-hop neighbors in the scenario considered above[5]. In both cases, the target rate mechanism has been applied. We can see that the 1-norm does not help to prevent the formation of the spatial wave observed without information sharing. In fact, it reinforces the wave's formation, as we can see by comparing Figures 6.13d on the previous page and 6.14d on the facing page. In contrast, the maximum norm successfully prevents the formation of the spatial wave as we can see in Figure 6.15d on the next page. In steady state, the transmission rates of the nodes located between 1500 m and 4500 m stay near 4.5 Hz and oscillate according to the selected AIMD parameters.

Note that while the maximum norm enables the protocol to converge towards global fairness, it comes at the price of an increased vulnerability to erroneous CBR measurements. Such measurements might originate from malfunctioning radio devices as well as from intentionally falsified data. In practice, it may thus be necessary to introduce *sanity checks* for received CBR information. For example, a receiving vehicle could check if it has received congestion indications from the reporting vehicle's neighbors as well (*spatial validation*). Similarly, the receiver could compare the new

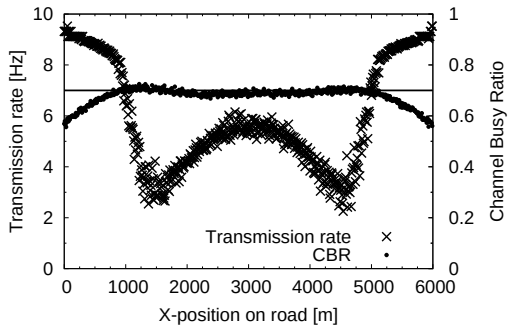
[5] Each node piggybacks its last CBR measurement which was taken at the last (synchronized) transmission rate adaptation event.



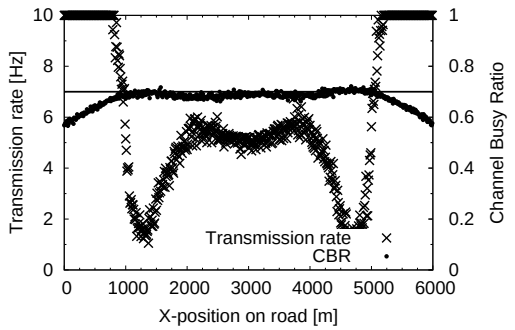
(a) Simulation time 4 s



(b) Simulation time 13 s

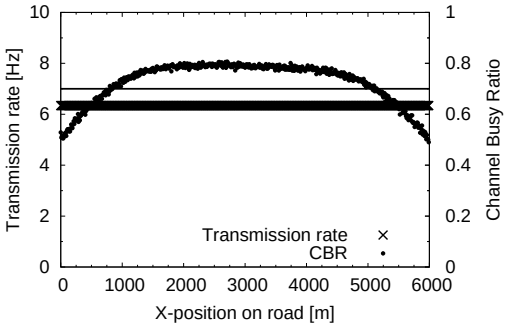


(c) Simulation time 30 s

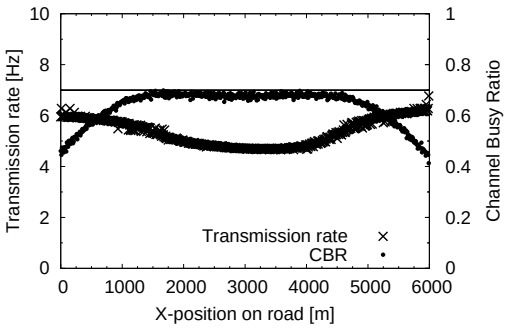


(d) Simulation time 179 s

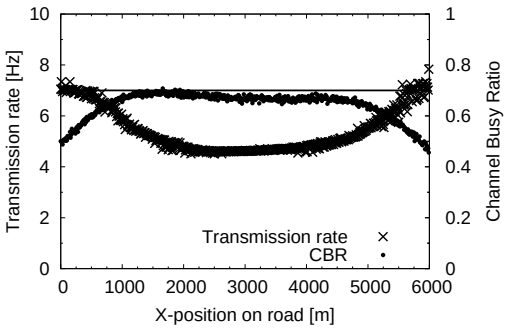
Figure 6.14: AIMD-based adaptation on a linear road (one-hop information, 1-norm, with target rate)



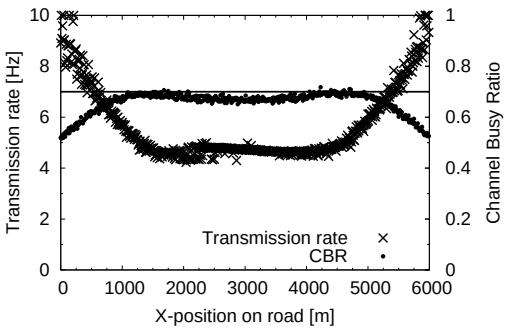
(a) Simulation time 4 s



(b) Simulation time 13 s



(c) Simulation time 30 s



(d) Simulation time 179 s

Figure 6.15: AIMD-based adaptation on a linear road (one-hop information, maximum norm, with target rate)

information against previously received measurements from the reporting vehicle (*temporal validation*). We leave a further exploration of this topic for future work.

6.4.3 Piggybacking Efficiency

So far, we focused on a linear road scenario in which the amount of interference at any location is limited due to the one-dimensionality of the setup. In the following, we thus use a crossing of two linear roads in order to create a single location with a high amount of interference. In the center of the intersection, interference aggregates from all four directions. It should be noted here that the studied scenario is intended to resemble two highways crossing over a bridge, cf. Figure 3.9 on page 72, rather than a city scenario in which buildings can obstruct the line of sight.

Performance of One-Hop Piggybacking in a Crossing Scenario

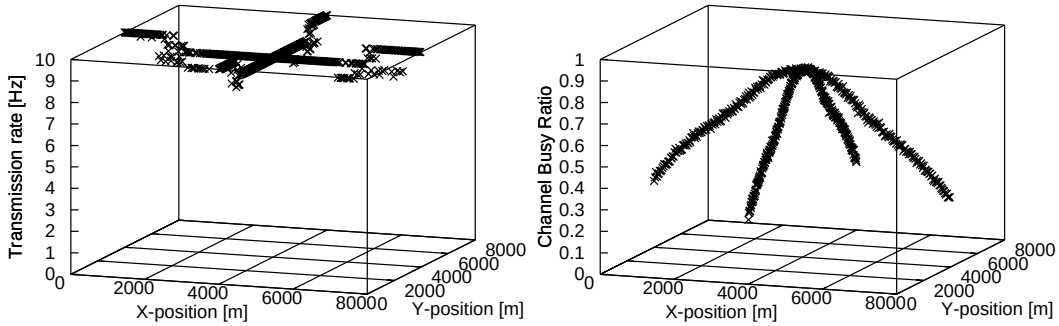
Figure 6.16 on the next page illustrates four simulation snapshots of AIMD with target rate and one-hop piggybacking of CBR measurements. Each sub-figure consists of two 3D plots whose z-axis shows the current transmission rate and measured CBR, respectively. The x-axis and y-axis of each plot represent the reporting node's location in the 2D scenario. The considered crossing consists of two linear roads, each with a length of 8 km and an inter-vehicle distance of 10 m. To better illustrate the convergence behavior, vehicles are placed statically, i.e., they do not move. Like before, the globally fair transmission rate in the center of the scenario is approximately 5.3 Hz, which is reached if all vehicles use this transmission rate.

Following AIMD's convergence from simulation time 1 s to 180 s in Figures 6.16a to 6.16d, we can see that the nodes at the intersection reduce their transmission rates up to close to the minimum transmission rate of 1 Hz[6], while those beyond approximately 1300 m from the intersection center transmit at 10 Hz, i.e., the maximum transmission rate. In other words, the global fairness principle is not fulfilled in steady state.

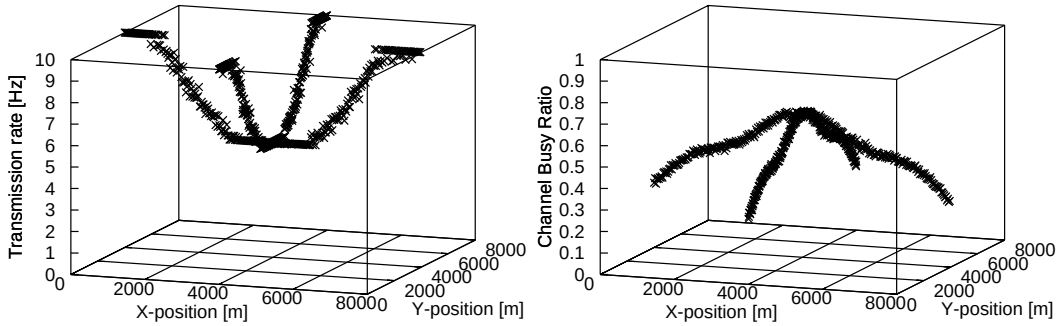
Probability of Receiving Congestion Information

To see why nodes at the intersection center get pushed down almost to their minimum transmission rate, consider Figure 6.17 on page 160. The figure illustrates the aggregated probability of receiving one out of n messages at a CBR of 0.7 with respect to sender-receiver distance, cf. Figure 4 in [25]. The aggregated probabilities were calculated based on the observed probability of receiving an individual message analogously to the IRT concept, cf. Section 4.1.2 on page 86. The figure further shows the carrier sensing probability at the configured transmit power of 20 dBm and carrier sense threshold of -95 dBm. The vertical lines correspond to the theoretical transmis-

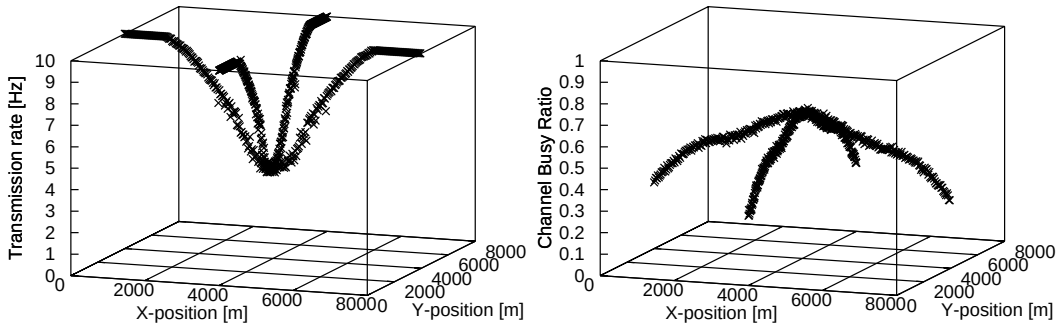
[6] Without the target rate mechanism, nodes at the center of the intersection converge to the minimum transmission rate.



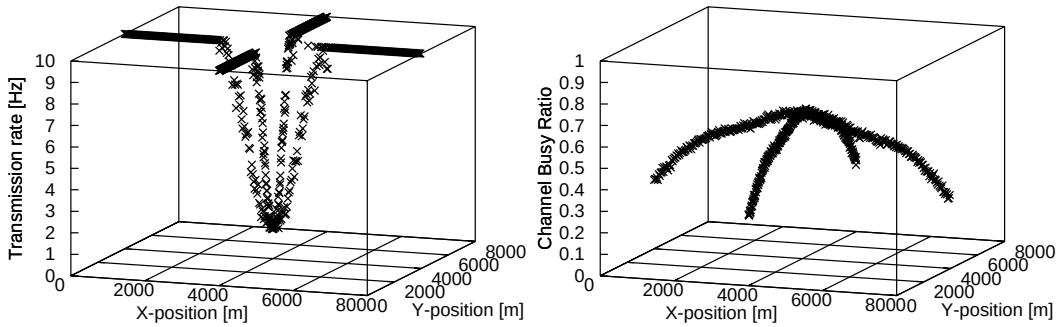
(a) Simulation time 1 s



(b) Simulation time 4 s



(c) Simulation time 13 s



(d) Simulation time 180 s

Figure 6.16: Convergence of AIMD with target rate using one-hop piggybacking in the crossing scenario

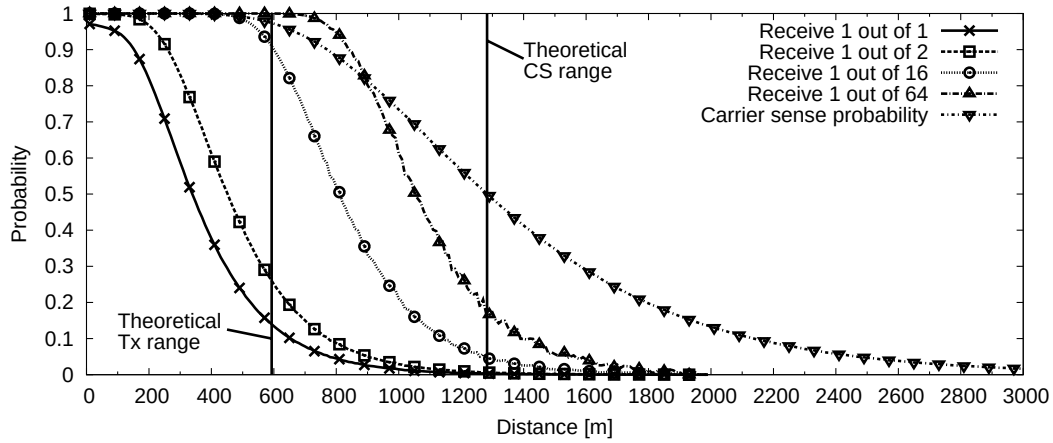


Figure 6.17: Piggybacking efficiency (probability of receiving one out of n messages) at transmit power 20 dBm, CBR 0.7

sion range and carrier sense range, respectively, i.e., without fading and interference. The basic assumption in the following line of reasoning is that if channel congestion occurs at a certain location, more than one vehicle will measure and report a CBR above the configured threshold, e.g., 0.7. From a remote interferer's point of view, it is then sufficient to receive one out of these n messages because of the applied maximum norm in the aggregation process.

Coming back to Figure 6.17, we can see that an interferer located at the theoretical carrier sense range of 1283 m still has a 50% chance of triggering another node into channel busy state[7]. On the other hand, the probability that this interferer receives one out of 16 messages congestion information is below 5%. When the number of messages is increased to 64, the probability is still below 20%. It is thus significantly more likely for a distanced interferer to contribute to congestion than to receive congestion information. In other words, the 1-hop piggybacking protocol does not cover the required *participation distance* sufficiently, cf. Section 5.3.3 on page 122. In the crossing scenario considered above, nodes located at the center of the intersection thus cannot disseminate their channel state information far enough and as a result reduce their transmission rates further and further.

6.4.4 Two-Hop Piggybacking

In order to extend the dissemination range of congestion information, PULSAR includes a two-hop piggybacking mechanism for CBR information [25]. In a nutshell, each node not only piggybacks its own CBR measurement, but also the highest

[7] Among other factors, this high probability results from the applied one-slope path loss model. Some authors have suggested the use of dual slope models, e.g., [218], which would result in less interference at far distances. However, it is out of the scope of this thesis to determine which approach is more realistic.

CBR information received from its neighbors within a pre-configured *hop distance* d_{hop}. A receiver can then add d_{hop} to its distance to the sender and thus infer the distance from within which the two-hop CBR information originated. If the sum exceeds the node's *participation distance* d_{part}, cf. Section 5.3.3 on page 122, the two-hop information is ignored. This way, the choice of d_{hop} can be used to limit the spatial impact of congestion control information if desired. Note that the two-hop piggybacking approach has similarities with the one applied in SOURC [202]. We will get back to this topic in Section 7.4.2 on page 206, when we compare PULSAR's performance against SOURC's.

Analogously to the limitation in the spatial dimension in the form of d_{hop}, the two-hop piggybacking mechanism uses a maximum time window t_{max} to limit the age of the CBR information taken into account in the rate control process. Due to the synchronization required for AIMD, cf. Section 6.3.2 on page 144, we simplify the temporal dimension by sharing CBR measurements taken at discrete points in time which are multiples of the control interval t_{adp}. This way, we ensure that all nodes react to the same state of the channel, albeit with a reaction delay. Note that for other increase/decrease algorithms, it may be an option to piggyback measurements taken at the time the respective message was generated. We leave such considerations for future work.

In the following, we introduce a simplified version of the two-hop piggybacking approach presented in [25]. While in [25], we applied $t_{max} = 2t_{adp}$, it turned out later that this approach does not seem to provide significant advantages over an application of $t_{max} = t_{adp}$ with a higher value of t_{adp}. We thus focus on the simple version here and refer the reader to [25] for the description of the originally published version.

Figure 6.18 on the following page illustrates the basic functionality of the two-hop piggybacking approach. The figure shows three vehicles A, B and C with an inter-vehicle distance of d_{hop}, depicted in the horizontal dimension. The transmission and carrier sense ranges in the example are equal to d_{hop} and $2 \cdot d_{hop}$, respectively. In other words, C contributes to congestion at A, but A cannot communicate with C directly. The vertical dimension of the plot corresponds to the vehicles' actions in temporal order. Three symbols, i.e., circle, square an triangle, represent the CBR measurements taken by each vehicle at three successive control intervals t_1, t_2 and t_3, respectively. The arrows between the vehicles correspond to transmissions with piggybacked information. Each packet has two header fields, i.e., the local CBR measurement at t_i ("Local (t_i)") and the maximum received CBR measurement from neighbors within d_{hop} and generated at t_i ("1Hop (t_i)"). Not explicitly shown in the figure is a timestamp which is also attached to each packet in order to facilitate mapping the contained information to the corresponding control interval. The right column of the plot beyond the dashed line illustrates how vehicle C aggregates the received information before adapting its transmission rate.

In our example, the dissemination process starts at t_1, when all vehicles perform a CBR measurement, represented by a circle. At $t_1 + 0.3 \cdot t_{adp}$, vehicle A broadcasts a message

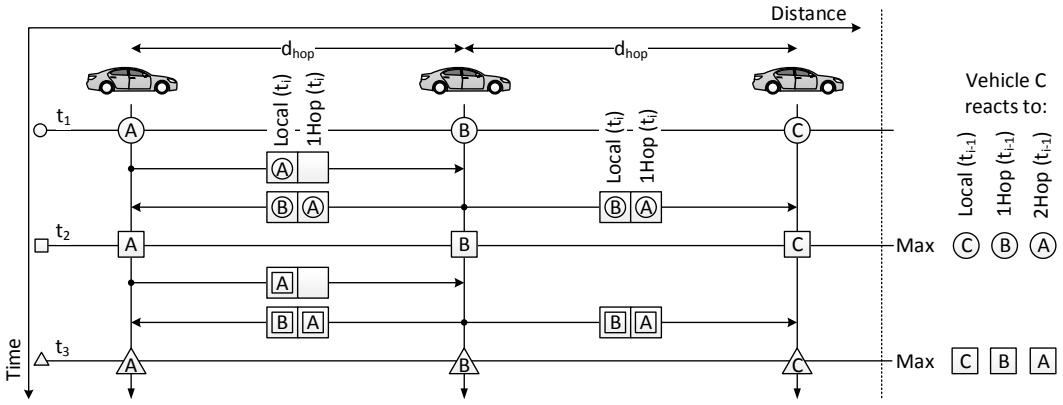


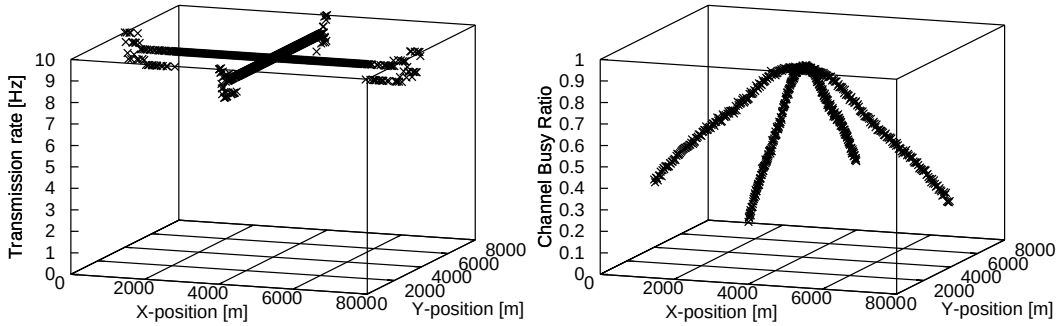
Figure 6.18: Illustration of two-hop piggybacking and CBR aggregation

and attaches its last CBR measurement. The one-hop field remains empty, since it has not received any CBR information from its neighbors within this adaptation period. At $t_1 + 0.6 \cdot t_{adp}$, vehicle B broadcasts a message and attaches its own measurement from t_1 as well as vehicle A's measurement from t_1, since it is the maximum value B has received from its neighbors within this period. Vehicle C receives B's message and is thus informed of the latest CBR measurements of both A and B. Note, however, that B's message is also received by A, which essentially includes a reflection of its own measurement. Since the origin of the 1-hop information is not included in the packet, A cannot detect this *bounce back effect*. To avoid a situation in which A would react twice to its own measurement while B and C only react once, we introduce a *reaction delay* which is illustrated using the example of node C. As shown in the right column of the figure, at t_2 vehicle C reacts to the maximum of its own measurement from t_1, the maximum received one-hop CBR, i.e., B's measurement, and the maximum received two-hop CBR, i.e., A's measurement. The same process is repeated at t_3, when C reacts to the measurements taken at t_2, represented by squares.

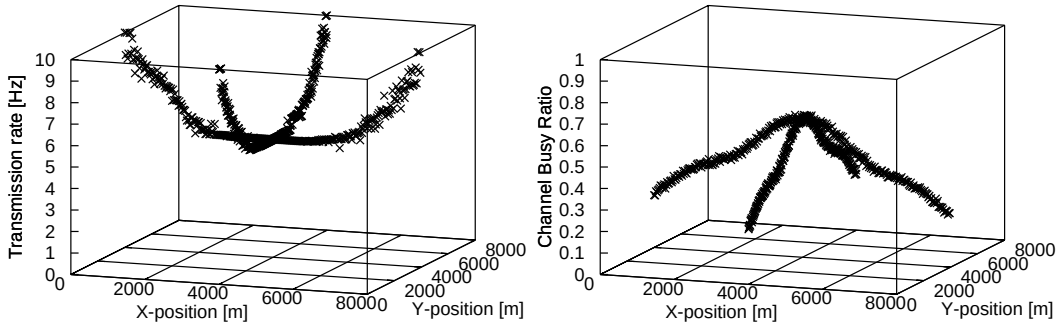
Note that the reaction delay is primarily introduced because of the binary nature of AIMD. For example, if A's measurement from t_1 were the only one indicating congestion among the measurements of A, B and C within t_1 and t_2, A would reduce its transmission rate twice, while B and C would only do so once. This would be unfair, since both B and C contribute to the channel congestion observed at A.

Convergence in the Crossing Scenario

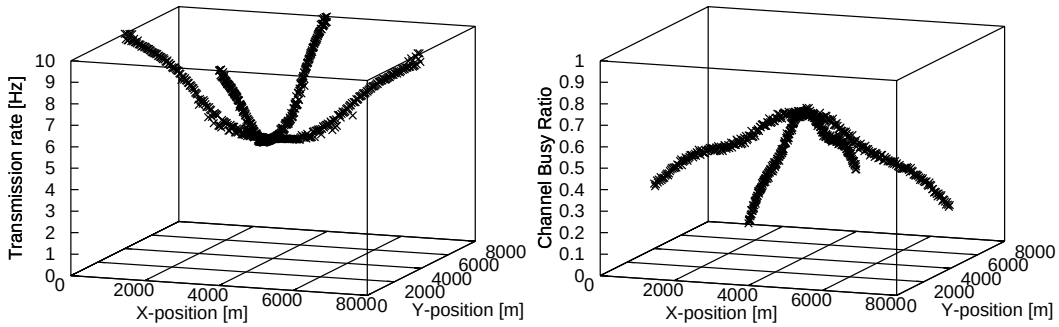
The impact of the two-hop piggybacking mechanism on the convergence of AIMD in the crossing scenario considered above is depicted in Figure 6.19 on the next page. We observe a significant extension of the distance at which nodes participate in congestion control. That is, the distance at which nodes transmit at the maximum rate of 10 Hz has been increased to 2000 m from the intersection center. The result is that in steady state, the transmission rate of the nodes near the intersection center stays near 4.5 Hz, cf. Figure 6.19d.



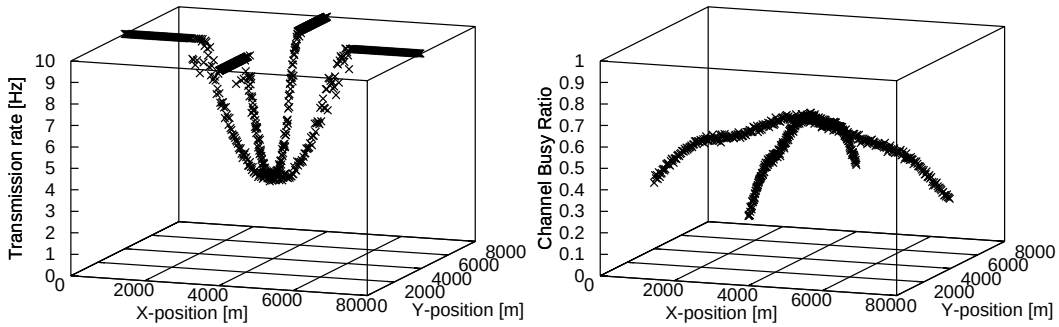
(a) Simulation time 1 s



(b) Simulation time 4 s



(c) Simulation time 13 s



(d) Simulation time 180 s

Figure 6.19: Convergence of AIMD with target rate using two-hop piggybacking in the crossing scenario

Overhead and Header Format

So far, we described the two-hop piggybacking mechanism based on the exchange of CBR measurements. Assuming a discretization in steps of 0.005, this approach requires 8 bit to encode each CBR measurement. In total, we thus need two bytes to exchange CBR information. However, due to the binary nature of AIMD, this overhead can be reduced significantly by piggybacking a *congestion bit* rather than the entire CBR measurement. That is, if the node's CBR measurement exceeds the CBR threshold, the bit is set to one. Else, it is set to zero. The resulting one-hop information to be forwarded is reduced to a single bit as well, while the aggregation and forwarding process does not need to be changed. Using this approach, the overhead can thus be reduced from 2 bytes to 2 bit.

Note that PULSAR requires two more header fields, i.e., a time stamp and the vehicle's current transmission rate. The former is used to map received information to control intervals, while the latter is required to calculate the target rate. Table 6.1 on the facing page summarizes PULSAR's header format.

6.5 Summary

In this chapter, we introduced PULSAR as a concrete congestion control protocol for VSC which meets the design principles formulated in Section 5.5 on page 131. Table 6.2 on the facing page provides an overview of which protocol mechanism is most important for the fulfillment of each design principle.

A key characteristic of the PULSAR protocol is its modular structure. In the course of this chapter, we discussed different options for each of its three building blocks, i.e., local channel load assessment, transmission rate adaptation and information sharing. It is interesting to note that PULSAR's structure essentially reflects the six basic components of a congestion avoidance scheme formulated by Jain and Ramakrishnan in [115], cf. Section 3.1 on page 46. Table 6.3 on the facing page illustrates how each of the six components can be matched to PULSAR's protocol mechanisms.

Field	Size [bit]	Meaning
Timestamp	8	Generation time of the message
TxRate	8	Transmission rate when generating the message
Local (t_i)	1 *or* 8	Local congestion state *or* CBR measurement generated at the last (synchronized) adaptation event t_i
1Hop (t_i)	1 *or* 8	Maximum congestion state *or* CBR measurement received from neighbors within 1-hop range and generated at t_i

Table 6.1: PULSAR's header format

Design principle	Protocol mechanism
Decentralization	Distributed algorithm
Local fairness	Target rate mechanism
Global fairness	Aggregation of CBR values (maximum norm)
Participation fairness	Two-hop piggybacking
Deference to safety applications	Adaptation of relative transmission rate

Table 6.2: Realization of design principles in PULSAR

Component	Protocol mechanism
Congestion detection	CBR measurements
Feedback filter	Self averaging
Feedback selector	Participation distance in information sharing
Signal filter	Aggregation of CBR values (maximum norm)
Decision function	Comparison of aggregated CBR against target
Increase/ decrease algorithm	AIMD with target rate mechanism

Table 6.3: PULSAR as a congestion avoidance scheme according to Jain et al. [115]

7

Evaluation and Discussion

In the first part of this chapter, we evaluate PULSAR's performance in terms of convergence, fairness and stability. We thereby increase the level of complexity of the considered scenarios from a zero-dimensional static topology in which all nodes share the same location to a mobile highway scenario with a high dynamic in vehicle density. In the second part, we evaluate the adaptation of the relative transmission rate. Our main objective is to show how PULSAR reacts to reasonably realistic changes in target distance and transmission rate boundaries and that the relative transmission rate concept can make a difference from the perspective of a safety application compared to a conventional congestion control which only focuses on channel load. In the last part of the evaluation, we compare PULSAR's performance against three other state of the art approaches, two of which appear to be very similar to PULSAR at first glance. We conclude this chapter by discussing two potential limitations of this work.

7.1 Responsiveness and Smoothness

As discussed in Section 3.1.3 on page 49, the convergence behavior of a linear control scheme can be characterized by its *responsiveness*, i.e., the time it takes to converge to efficiency, and its *smoothness*, i.e., the amplitude of the oscillations [120]. In PULSAR, smoothness and responsiveness are determined by the choice of the Additive Increase Multiplicative Decrease (AIMD) parameters on the one hand and the applications of additional mechanisms on the other hand, e.g., target rate and two-hop piggybacking. In the following, we study and quantify the impact of each of these factors.

7.1.1 AIMD Parameter Choice

In [120], Chiu and Jain calculate the responsiveness and smoothness of a linear control scheme according to Equations (3.5) and (3.6) on page 51, respectively. In the following, we use these results to quantify the responsiveness and smoothness of AIMD in the context of Vehicle Safety Communications (VSC).

Responsiveness When Converging From Underutilization

Due to the binary nature of AIMD, its responsiveness depends on the initial channel load U_o compared to the targeted channel load U_t. If $U_o < U_t$, the multiplicative component is zero and Equation (3.5) becomes

$$t_e^- = \frac{U_t - U_o}{\alpha_I \cdot n} \tag{7.1}$$

where U_o and U_t denote the initial and targeted channel load, respectively, α_I is the additive increase parameter and n is the number of nodes. Since in PULSAR, α_I is specified in the unit of Hertz, it needs to be converted to Channel Busy Ratio (CBR) by multiplying it with the packet duration, i.e., 584 μs in our setup, cf. Section 5.2 on page 111. For example, if $\alpha_I = 0.05$ Hz, $n = 200$, $U_o = 0.11$ and $U_t = 0.7$, we get $t_e^- = \lceil (0.7 - 0.11)/(0.05 \cdot 0.000584 \cdot 200) \rceil = 101$ adaptation steps. Analogously, we get 6 steps if $\alpha_I = 1$ Hz.

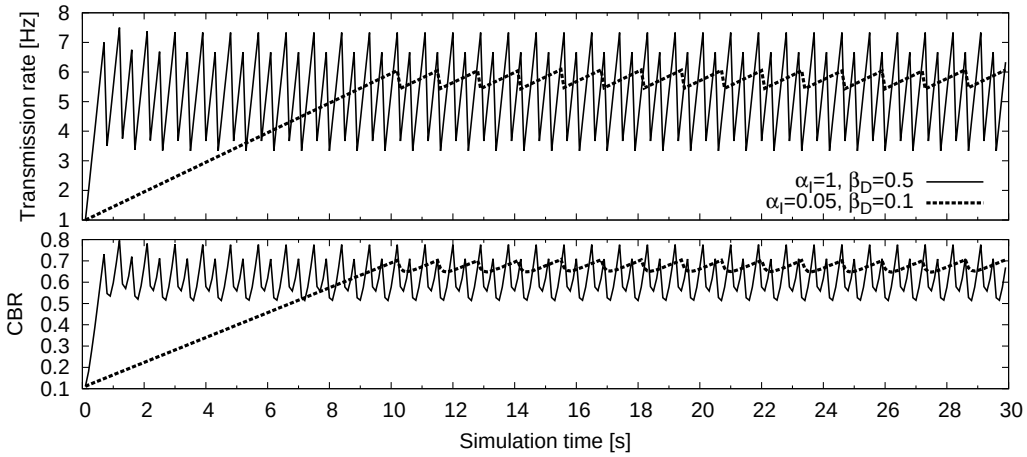
Figure 7.1a on the facing page illustrates the convergence of AIMD in the 200 node scenario considered before where all nodes share the same location. To minimize random effects, the initial transmissions of all nodes are equally spaced and no random jitter is applied. The figure shows the resulting transmission rate and (self-averaged) CBR for two AIMD configurations with $\alpha_I = 0.05$ Hz, $\beta_D = 0.1$ and $\alpha_I = 1$ Hz, $\beta_D = 0.5$. As illustrated by the figure, the target CBR of 0.7 is first crossed at simulation time 10.2 s for the former configuration and 0.7 s for the latter configuration, which corresponds 101 and 6 completed adaptation intervals of 100 ms, respectively. Thus, the observed values match the theoretically calculated ones.

Responsiveness When Converging From Overutilization

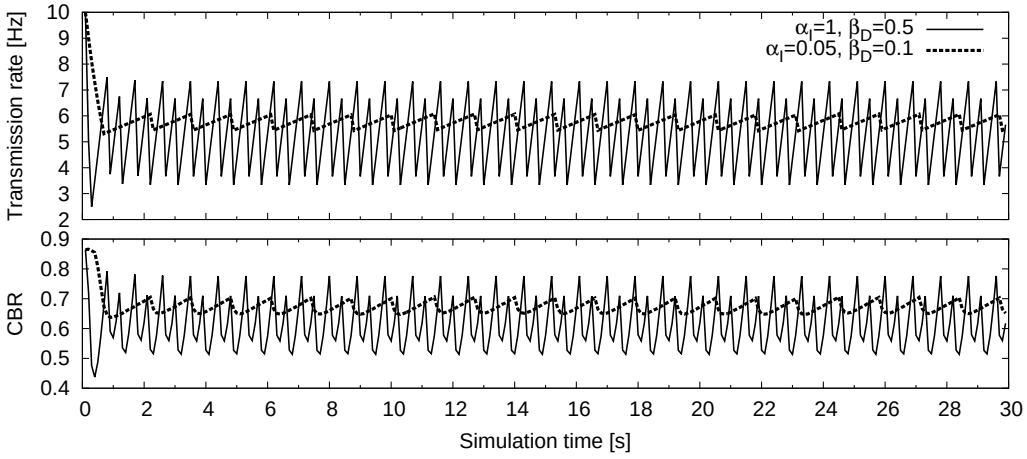
When starting from a channel overload situation, i.e., if $U_o > U_t$, the additive component is zero and Equation (3.5) becomes

$$t_e^+ = \frac{log(U_t/U_o)}{log(1 - \beta_D)} \tag{7.2}$$

where $1 - \beta_D$ is the multiplicative parameter of AIMD. When applying this equation, however, it has to be taken into account that the relationship between communication density and CBR becomes nonlinear beyond a CBR of approximately 0.7, cf. Section 5.2 on page 111. For example, if $U_o = 0.875$, $U_t = 0.7$ and $\beta_D = 0.1$, Equation (7.2) yields



(a) Starting from 1 Hz



(b) Starting from 10 Hz

Figure 7.1: Convergence of AIMD for two different parameter choices

$t_e{}^+ = \lceil log(0.8)/log(0.9) \rceil = 3$ adaptation steps as a result. However, as illustrated by Figure 7.1b the target CBR of 0.7 is first crossed at simulation time 0.7 s, i.e., after 6 adaptation steps. In fact, the CBR stays above 0.85 up to simulation time 0.4 s, i.e., for 3 adaptation steps. For $\beta_D = 0.5$, the threshold of 0.7 CBR is first crossed at simulation time 0.3 s, i.e., after 2 adaptation steps, while the theoretically calculated value is 1 step.

Smoothness: Maximum Overshoot

In [120], Chiu and Jain quantify smoothness in terms of the maximum overshoot, cf. Equation (3.6) on page 51. For AIMD, the maximum overshoot occurs when all nodes perform an increment operation starting from just below U_t. The maximum overshoot results as

$$s_e{}^+ = \alpha_I \cdot n \tag{7.3}$$

Normalizing α_I by 584 μs and setting $n = 200$ nodes, we get $s_e{}^+ = 0.00584$ CBR and $s_e{}^+ = 0.1158$ CBR for $\alpha_I = 0.05$ Hz and $\alpha_I = 1$ Hz, respectively. In Figure 7.1 on the preceding page, the maximum observed overshoot is 0.0061 and 0.077, respectively. Note that the shown CBR values are time-averaged using the self-averaging concept introduced in Section 6.2.1 on page 137. The maximum 100 ms Simple Moving Average (SMA) values observed before averaging match the theoretically computed maximum overshoot but are not shown in the figure for presentation clarity.

Smoothness: Maximum Undershoot

The maximum undershoot in AIMD happens when all nodes decrease their transmission rates from just above the CBR threshold. That is, it can be calculated as

$$s_e{}^- = \beta_D \cdot U_t \tag{7.4}$$

For $\beta_D = 0.1$ and $\beta_D = 0.5$ we get $s_e{}^- = 0.07$ CBR and $s_e{}^- = 0.35$ CBR, respectively. In Figure 7.1, the observed maximum undershoot in steady state is 0.053 CBR and 0.186 CBR, respectively. Before self-averaging, the minimum observed 100 ms SMA values match the theoretically computed maximum undershoot.

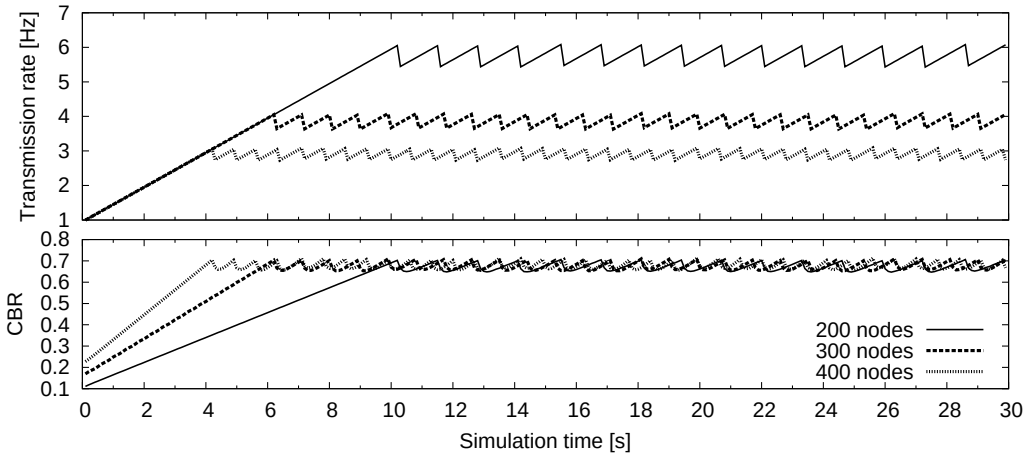
7.1.2 Other Influencing Factors

PULSAR's responsiveness and smoothness is influenced by a number of additional factors which need to be taken into account when selecting AIMD parameters. In the following, we discuss the influence of the number of nodes, the adaptation interval as well as of the piggybacking and target rate mechanisms.

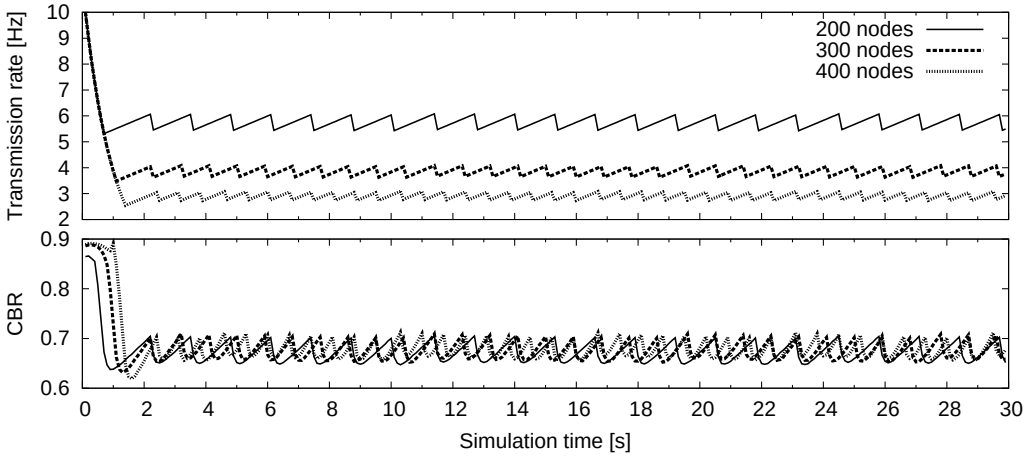
Number of Nodes

In Equation (7.3) on the preceding page, we saw that the maximum overshoot of AIMD depends on the number of nodes. That is, if the additive increase parameter is set too high, the combined increment of all nodes may cause large oscillations in the observed CBR up to total channel congestion. For example, if $\alpha_I = 1$ Hz and $n = 1000$ nodes, a synchronized increment operation by all nodes would result in a total increase of $584\mu s \cdot 1\text{Hz} \cdot 1000 = 0.584$ CBR, which corresponds to more than 50% of the channel capacity in one step. In comparison, $\alpha_I = 0.05$ Hz results in a total increase of 0.029 CBR for $n = 1000$. The maximum number of nodes therefore needs to be taken into account when selecting the additive increase parameter.

Figure 7.2 on the next page illustrates the impact of the number of nodes for $\alpha_I = 0.05$ Hz and $\beta_D = 0.1$. In the example, the smoothness of AIMD is dominated by the multiplicative parameter, i.e., $s_e{}^- > s_e{}^+$. Thus, the observed smoothness in terms of CBR does not depend on the number of nodes. However, the oscillations in the transmission rate are higher for a smaller number of nodes, since AIMD converges



(a) Starting from 1 Hz



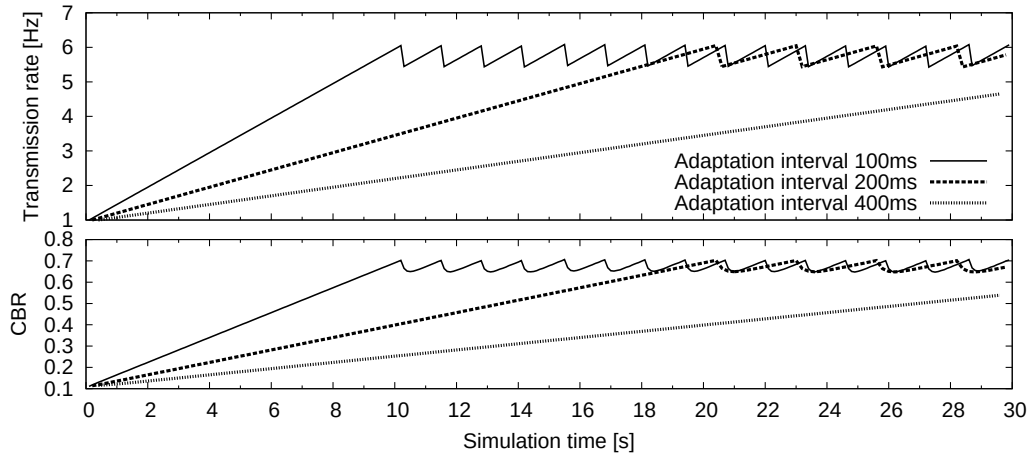
(b) Starting from 10 Hz

Figure 7.2: Impact of the number of nodes on responsiveness and smoothness

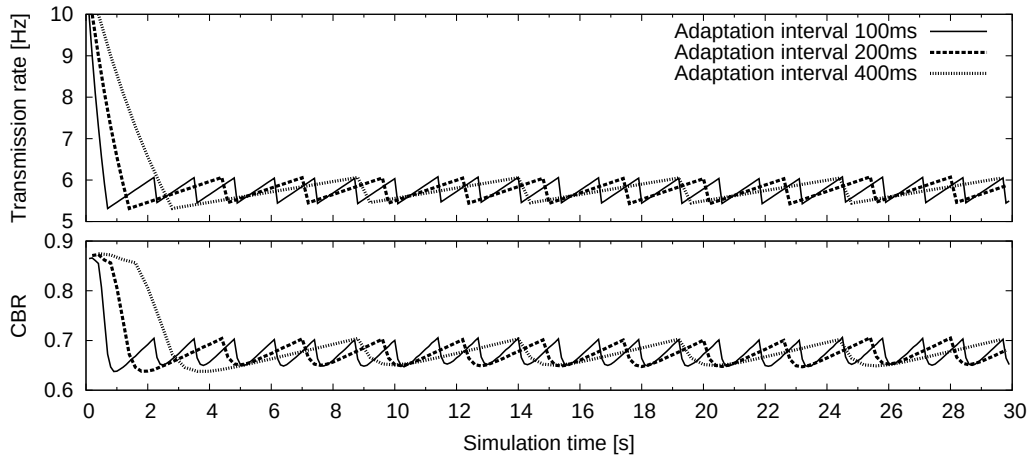
to a higher transmission rate. We can further observe that the convergence time of AIMD decreases with the number of nodes, since the accumulated rate increment of all nodes fills up the channel faster. This is also reflected by a lower transmission rate at which the algorithm converges.

Adaptation Interval

The adaptation interval t_{adp}, i.e., the period between two transmission rate adaptation events, influences PULSAR's responsiveness. For example, the previously calculated 101 adaptation steps required by AIMD with $\alpha_I = 0.05$ Hz, $\beta_D = 0.1$ to converge from 0.1 CBR to 0.7 CBR results in $t_e^- = 10.2$ s for $t_{adp} = 100$ ms, $t_e^- = 20.4$ s for $t_{adp} = 200$ ms, and so on. The adaptation process is depicted in Figure 7.3 on the following page. We can also observe that the adaptation interval does not influence smoothness.



(a) Starting from 1 Hz

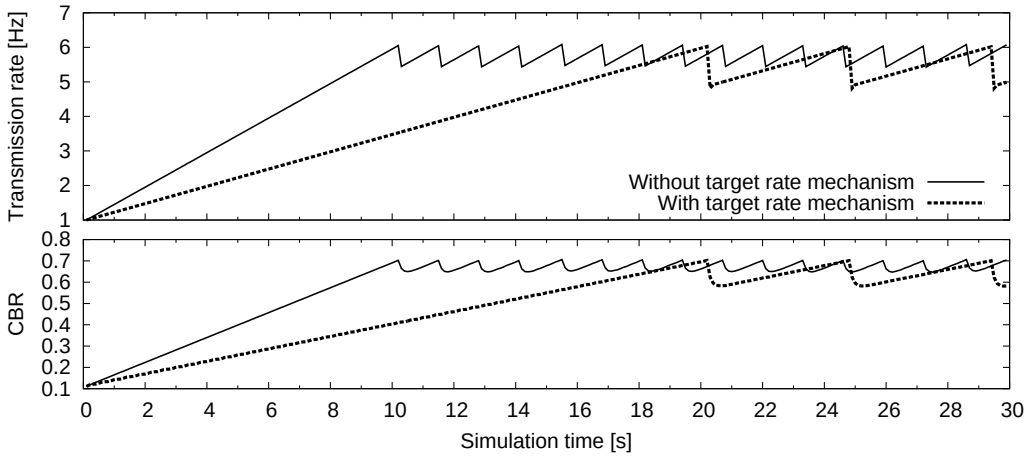


(b) Starting from 10 Hz

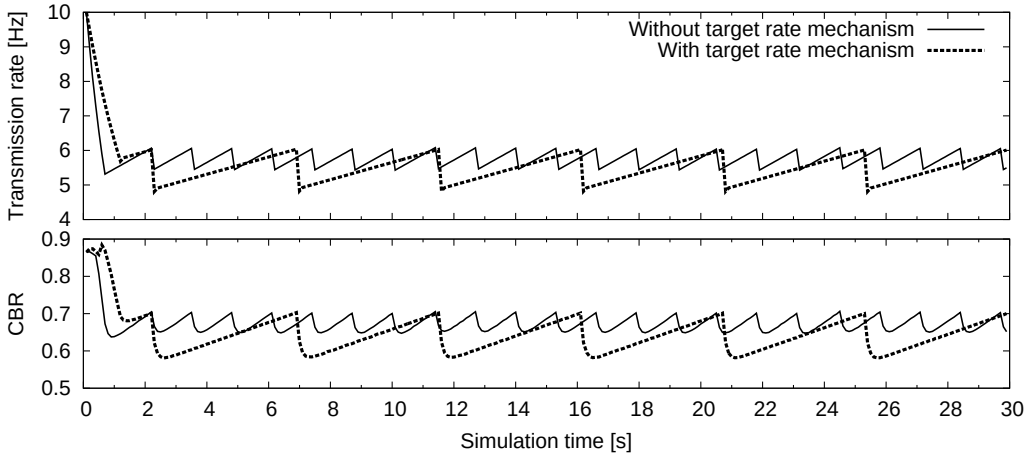
Figure 7.3: Impact of the adaptation interval on responsiveness and smoothness

Target Rate Mechanism

Figure 7.4 on the next page illustrates the convergence of AIMD with and without the target rate mechanism in the scenario considered before. We can see that the target rate approximately doubles the amplitude of the oscillations as expected. However, we also observe that it doubles responsiveness as well. This effect is actually an artifact of the considered scenario in which all nodes use exactly the same transmission rate. What happens is that when an increment decision is taken, nodes' transmission rates are typically slightly above target rate due to the calculation of the target rate as an Exponentially Weighted Moving Average (EWMA). The binary decision then results in a halving of the increment, which in turn doubles the convergence time to efficiency. Similarly, when decreasing rate, nodes' transmission rates are typically slightly above target rate and thus the decrement is doubled, which doubles the observed oscillations. The figure shows that the target rate mechanism should be further evaluated and improved in future work.



(a) Starting from 1 Hz

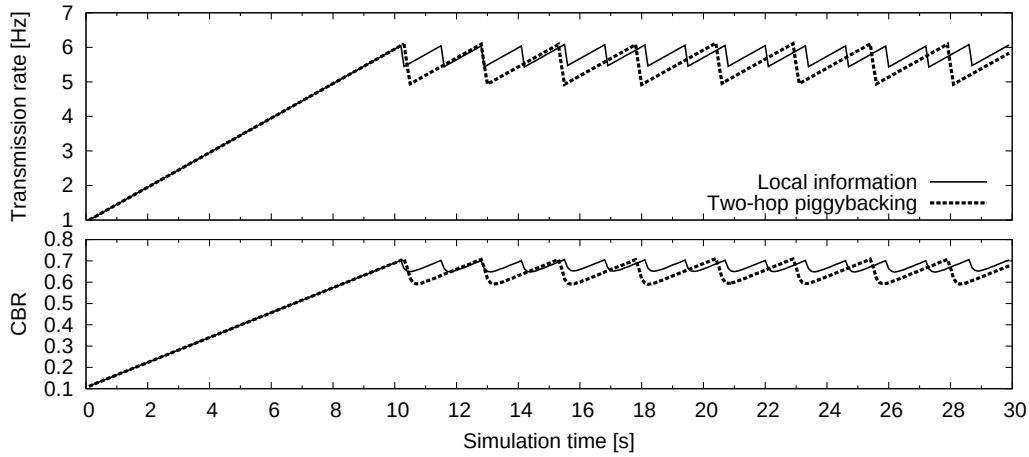


(b) Starting from 10 Hz

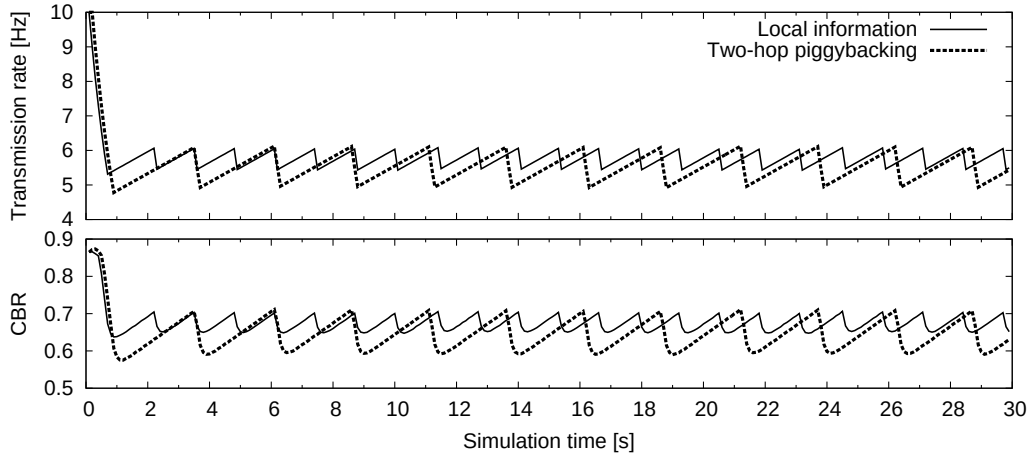
Figure 7.4: Impact of the target rate mechanisms on responsiveness and smoothness

Piggybacking Mechanism

The two-hop piggybacking mechanism introduced in Section 6.4.4 on page 160, introduces a reaction delay of one adaptation interval in order to prevent AIMD from reacting to the same information more than once, which might result in unfair transmission rate allocations. The impact of these delays on smoothness and responsiveness is illustrated by Figure 7.5 on the next page. The AIMD parameters used in the figure are $\alpha_I = 0.05$ Hz and $\beta_D = 0.1$. We can see that while responsiveness remains the same compared to an adaptation based on local knowledge, oscillation size is approximately doubled due to the introduced feedback delay.



(a) Starting from 1 Hz



(b) Starting from 10 Hz

Figure 7.5: Impact of two-hop piggybacking on responsiveness and smoothness

7.1.3 Resulting Trade-Offs

The choice of AIMD parameters is a trade-off between smoothness and responsiveness [120]. In the example in Figure 7.1 on page 169, the short convergence time using $\alpha_I = 1$ Hz and $\beta_D = 0.5$ comes at the price of high oscillations. On the other hand, $\alpha_I = 0.05$ Hz and $\beta_D = 0.1$ results in lower oscillations but a longer convergence time.

From a safety point of view, not only the system wide perspective in terms of channel load but rather the resulting transmission rate of an individual vehicle is of interest. If the oscillations of the allowed transmission rate are high, a vehicle may not able to transmit enough messages in a critical traffic situation. For this reason, we choose AIMD parameters with low oscillations in this thesis to avoid high undershoot. To compensate for the resulting longer convergence time, additional mechanisms could be introduced, similar to TCP's slow start mechanism, cf. Section 3.2 on page 53. For example, different AIMD parameters could be applied below a certain CBR

threshold, e.g., 0.5. Such a measure could be especially useful in the context of dynamic transmission rate boundaries and/or transmit power. In Section 7.3 on page 191, we apply different AIMD parameters based on the observed CBR value. Our preliminary results indicate that this approach can significantly improve responsiveness while allowing for comparably low oscillations near the target CBR. We leave a further exploration of this topic for future work.

In PULSAR, additional trade-offs between convergence and fairness arise from the choice of the adaptation interval t_{adp} and the piggybacking mechanism. From a responsiveness point of view, it would be desirable to set t_{adp} as low as possible and to use one-hop piggybacking of CBR information. However, as discussed in Section 6.4.4 on page 160, this leads to a significant decrease in global fairness.

To conclude, Table 7.1 on the following page summarizes the influence of each factor on PULSAR's smoothness and responsiveness.

7.2 Fairness and Stability

In the design of PULSAR, we laid special emphasis on fairness, i.e., local, global and participation fairness. In this section, we evaluate PULSAR's convergence behavior in different scenarios designed to stress the different protocol mechanisms. First, we focus on its long-term behavior in the static crossing scenario considered before, creating a single location with high interference. Then, we evaluate a highway scenario in which two groups of vehicles pass each other, creating a rapid change in vehicle density. For a better comparability of the results, we use uniform transmit power settings as well as transmission rate boundaries. These restrictions will be removed in the next section.

7.2.1 Long-Term Convergence in a Static Scenario

The scenario considered in this subsection consists of two 8 km long roads crossing each other at 4 km. Vehicles are placed statically on each road with a fixed inter-vehicle distance of 25 m. The intention behind this scenario is to create a single location in which interference from four directions accumulates, thus stressing the protocol's ability to maintain global fairness over a longer period of time. Note that the scenario is intended to resemble two highways crossing each other over a bridge rather than an intersection with conflicting traffic flows. In the latter case, nonuniform settings for transmit power and transmission rate boundaries would likely be required, which we avoid here for presentation clarity.

Influencing factor	Convergence time t_e	Oscillation size s_e
AI parameter α_I	If $U_o < U_t$, Eq. (7.1)	Max. overshoot, Eq. (7.3)
MD parameter β_D	If $U_o > U_t$, Eq. (7.2)	Max. undershoot, Eq. (7.4)
Number of nodes n	${t_e}^- \propto 1/n$	${s_e}^+ \propto n$
Adapt. interval t_{adp}	$\{{t_e}^-, {t_e}^+\} \propto t_{adp}$	-
Target rate	Less than ×2	Less than ×2
One-hop piggybacking	-	-
Two-hop piggybacking	-	Less than ×2

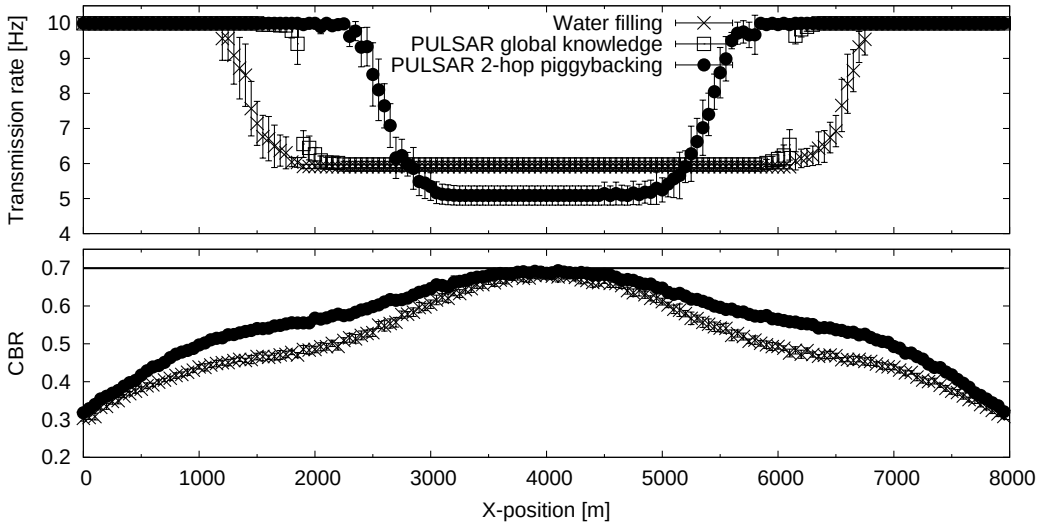
Table 7.1: Factors influencing PULSAR's responsiveness and smoothness

Comparison With Water Filling

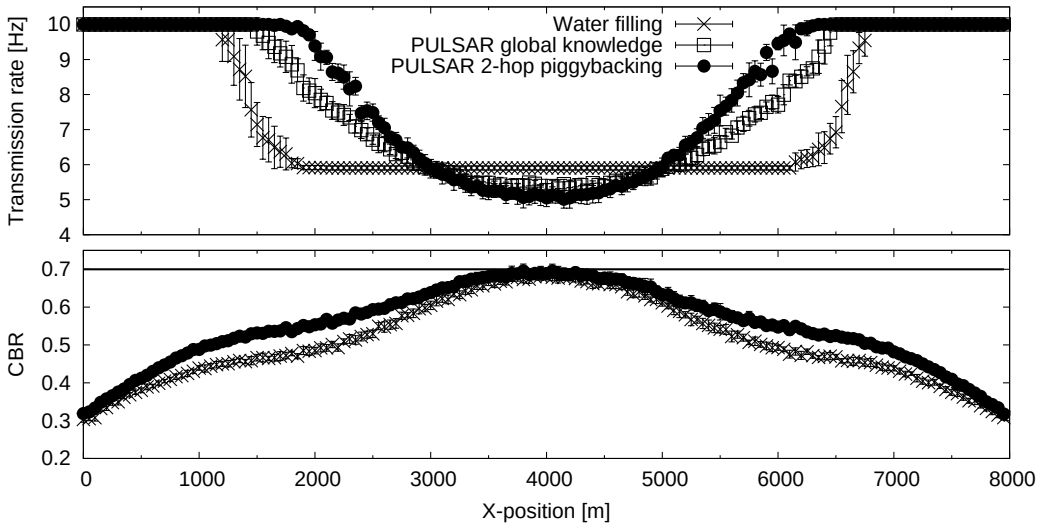
In Section 5.3.2 on page 121, we defined global fairness as a best-effort approach to achieve max-min fairness. Due to the unbounded nature of wireless signal propagation, max-min fairness is difficult to define and to evaluate in a wireless context. To approximate the max-min fair transmission rate allocation in our scenario, we use a water filling approach, cf. "progressive filling" in [119], in combination with the participation range as defined by Equation (5.2) on page 122. In the following, the transmit power is set to 20 dBm unless specified otherwise, resulting in a participation range of 2138 m.

We implement the water filling approach as follows. Starting the from a transmission rate of 1 Hz, each node increases its transmission rate by 0.05 Hz every 100 ms until at least one node within its participation range measures a CBR above the CBR threshold. This condition is evaluated in real time based on global knowledge from the simulator, i.e., not from the information piggybacked in the exchanged messages. If the condition is met, the node decreases its transmission rate by one step, i.e., 0.05 Hz, and then stops the adaptation process. To account for the fact that PULSAR reacts very sensitively to any measurement exceeding the CBR threshold, we reduce the CBR threshold to 0.69 for water filling.

Figure 7.6 on the next page compares PULSAR's transmission rate allocation after 1500 adaptation steps of t_{adp} = 400 ms with and without the target rate mechanism to the max-min fairness approximation resulting from the water filling approach. The shown values correspond to the average of five independent simulation runs. Additionally, the 95% confidence intervals are indicated. The same statistical background is true for the remainder of this chapter unless indicated otherwise. Due to the symmetry of the considered scenario, the figure shows the observed CBR and resulting transmission rate allocation for only one of the two roads. The highest load is observed at x-position 4000 m, where the second road crosses.



(a) Without target rate mechanism



(b) With target rate mechanism

Figure 7.6: Comparison of the convergence of PULSAR and water filling in the static crossing scenario

In Figure 7.6a, we can see that the water filling approach results in a transmission rate of approximately 5.9 Hz in the center of the intersection. In the following, we use this value as a reference for max-min fairness. With global knowledge from the simulator, i.e., the real-time measurements of all nodes within participation range, PULSAR without target rate converges to approximately 5.97 Hz at the intersection center, a result which lies within the confidence interval of the water filling outcome. The difference is thus not statistically significant. With two-hop piggybacking, we observe a minimum transmission rate of 5.09 Hz, which corresponds to 86% of the water filling result. This difference results from the limited dissemination range of

the piggybacking approach, as we discuss later in this section when we evaluate the efficiency of the information dissemination process.

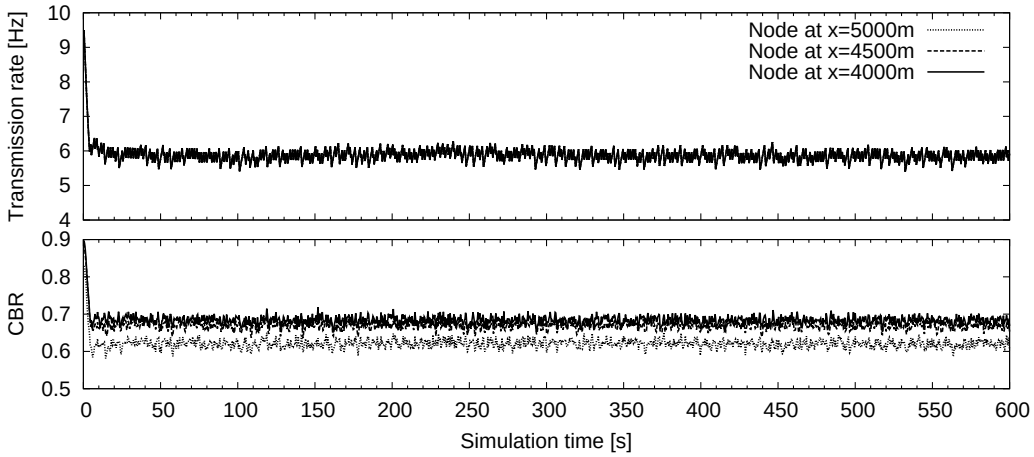
Impact of the Target Rate Mechanism

Figure 7.6b illustrates the impact of the target rate mechanism on PULSAR's spatial convergence compared to the water filling approach. We can see that compared to without target rate, the resulting transmission rate allocation has more the shape of a semicircle than of a rectangle. We further observe that the minimum transmission rate with global knowledge is now 5.27 Hz compared to 5.9 Hz before. While the participation range remains the same, the target rate mechanism causes nodes further away from the intersection to increase their transmission rates and thus to generate more load at the intersection center. As a result, nodes in the intersection center converge to a lower rate. With two-hop piggybacking on the other hand, the minimum observed transmission rate is 5.05 Hz, which lies within the confidence interval of the result without target rate. In this case, the target rate mechanism stabilizes distanced nodes at a lower rate than without the target rate mechanism. Since these nodes receive congestion information with lower probability, they tend to increase their rates more often than those in the intersection center. The target rate mechanism works against this trend by doubling the increment of the most congested nodes in the intersection center and halving those of more distanced nodes.

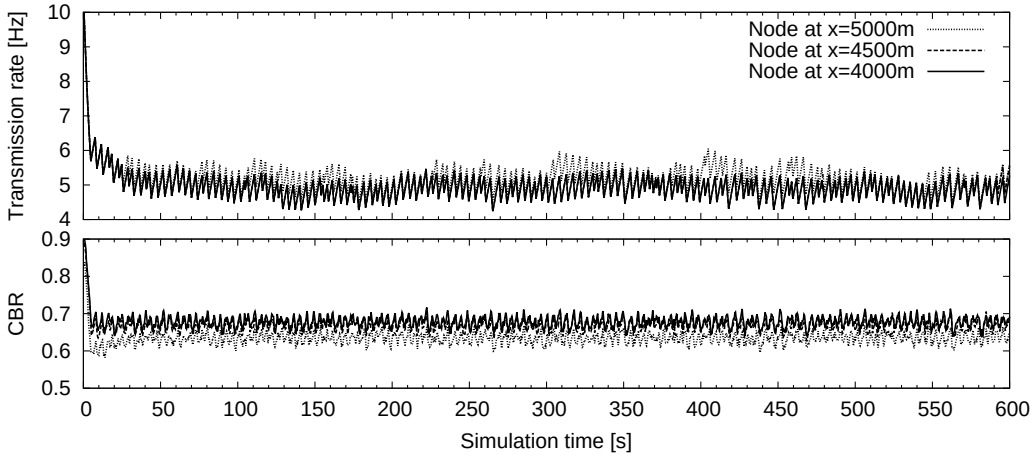
Figure 7.7 on the facing page illustrates the observed CBR and resulting transmission rate with respect to simulation time for three exemplary nodes located at x-positions 4000 m, 4500 m and 5000 m, respectively. For better comparability, all plots shown in the figure use the same seed of the random number generator. Note that the final transmission rates and CBR values at simulation time 600 s contributed to the snapshots shown in Figure 7.6 on the previous page.

In Figure 7.7a, we can see that with global knowledge and without the target rate mechanism, all three considered nodes reduce their transmission rates to approximately 6 Hz and then stay there throughout the duration of the simulation. Figure 7.7b shows that without global knowledge, the three nodes initially stop reducing their transmission rates near 6 Hz but then go further down to approximately 5 Hz. At this equilibrium point, sufficient congestion feedback from the intersection reaches the distanced nodes to prevent them from further increasing their transmission rates. Note that compared to global knowledge, oscillations are higher in amplitude due to the delay in feedback. In addition, the average transmission rate over time is not as constant, especially for the node at x-position 5000 m. In Figure 7.7c, we can see that the target rate mechanism has a stabilizing effect on the average transmission rate. Especially the node at 5000 m now shows a more stable behavior.

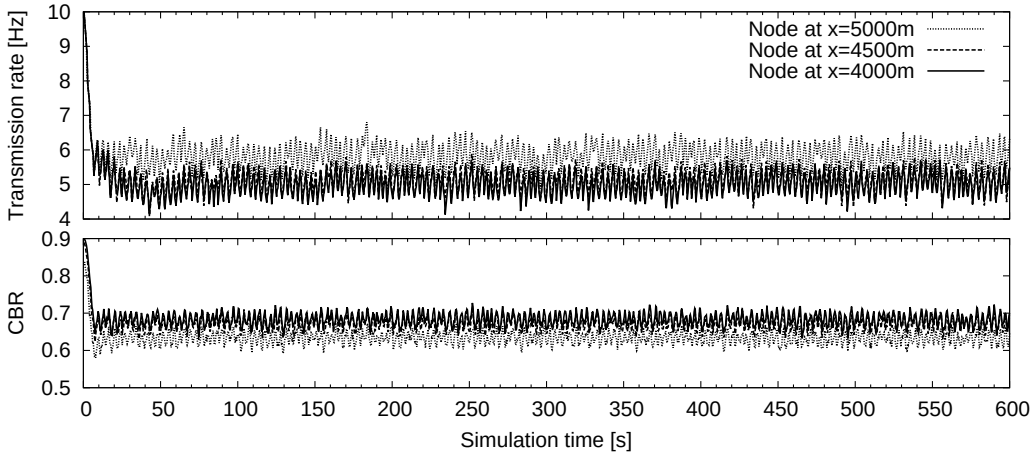
The stabilizing effect of the target rate mechanism is further illustrated by Figure 7.8 on page 181, which shows the temporal convergence of PULSAR in the same setup but with an adaptation interval of t_{adp}=200 ms compared to t_{adp}=400 ms in Figure 7.7.



(a) Global knowledge without target rate mechanism



(b) Two-hop piggybacking without target rate mechanism



(c) Two-hop piggybacking with target rate mechanism

Figure 7.7: PULSAR's rate adaptation over time in the crossing scenario, t_{adp} = 400 ms

We observe that without the target rate mechanism, the transmission rates of all three nodes show long-term oscillations. With target rate on the other hand, a more stable average transmission rate is maintained after a steady state has been reached at approximately simulation time 45 s. We can further see that the node at the center of the intersection now converges to approximately 4 Hz compared to approximately 5 Hz with t_{adp}=400 ms. We evaluate why this is the case in the following.

Impact of the Adaptation Interval

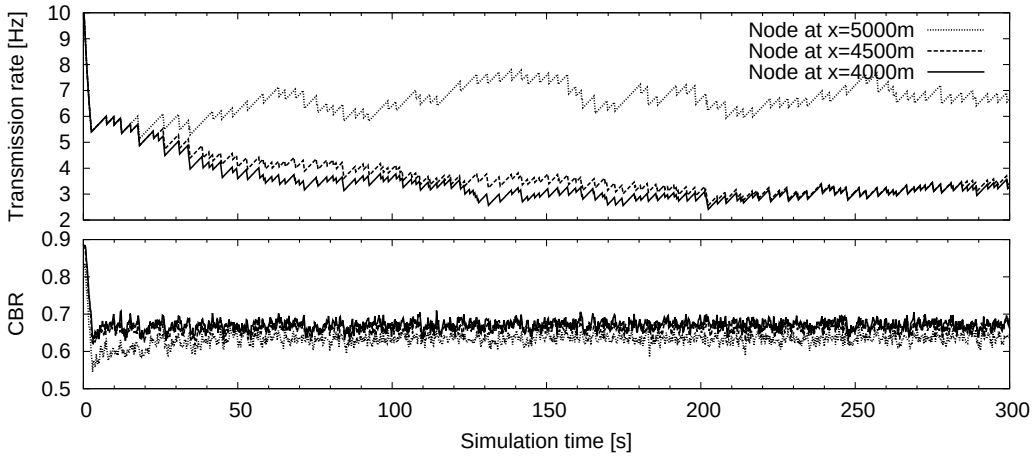
Figure 7.9 on page 182 illustrates the spatial convergence of PULSAR after 1500 adaptation steps for different adaptation intervals t_{adp} compared to the outcome using global knowledge. In Figure 7.9a, we observe a significant difference between the resulting transmission rates at the intersection center. While for t_{adp}=400 ms, the average transmission rate of the node at x-position 4000 m is 5.09 Hz, it is 3.68 Hz and 2.68 Hz for t_{adp}=200 ms and t_{adp}=100 ms, respectively. Note here that the adaptation interval t_{adp} also determines the maximum age of the congestion information taken into account in the adaptation process. Within a longer period of time, more messages containing congestion information can be generated and forwarded, resulting in a higher dissemination range, cf. Section 6.4.3 on page 158.

As shown in Figure 7.9b, the target rate mechanism reduces the difference between the outcomes for the considered adaptation intervals. The resulting average transmission rate at the center of the intersection amounts to 5.05 Hz, 4.54 Hz and 4.13 Hz for t_{adp}=400 ms, t_{adp}=200 ms and t_{adp}=100 ms, respectively. In other words, the target rate mechanism prevents the nodes in the most congested location from reducing their transmission rates too far and is thus able to mitigate the impact of a reduced reception probability at further distances from the congested location.

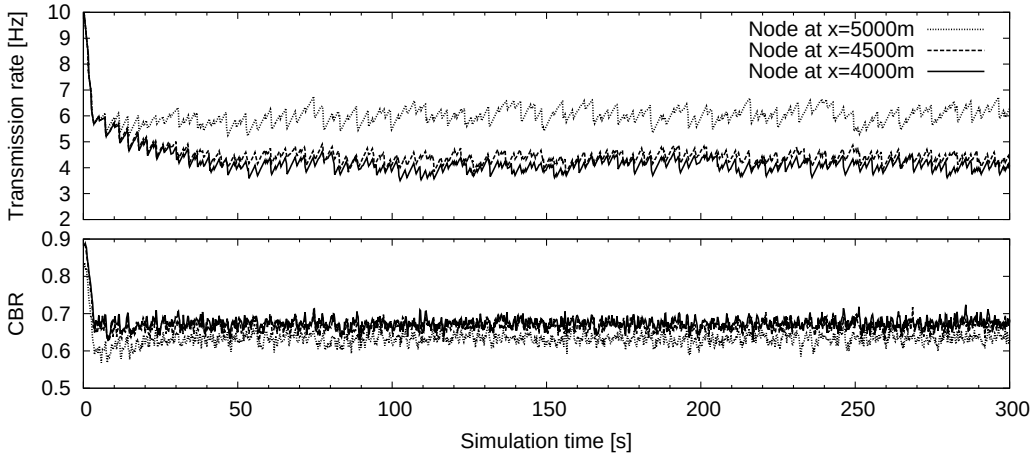
To quantify the probability of receiving congestion information within a certain amount of time, we set up a scenario consisting of a linear road of 6 km length with a fixed inter-vehicle distance of 10 m and no node mobility. We set the transmission rate to 5.5 Hz, resulting in a CBR of approximately 0.7 between x-positions 1500 m and 4500 m. We configured all nodes between x-positions 1700 m and 2000 m to piggyback a congestion bit in their messages. Figure 7.10 on page 183 illustrates the probability of receiving one of these piggybacked bits with respect to the distance to the end of the reporting region, i.e., to x-position 2000 m.

Figure 7.10a illustrates that while a longer time interval helps to extend the dissemination range of congestion information, one-hop piggybacking is not sufficient to reach up to the desired participation range of 2138 m for a transmit power of 20 dBm. This result confirms the theoretical considerations in Section 6.4.3 on page 158.

Figure 7.10b illustrates the outcome for two-hop piggybacking with a participation distance of 2138 m and a hop distance of 1100 m. We observe significant differences between the dissemination distances for different time intervals. Between 400 ms and 500 ms, however, the increase in dissemination range is much smaller. At this



(a) Without target rate mechanism

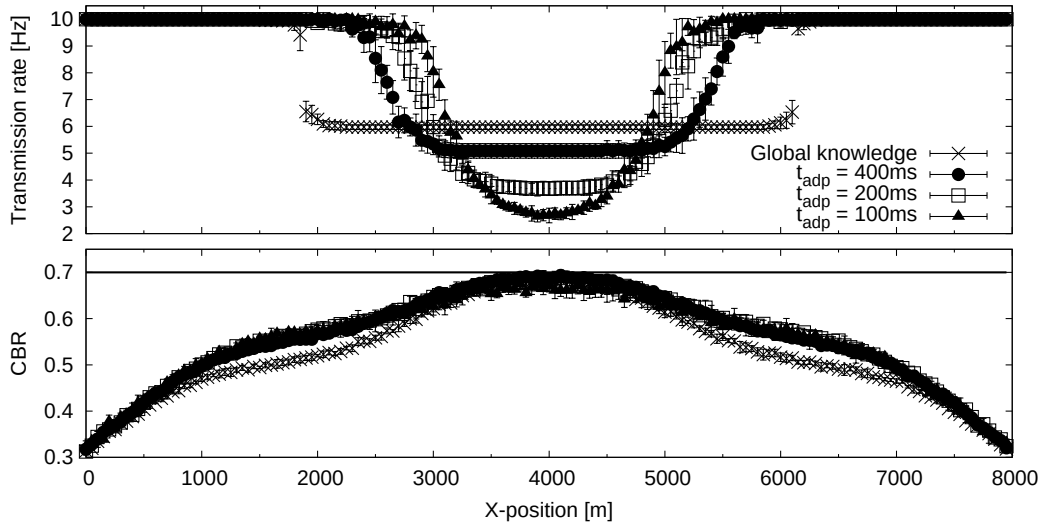


(b) With target rate mechanism

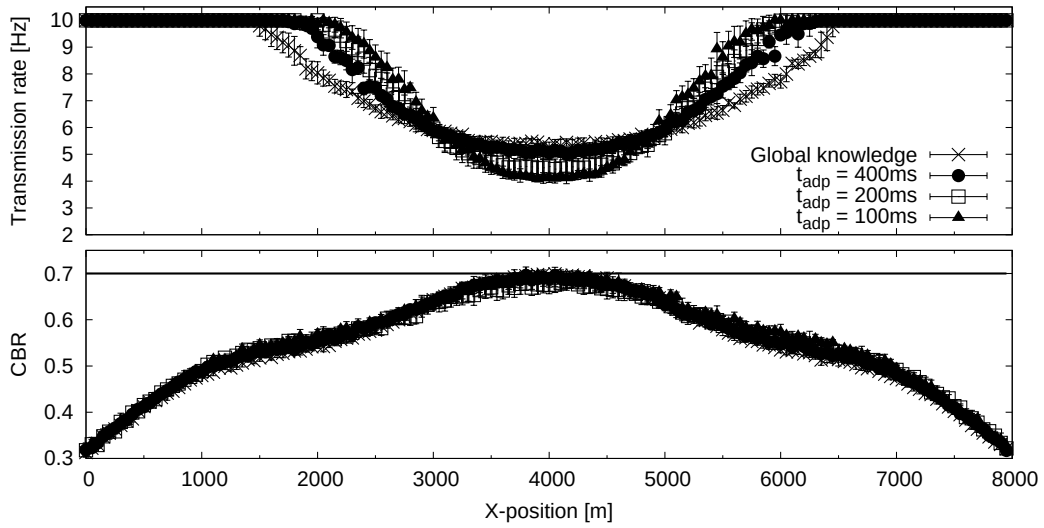
Figure 7.8: Stabilizing effect of target rate mechanism - PULSAR's rate adaptation over time in crossing scenario, t_{adp} = 200ms

point, the probability of receiving an individual message is too low to further increase the aggregated reception probability significantly, cf. Figure 6.17 on page 160. We also notice that the reception probability beyond the configured participation range is reported as zero, which corresponds to the probability as seen by the node but not the actual reception probability at this distance.

This aspect is illustrated by Figure 7.10c, which shows the reported reception probability for different values of participation distance d_{part} and hop distance d_{hop}. When both are set to very large distances, e.g., 10 km and 5 km, respectively, we can see that the information can be propagated up to a distance of 3000 m during the considered time interval of 400 ms. The aggregated probability of reception at the participation distance is 37.8%, in contrast to close to 0% when d_{part}=2138 m is applied. This discrepancy results from the fact that a receiving node ignores two-hop feedback whose sum of sender-receiver distance and d_{hop} exceeds d_{part}. The set of potential



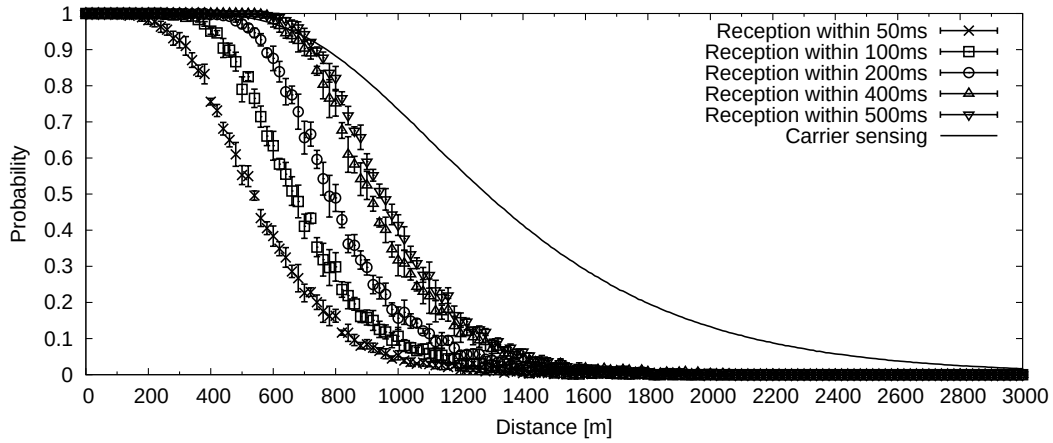
(a) Without target rate



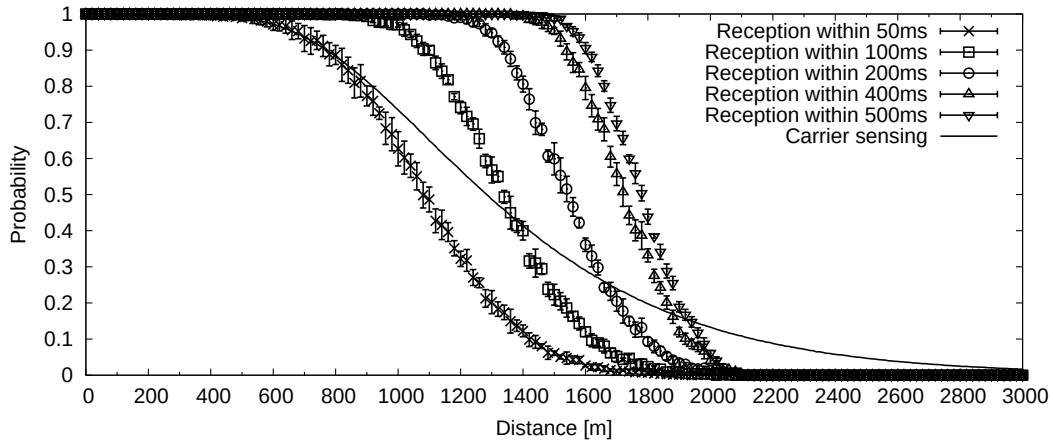
(b) With target rate

Figure 7.9: Impact of the adaptation interval on PULSAR's convergence in the crossing scenario

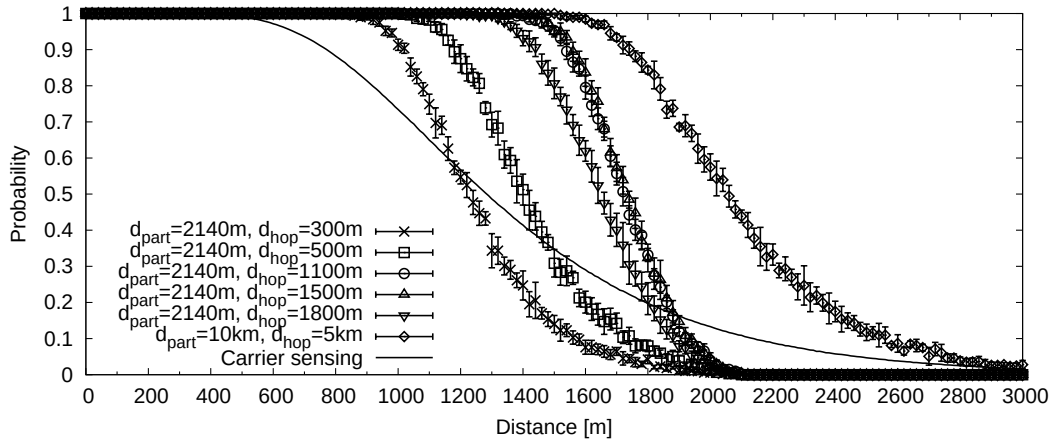
sender-forwarder pairs to result in an aggregated distance near d_{part} is thus small, resulting in a low probability. The figure further shows that for best results, d_{hop} should be chosen near 50% of d_{part}. Hop distances which are significantly smaller or significantly larger result in a decreased dissemination distance. In future work, the hop distance concept could be extended to reflect the congestion state in a distance-dependent way. For example, different congestion bits could be used to convey the congestion state within multiple distances from the transmitter.



(a) Impact of t_{adp} on one-hop piggybacking



(b) Impact of t_{adp} on two-hop piggybacking (d_{part}=2138 m, d_{hop}=1100 m)



(c) Impact of d_{hop} and d_{part} on two-hop piggybacking (t_{adp}=400 ms)

Figure 7.10: Probability of receiving a piggybacked bit within a certain time limit; linear road with transmit power 20 dBm, CBR 0.7, carrier sense threshold -95 dBm

Impact of the Participation Distance

Figure 7.11 on the facing page illustrates the impact of the participation distance on the spatial convergence of PULSAR using two-hop piggybacking and t_{adp}=400 ms. We can see that there is not significant difference between the resulting transmission rates at the intersection center for d_{part}=2140 m and d_{part}=10 km. However, in the latter case it can occur that nodes located up to 3 km away from the intersection participate in congestion control, while their contribution to congestion is negligible. The example thus illustrates the benefit of using a limitation of the participation distance. On the other hand, if the participation distance is chosen too small, we observe a significant decrease of the transmission rates near the center of the intersection, which contradicts the global fairness principle. Comparing Figures 7.11a and 7.11b, we can see that the target rate mechanism again mitigates the impact of a smaller participation distance on global fairness.

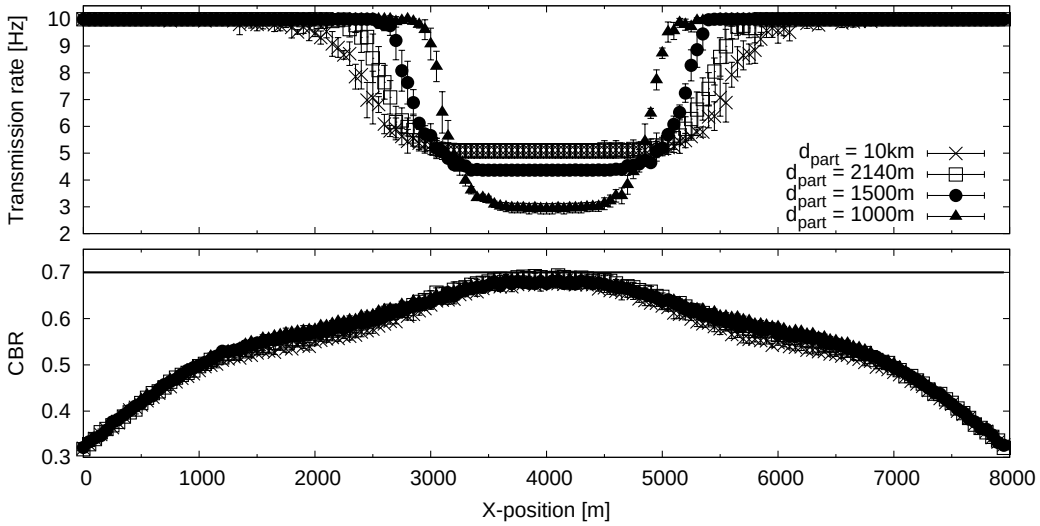
Impact of AIMD Parameters

As the last aspect in this subsection, we study the impact of the choice of AIMD parameters on the spatial convergence of PULSAR. Figure 7.12 on page 186 illustrates the outcome of three different AIMD parameter configurations with and without the target rate mechanism. The figure shows that for global fairness, it is beneficial to use a larger additive increase parameter α_I over a smaller one and a smaller multiplicative decrease parameter β_D over a larger one. To explain this result, consider again Figure 7.10b on the preceding page which shows that the probability of receiving congestion feedback is significantly higher near the origin of the information. In other words, nodes near the congested location are much more likely to receive a congestion feedback and thus to reduce their transmission rates, while those further away might not receive it and thus increase their transmission rates instead. A relatively high increment parameter and relatively low decrement parameter benefit the nodes in the most congested location and are thus beneficial for global fairness.

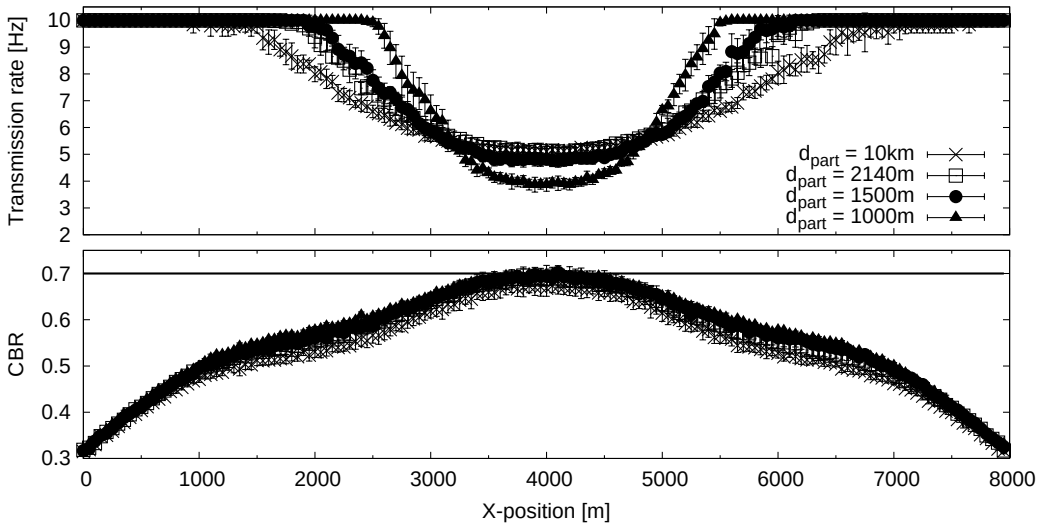
7.2.2 Performance in a Dynamic Highway Scenario

While so far, we focused on static scenarios only in order to better illustrate PULSAR's long-term behavior, we now add node mobility to study its behavior in a more realistic highway environment. The considered scenario consists of a total of 451 vehicles split into two groups passing each other on a five-lane highway at approximately 100 km/h. The scenario was designed to create a rapid change in vehicle density, challenging the protocol's ability to adapt quickly while maintaining local and global fairness. The vehicle movement traces for the simulation were created using the microscopic traffic simulator VISSIM[1].

[1] *http://www.ptv.de*



(a) Without target rate

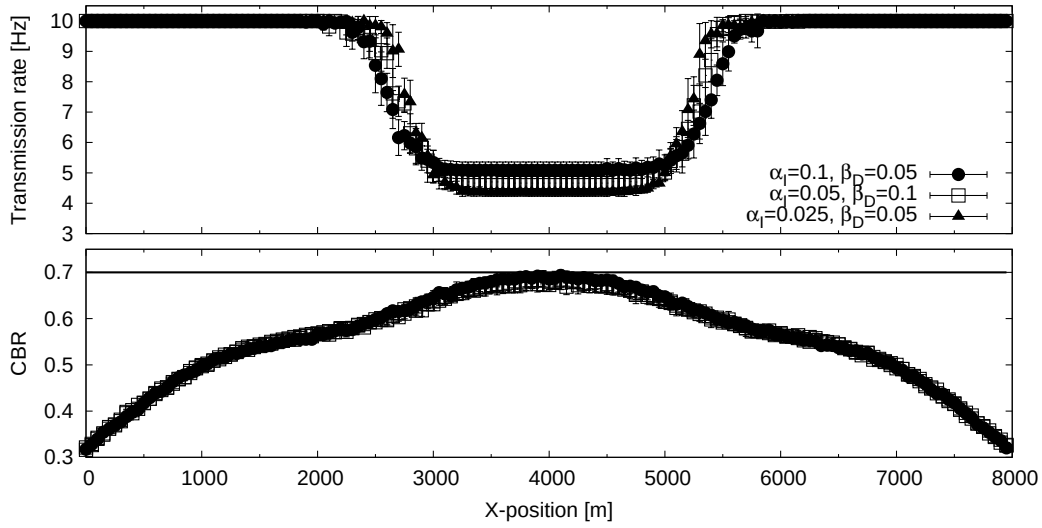


(b) With target rate

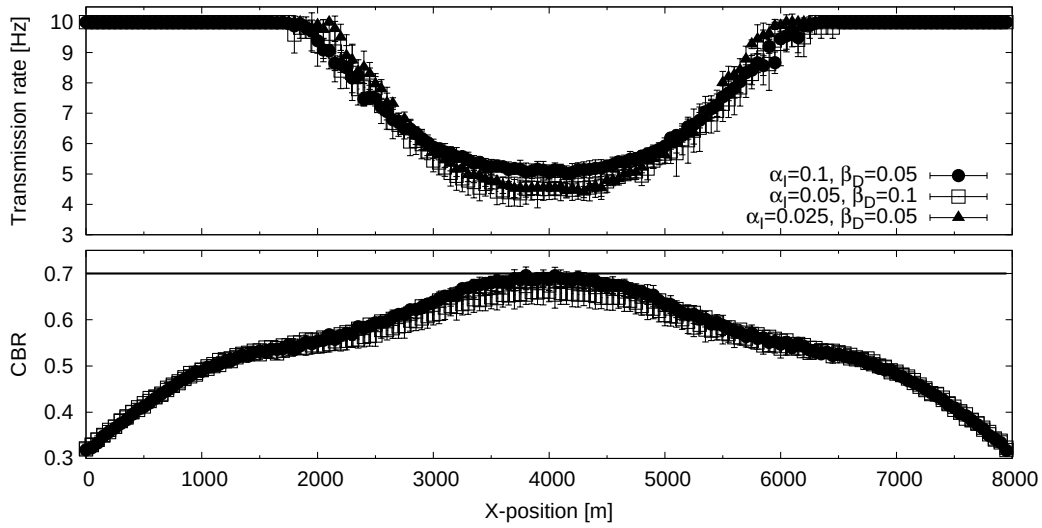
Figure 7.11: Impact of the participation distance d_{part} on PULSAR's convergence in the crossing scenario

Figure 7.13 on page 187 illustrates the vehicles' positions on the highway in six snapshots taken in 25 s to 30 s intervals during the course of the simulation. The highest vehicle density is reached between simulation times 60 s and 90 s, when the two groups are on level with each other.

Figure 7.14 on page 188 illustrates the measured CBR and resulting transmission rate for each snapshot with respect to the x-position on the road. Like before, the transmit power was set to 20 dBm and the minimum and maximum transmission rate is 1 Hz and 10 Hz, respectively. The adaptation interval t_{adp} was set to 200 ms and the target rate mechanism was applied. Each subfigure compares PULSAR's transmission rate



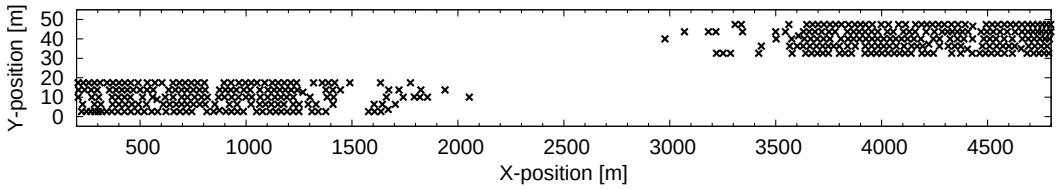
(a) Without target rate



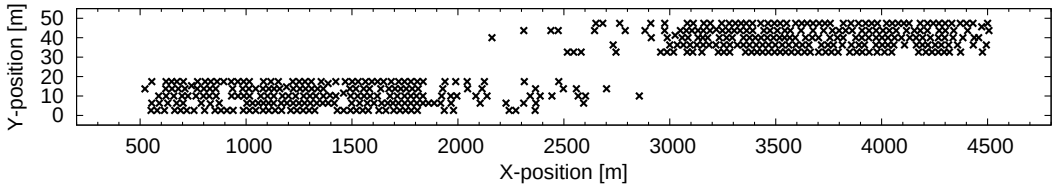
(b) With target rate

Figure 7.12: Impact of AIMD parameters on PULSAR's convergence in the crossing scenario

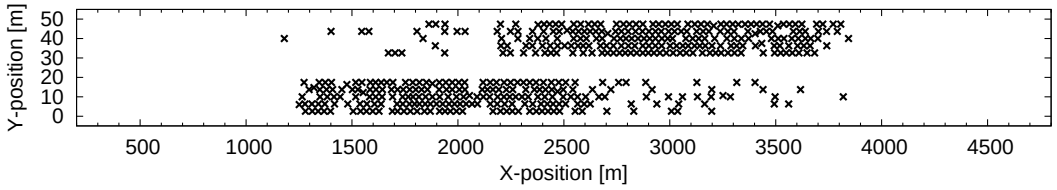
allocation and CBR value against the outcome for the water filling approach. To generate the latter values, we used static snapshots of the vehicle positions at the respective simulation times. The shown values are averages of five independent simulation runs. We observe that there is typically no statistically significant difference between the outcomes for PULSAR and water filling. Sometimes, PULSAR's transmission rate allocation lies above the one for water filling, which originates from the reduced CBR target of 0.69 for water filling. We conclude that the global fairness principle, i.e., a best effort approach to fulfill max-min fairness, has been fulfilled.
Figure 7.15 on page 190 illustrates the impact of the target rate mechanism on PULSAR's convergence behavior under mobility. The figure shows a further snapshot



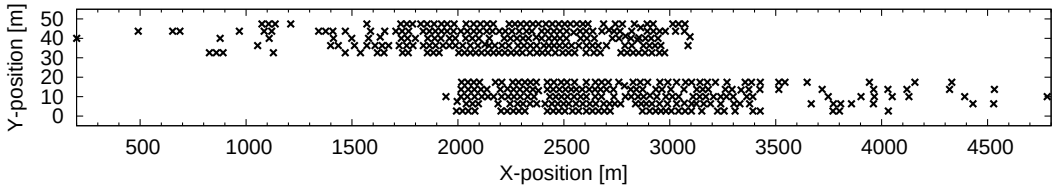
(a) Simulation time 5 s



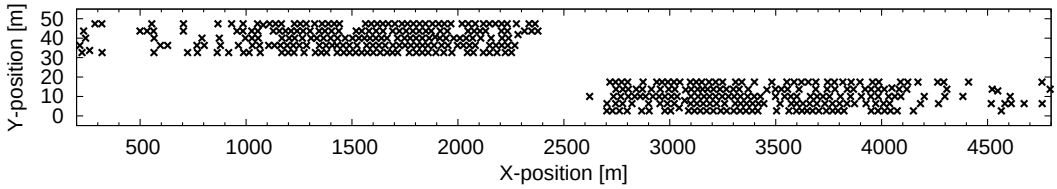
(b) Simulation time 30 s



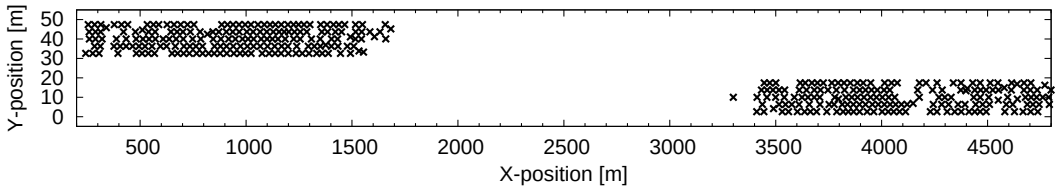
(c) Simulation time 60 s



(d) Simulation time 90 s

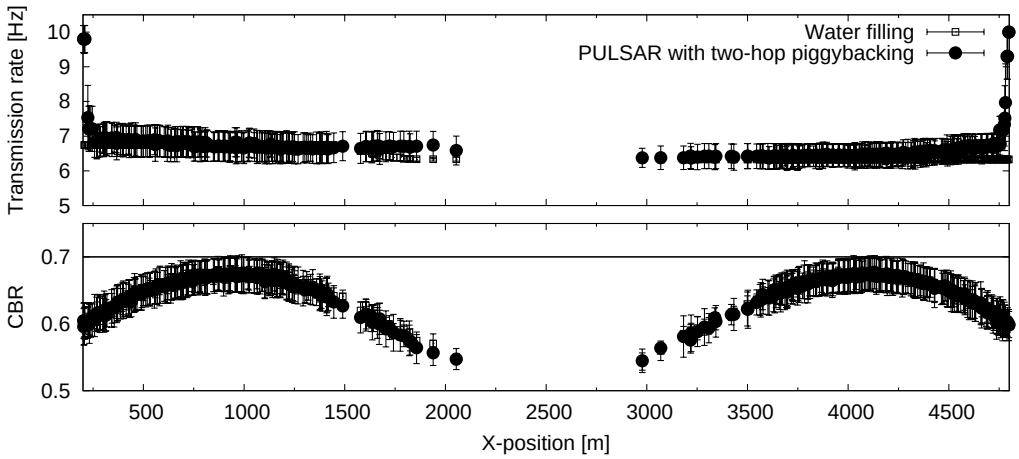


(e) Simulation time 120 s

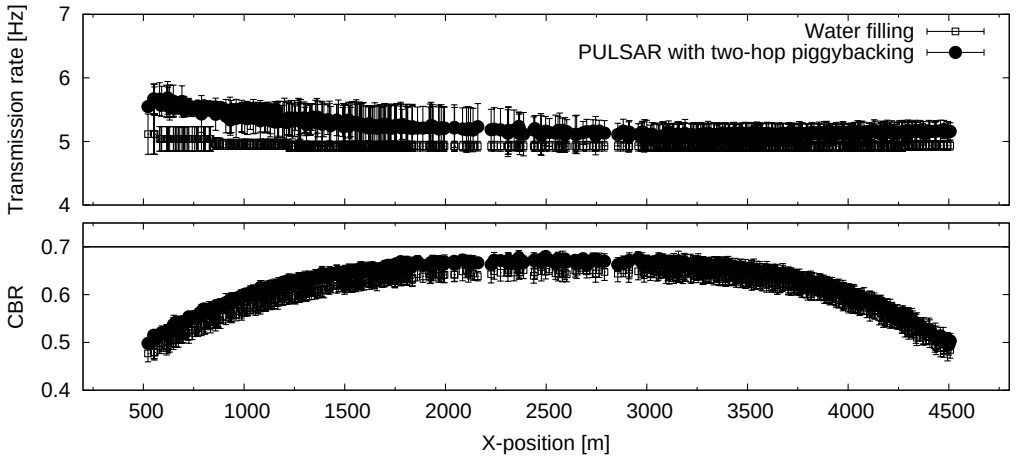


(f) Simulation time 150 s

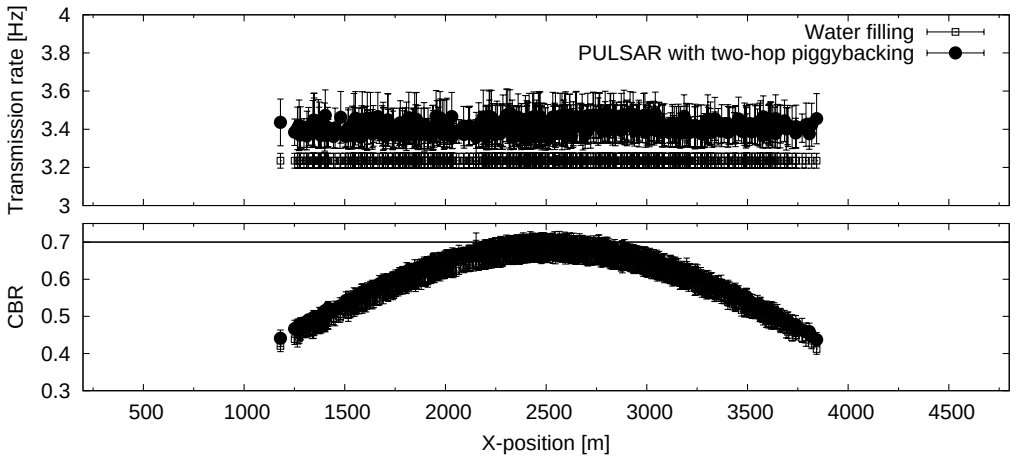
Figure 7.13: Snapshots of vehicle positions in scenario with two approaching groups of vehicles



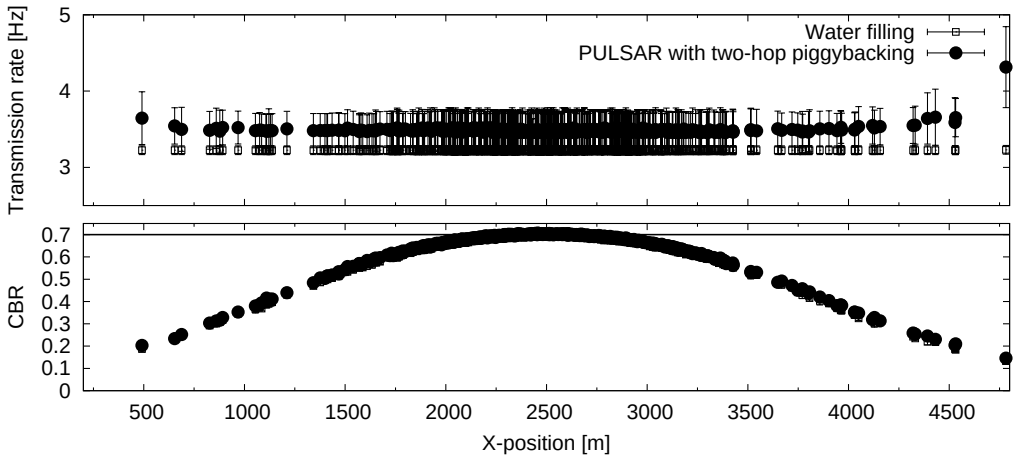
(a) Simulation time 5 s



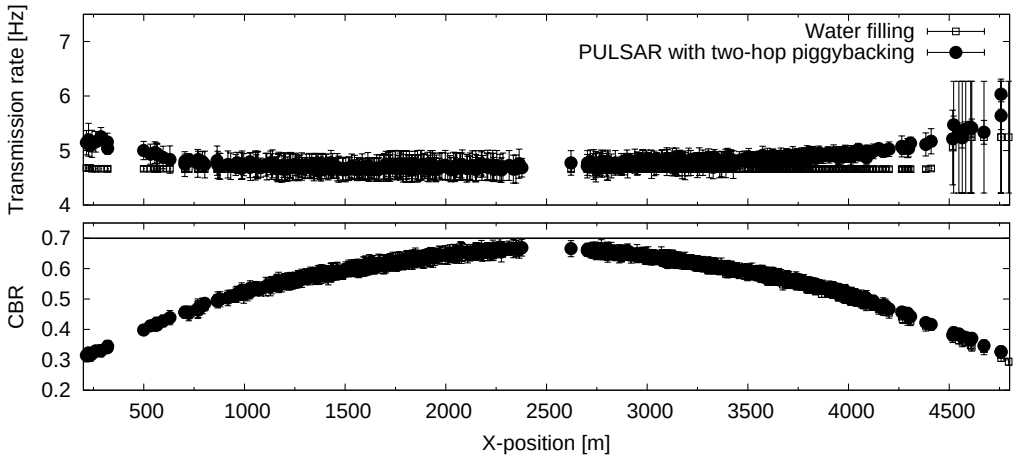
(b) Simulation time 30 s



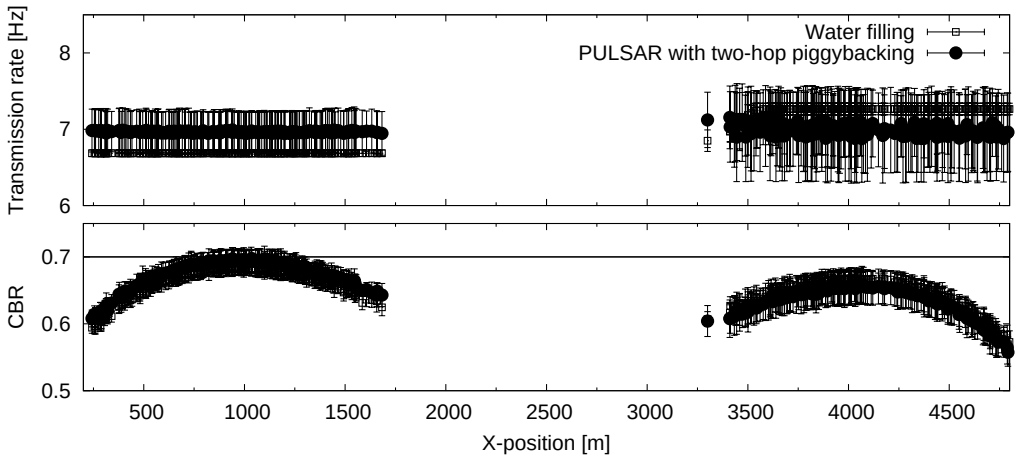
(c) Simulation time 45 s



(d) Simulation time 60 s

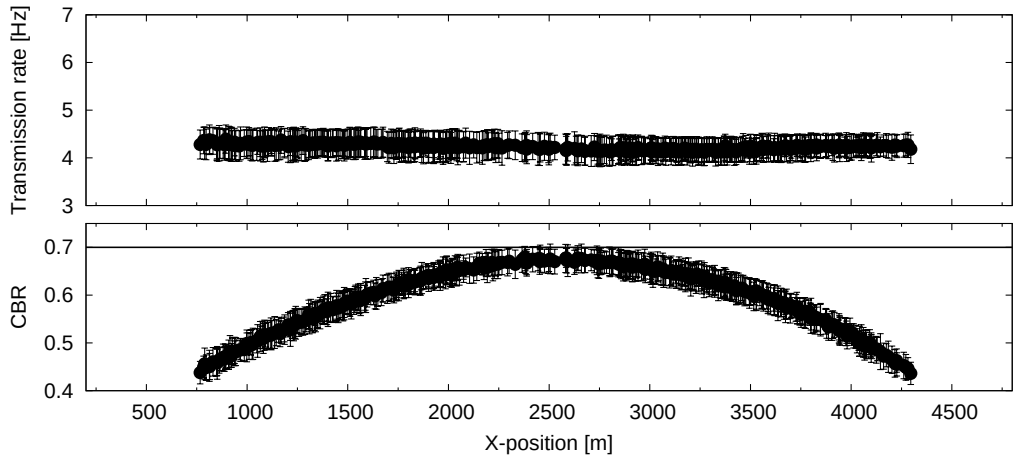


(e) Simulation time 90 s

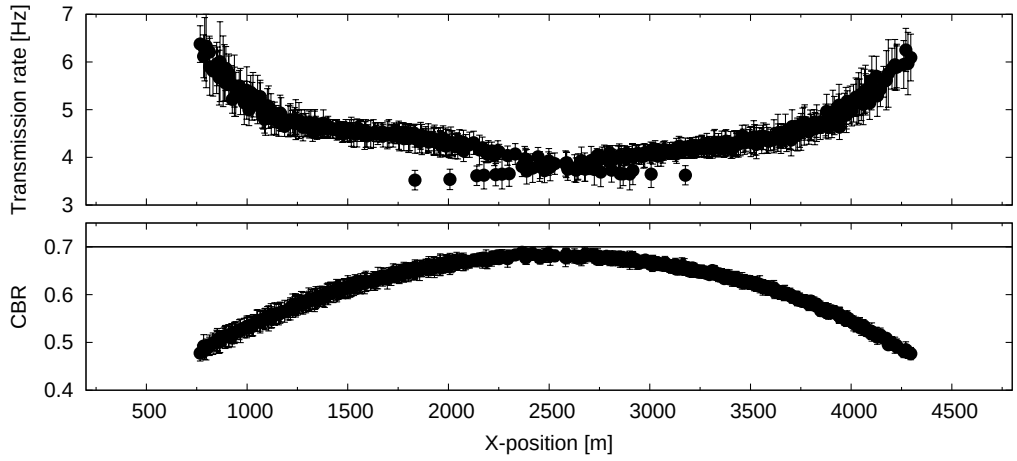


(f) Simulation time 150 s

Figure 7.14: Snapshots of PULSAR's transmission rate and observed channel load in a scenario with two passing groups of vehicles



(a) With target rate



(b) Without target rate

Figure 7.15: Impact of the target rate mechanism on PULSAR's convergence in a scenario where two groups of vehicles pass each other (simulation time 40 s)

of the simulation at simulation time 40 s, at which the two groups of vehicles are in the process of merging into one group from a channel perspective. In Figure 7.15a, we can see that with target rate, a homogeneous transmission rate of approximately 4.3 Hz is assigned to all vehicles. In contrast, Figure 7.15b shows that without the target rate mechanism, the convergence to fairness is not fast enough to keep up with vehicle mobility. We observe an X-shaped transmission rate allocation which violates the local fairness principle. In addition, some vehicles reduce their transmission rates to approximately 3.5 Hz, which is 0.8 Hz lower than with the target rate mechanism. The example highlights the importance of the target rate mechanism to maintain local and global fairness.

To conclude this subsection, Figure 7.16 on page 192 illustrates the transmission rate and measured CBR as seen by an individual vehicle traversing the scenario from x-position 4778 m at simulation time 0 s to x-position 260 m at simulation time 180 s.

The figure compares the outcome for two-hop piggybacking against the outcome using global knowledge from the simulator. We can see that the rate allocation with piggybacking follows the one with global knowledge with somewhat increased oscillations due to the resulting information dissemination delay.

7.2.3 Dynamic Change of Minimum and Maximum Rate

So far, we evaluated PULSAR's convergence behavior based on uniform minimum and maximum transmission rates for all nodes. In the following, we remove this restriction and perform a stress test of PULSAR's ability to adapt the relative transmission rate instead of the absolute one, cf. Section 6.3.4 on page 151. For this purpose, we use the previously introduced scenario in which 200 nodes share the same location. In intervals of 1 s, every node randomly chooses a new minimum and maximum transmission rate from the interval [1 Hz; 20 Hz]. The underlying distribution is exponential with a mean of 5 Hz. We set the adaptation interval t_{adp} to 200 ms and the AIMD parameters to $\alpha_{I,rel} = 0.01$ and $\beta_{D,rel} = 0.05$.

Figure 7.17 on page 193 illustrates the measured CBR and applied relative transmission rate by all nodes as well as the absolute transmission rates and rate boundaries of three exemplary nodes A, B and C. In Figure 7.17a, we can see that despite the randomly changing minimum and maximum transmission rates exmplarily illustrated by Figures 7.17c to Figure 7.17e, PULSAR is able to maintain the CBR near the target of 0.7. In Figure 7.17b, we observe that after an initial decrement phase, the relative transmission rate stays between 0.25 and 0.45.

7.3 Impact on an Exemplary Safety Application

One of PULSAR's five design principles is to operate within the limits provided by safety applications. For this purpose, the protocol expects input from the application layer in the form of a minimum and maximum transmission rate as well as a transmit power. Since the definition of safety applications' requirements is a research topic of its own and out of the scope of this thesis, in the following we make a number of simplifying assumptions based on the work of An et al. [24] with the objective to create a reasonably realistic input for PULSAR.

First of all, the considered safety application is hypothetical in the sense that it is a Forward Collision Warning (FCW) application implemented with global knowledge from the simulator. In other words, the selection of the vehicle for which transmission parameters are optimized is optimal. The results for an actual FCW implementation based on received messages may vary and could be explored in future work. In particular, the selection of transmission parameters and the identification of the targeted vehicle may influence each other in an actual implementation.

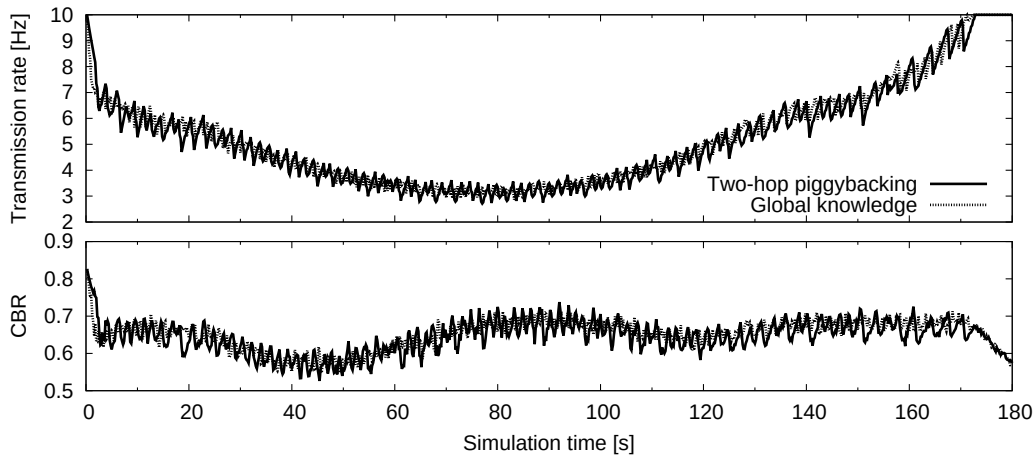


Figure 7.16: Adaptation by an exemplary vehicle in a scenario where two groups of vehicles pass each other

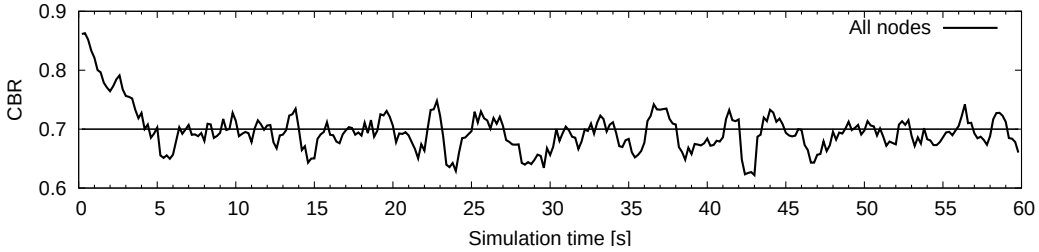
Second, the applied metric of Inter-Reception Distance (IRD) may result in a simplified view of safety, since it assumes a uniform acceleration of both vehicles. In contrast to the approach by An et al. [24], it thus does not specify how long a remote vehicle can rely on the received information. Note that while in the following, we select transmission parameters based on channel load and observe the resulting IRD, An et al. selected transmission parameters to meet a specific reception requirement and observed the resulting *channel load*. In future work, it could be evaluated to which extent the considered metric reflects the actual requirements of a safety application. Further, it could be explored how other awareness metrics can be integrated into the concept of relative transmission rate.

For these reasons, the results presented in this section are intended as first results to demonstrate the potential of the relative transmission rate concept.

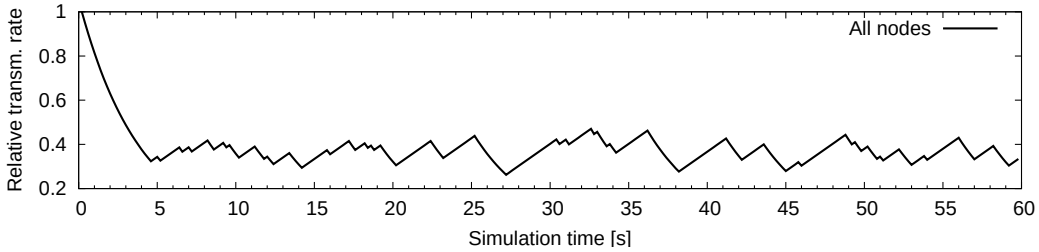
7.3.1 Considered Safety Application

For simplicity and to make the evaluation less dependent on the actual design of the application, we implement an FCW-like[2] application using global knowledge of the simulator. More specifically, each *ego vehicle*, i.e., the one running the FCW application, tracks two of its neighbors, i.e., the one directly in front of it in the same lane and the one directly behind it in the same lane, which we denote as *leading vehicle* and *trailing vehicle*, respectively. Note that for an actual FCW implementation, it may make sense to extend the tracking range to the ones after the leading and trailing vehicle, respectively, to account for dangerous situations resulting from lane changes. However, our main purpose in this evaluation is to see how PULSAR reacts

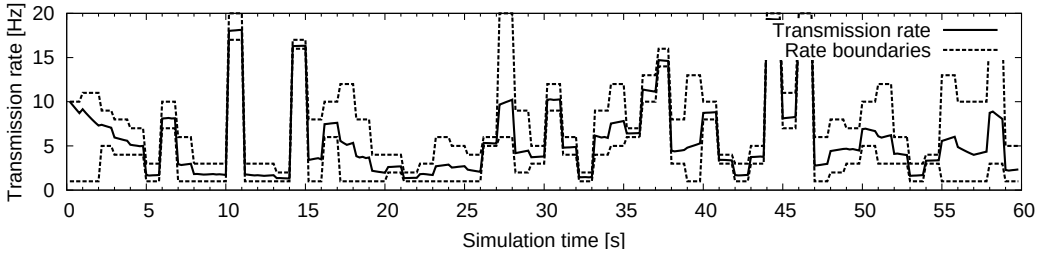
[2] We chose the FCW application since it tracks only a single vehicle, cf. Figure 2.6 on page 27, which reduces complexity compared to other applications.



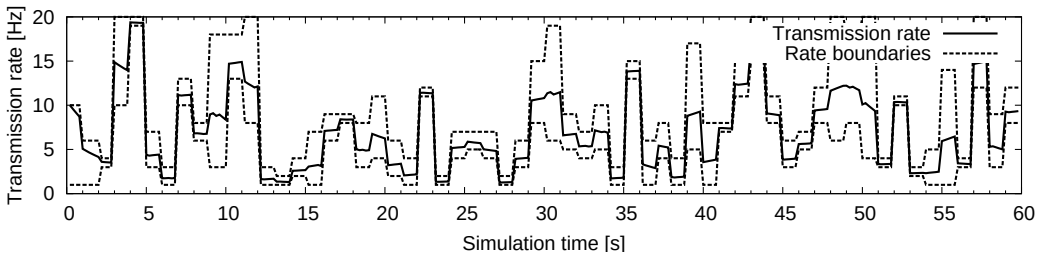
(a) Channel Busy Ratio (all nodes)



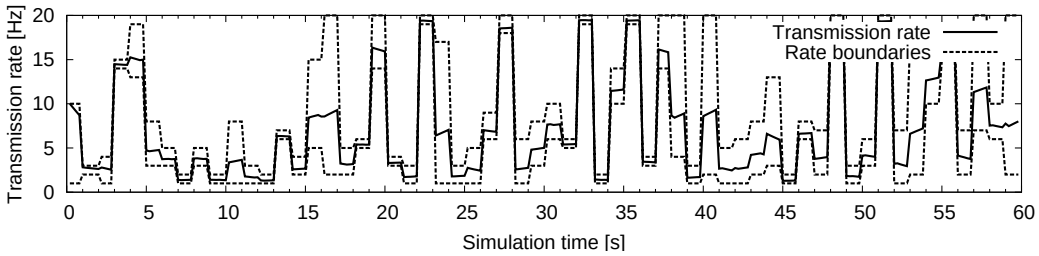
(b) Relative transmission rate (all nodes)



(c) Adaptation by exemplary node A



(d) Adaptation by exemplary node B



(e) Adaptation by exemplary node C

Figure 7.17: Stress test of dynamic change in minimum and maximum rate

to somewhat realistic changes in target distance and relative speed. For this purpose, we consider our implementation to be sufficient.

The ego vehicle optimizes its transmission parameters to communicate with the trailing vehicle, performing a recalculation every 50 ms. It sets its target distance d_t to the distance between itself and the trailing vehicle. If no trailing vehicle is known, it assumes the worst case of a vehicle approaching at the configured maximum speed $v_{max} = 200$ km/h and sets d_t to this imaginary vehicle's braking distance according to the lower part of Equation (4.1) on page 85 using $t_{rct} = 1.5$ s, $t_{sys} = 0.5$ s and $a = -5$ m/s².

The ego vehicle further calculates its transmission rate boundaries according to Equation (5.5) on page 127 with $d \in \{0.2\,\text{m}, 4\,\text{m}\}$, $k = 0.999$ and $p_{Rx} = 0.8$, where p_{Rx} corresponds to an assumed worst-case reception probability based on the results in Section 4.2.2 on page 89. Examples for corresponding IRD curves are depicted in Figure 5.11b on page 129. In the equation, v_{rel} is set to the *absolute* value of the *relative* speed between the ego vehicle and the trailing vehicle. That is, if the trailing vehicle decelerates and the relative speed becomes negative, we assume that it should get an update as well in order to stop warning the driver if necessary[3]. In the case that no trailing vehicle is known to the ego vehicle, v_{rel} is set to the difference between the ego vehicle's speed and the maximum speed v_{max}. During lane changes, the worst case situation is assumed to increase the probability of reception for the new trailing vehicle.

When the ego vehicle receives a message from the leading vehicle, it records how much the relative distance between itself and the leading vehicle has changed since the last message reception from this particular neighbor. This distance corresponds to the evaluated IRD.

7.3.2 Scenario and Setup

In order to achieve a high diversity in speed and inter-vehicle distance, we evaluate PULSAR in a scenario designed to resemble a German Autobahn with a forming traffic jam in one driving direction and a free-flow situation in the other driving direction. The corresponding vehicle movement traces were created using the traffic simulator VISSIM.

Figure 7.18 on the facing page illustrates the instantaneous speeds of the 815 vehicles in the scenario at simulation time 50 s with respect to their positions on the road. The snapshot is representative of the distribution of vehicle positions and speeds throughout the simulation time of 60 s. The traffic jam is located approximately between x-positions 500 m and 2200 m in the direction driving from left to right,

[3] This approach is based on the ideas expressed by An et al. in [24], where the authors argue that a change in the relative distance between is decisive to determine in which state the application is in, i.e., whether it has to give a warning or not. We chose this approach with the objective to address the state of the art in awareness control. However, our main focus is on the performance of PULSAR in terms of fairness and stability.

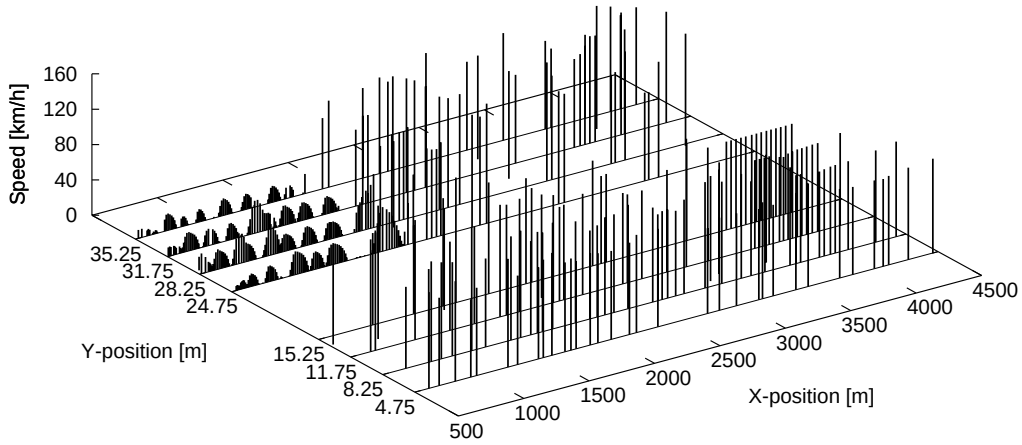


Figure 7.18: Vehicle speed distribution at simulation time 50 s in the traffic jam scenario

i.e., between y-positions 24 m and 36 m. Within the traffic jam, vehicles move in a stop-and-go fashion, creating a wave-like shape in the speed distribution. Their speeds lie between 0 km/h and 40 km/h, while the speeds of the vehicles in the free-flow environment lie between 90 km/h and 155 km/h.

In the following, we evaluate four variants of executing PULSAR, i.e.,

1. *Fixed* transmit power and *fixed* transmission rate boundaries,
2. *Fixed* transmit power and *dynamic* transmission rate boundaries,
3. *Dynamic* transmit power and *fixed* transmission rate boundaries, and
4. *Dynamic* transmit power and *dynamic* transmission rate boundaries.

In the fixed case, we set the transmit power to 20 dBm and the minimum and maximum transmission rates to 1 Hz and 20 Hz, respectively. In the dynamic case, we use a look-up table according to Figure 4.5a on page 94 to determine the transmit power optimizing reception performance at the vehicle's current target distance. Minimum and maximum transmission rates are calculated as described above based on the relative speed to the trailing vehicle. We additionally apply a global minimum rate of 1 Hz and a global maximum rate of 20 Hz as discussed in Section 5.4.1 on page 124. We thereby apply option (2), i.e., a vehicle's actual minimum and maximum transmission rates can lie within the interval [0 Hz, 100 Hz], while the global rate boundaries are used as cut-off values.

7.3.3 Convergence and Stability of PULSAR

In the following, we evaluate how PULSAR reacts to the dynamic adaptation of transmission rate boundaries and transmit power. Note that we can only cover a

small fraction of the evaluation possibilities of the relative rate adaptation approach. The following results are preliminary and represent the current state of development at the time of writing this thesis and are intended to show that PULSAR is able to maintain a stable channel load despite a reasonably realistic change in transmission rate boundaries and/or transmit power.

We parameterize PULSAR using t_{adp} = 200 ms, $\alpha_{I,rel}$ = 0.01 and $\beta_{D,rel}$ = 0.05. To ensure fast convergence from an overload situation, we use $\beta_{D,rel}$ = 0.2 if $U > 0.85$ and $\alpha_{I,rel}$ = 0.03 if $U < 0.3$. The latter values were chosen to ensure that the cumulative adaptation of all vehicles does not result in an oscillation between the two threshold CBR values.

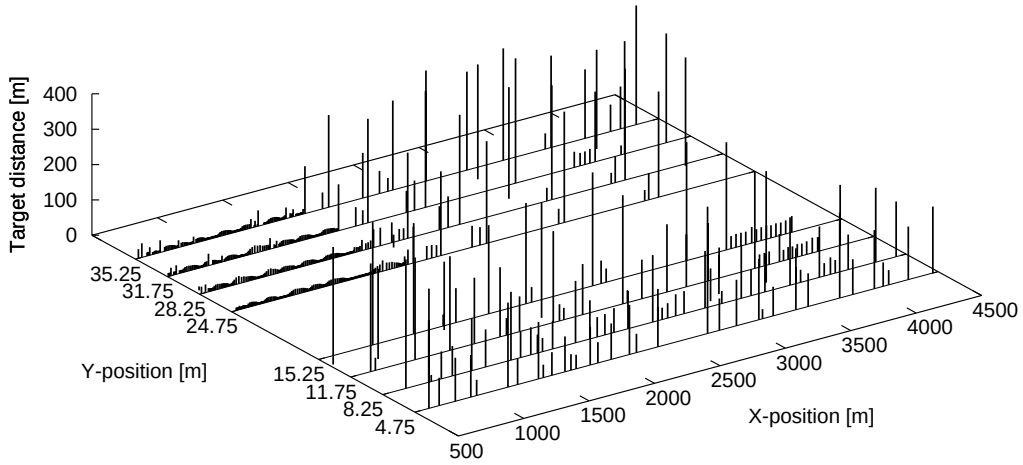
Discussion of a Representative Simulation Snapshot

To illustrate what the input to PULSAR looks like, we continue the discussion of the simulation snapshot shown in Figure 7.18 on the previous page. Figure 7.19a on the facing page illustrates the target distances calculated by each vehicle at simulation time 50 s, while Figure 7.19b shows the resulting transmit power values. We observe that typically the last vehicle within a group traveling in the same direction has a higher target distance and thus transmit power than the other members of the group. The lowest target distances occur within the traffic jam, while the highest target distances are not concentrated in a single area. As a result, vehicles within the traffic jam use a transmit power between 6 dBm and 10 dBm, while the other vehicles apply a value between 10 dBm and 26 dBm.

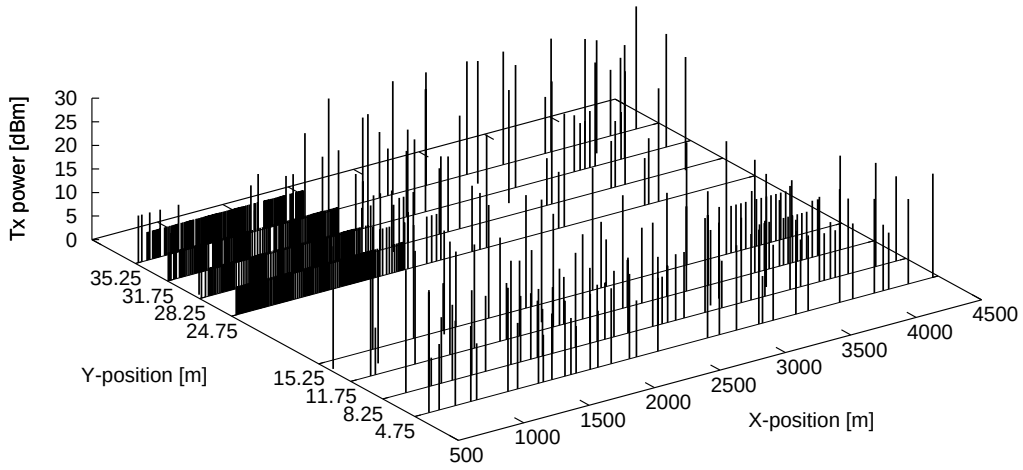
Figure 7.19c illustrates the relative speed each vehicle calculated based on its own speed and the one of its trailing vehicle. We observe that relative speeds are generally low and may also be negative if the trailing vehicle is slower than the ego vehicle. High relative speeds occur mostly at the end of the traffic jam and near the borders of the scenario where new vehicles arrive which do not have a trailing vehicle yet. In addition, one vehicle near x-position 3000 m is currently changing lanes and thus maximizing its transmit parameters in order to increase the reception probability for its new trailing vehicle. As a result of the assumed close relationship between relative speed and required transmission rate boundaries, the minimum transmission rates presented in Figure 7.19d show a similar pattern as the observed relative speeds. The corresponding maximum transmission rates are shown in Figure 7.19e. We observe a similar pattern, but a higher absolute value. Figure 7.19f illustrates the resulting *absolute* transmission rates according to the relative rate adaptation. Notice that no vehicle applies a higher value than the global maximum rate of 20 Hz.

Transmission Rate Adaptation of a Sample Vehicle

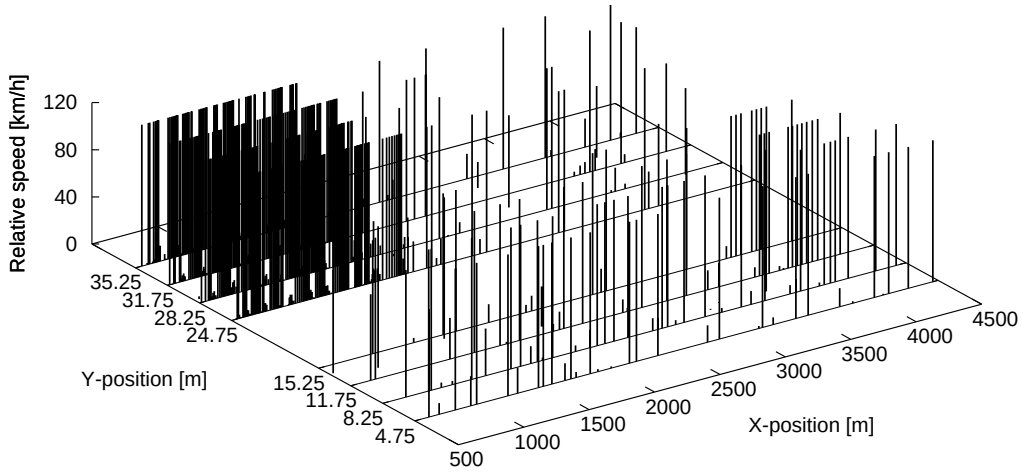
To give an example of what the rate adaptation looks like from an individual vehicle's perspective, Figure 7.20 on page 200 illustrates the transmission rate boundaries and absolute transmission rate with respect to simulation time for a vehicle traveling from



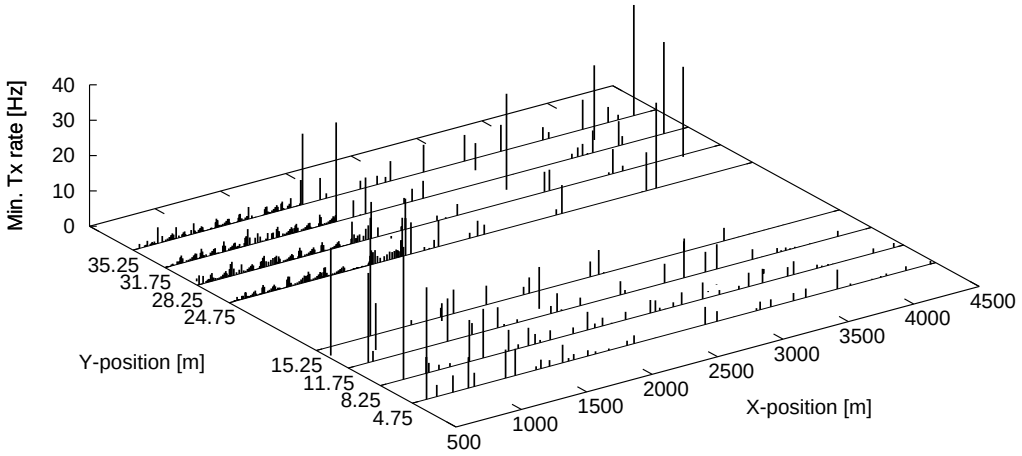
(a) Target distance



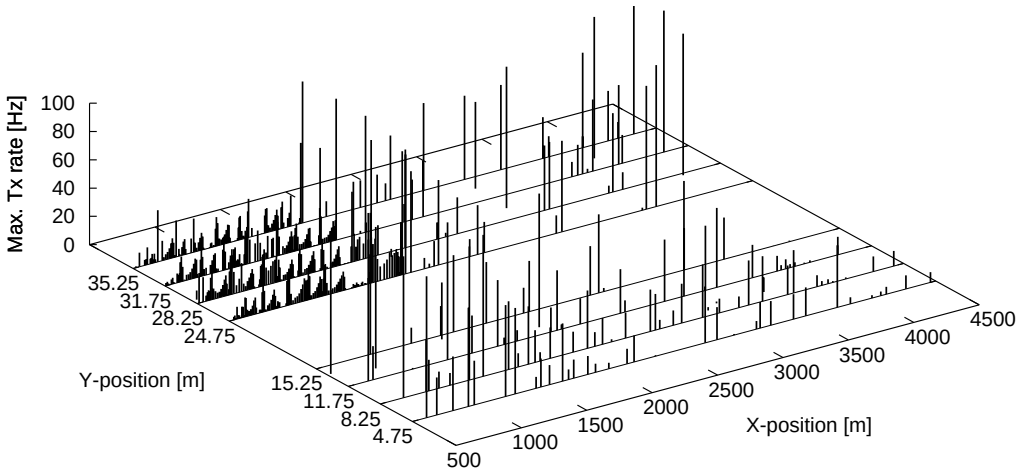
(b) Transmit power



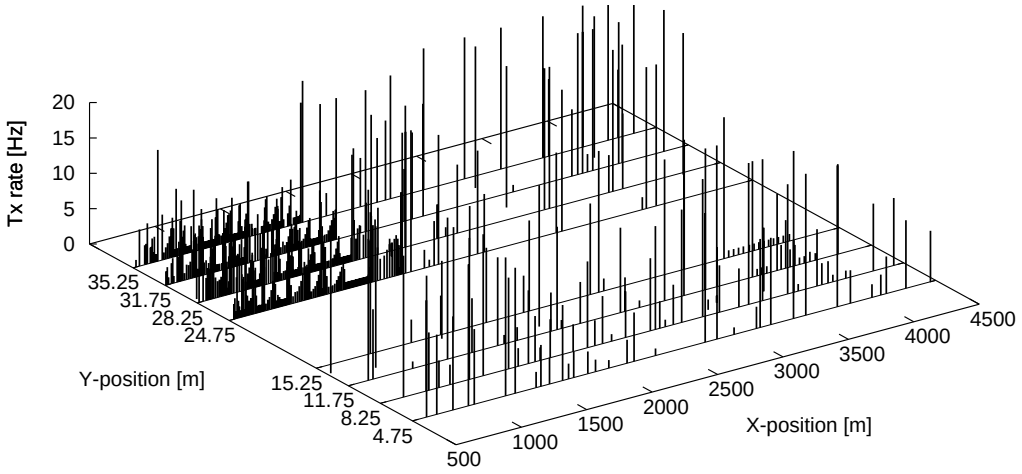
(c) Relative speed between ego vehicle and trailing vehicle



(d) Minimum transmission rate



(e) Maximum transmission rate (effectively capped at 20 Hz)



(f) Transmission rate (Case 4: Dynamic rate boundaries and dynamic transmit power)

Figure 7.19: Representative snapshot of PULSAR's performance in the traffic jam scenario

x-position 3200 m to x-position 1800 m over the course of the simulation. The vehicle approaches the traffic jam and reduces its speed from 110 km/h to 0 km/h between simulation times 42 s and 53 s. That is, it becomes the last vehicle in the traffic jam. Accordingly, we can see that its transmission rate boundaries as well as its actual transmission rate are increased. The latter, however, is capped at 20 Hz.

At simulation times 20 s and 42 s, the vehicle initiates lane changes which result in spikes in its rate adaptation. The reason for this behavior is that we configured our hypothetical application to assume worst-case conditions, i.e., a vehicle approaching from behind at 200 km/h, during lane change maneuvers.

Between simulation times 5 s and 22 s as well as between 28 s and 40 s, we can see that the vehicle adapts its transmission rate within the given rate boundaries according to the relative AIMD adaptation approach.

Relative Transmission Rate Adaptation

Figure 7.21 on page 201 illustrates the relative transmission rate and observed CBR value at simulation time 50 s, i.e., the snapshot considered before. The values shown in the figure are averages of 10 independent simulation runs. Each subfigure corresponds to one of the four considered cases as described above. What we would expect is that PULSAR has converged to a locally and globally fair allocation of the relative transmission rate. In Figures 7.21a and 7.21b, we observe that for fixed transmit power, PULSAR has converged to an S-shaped allocation of the relative transmission rate with reasonably close allocations for neighboring vehicles.

In Figures 7.21c and 7.21d on the other hand, the result appears to be more random. What happens here is that different transmit power levels imply a different participation range, cf. Section 5.3.3 on page 122. Thus, the aggregated CBR of one vehicle with a high transmit power may exceed the CBR threshold, while the aggregated CBR of its neighbor with a low transmit power does not. In fact, the target rate mechanism as described before may create an unfair (relative) transmission rate allocation with non-uniform transmission power, since it pulls up the transmission rates of high transmit power vehicles and pulls down those of low transmit power vehicles. However, compared to the execution of AIMD without target rate, the outcome in terms of IRD was still slightly better when the target rate mechanism was applied.

The example illustrates that a further development of the local fairness principle as well as of the target rate mechanism is required to account for different transmit power levels. In future work, the following approach could be investigated: Reduce the granularity of the applicable transmit power levels, e.g., to “high”, “medium” and “low”. Then, create one target rate each for every transmit power level and have nodes adapt their transmission rates towards the target rate corresponding to their respective transmit power.

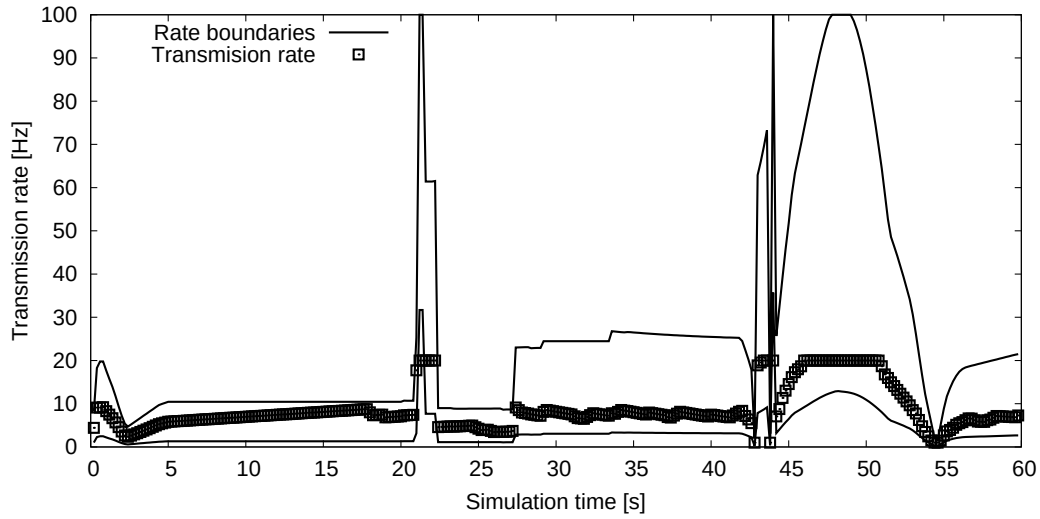


Figure 7.20: Transmission rate and rate boundaries of an exemplary vehicle applying relative rate adaptation in the traffic jam scenario; the vehicle approaches the traffic jam and performs lane changes at 20 s and 42 s

Congestion Control

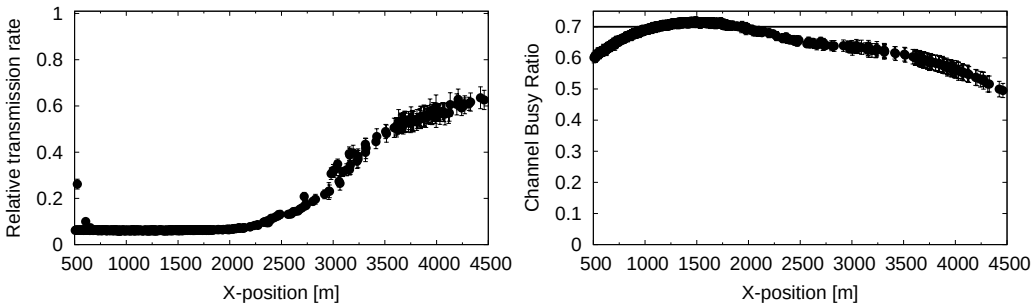
The CBR allocations in Figure 7.21 on the facing page indicate that PULSAR has converged to a channel load at or below the configured CBR target of 0.7 CBR. Figure 7.22 on page 202 illustrates the development of the CBR with respect to the x-position on the road over the course of an exemplary simulation run. The figure is based on case 4, i.e., dynamic rate boundaries and a dynamic transmit power. To help distinguish overload from underload, the figure contains a plane at 0.7 CBR. We can see that only at the beginning of the simulation, the CBR is significantly above the target value. After that, the CBR does not exceed 0.75 throughout the scenario.

Conclusion

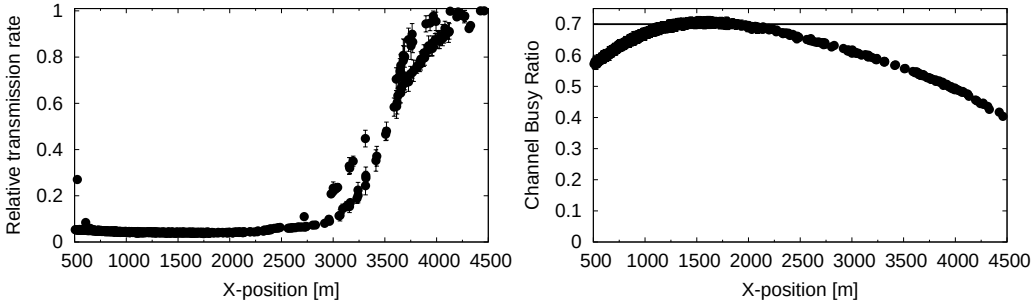
Based on the results presented in this section, we conclude that PULSAR is able to fulfill its main objective, i.e., to maintain a stable channel load, even when subject to frequent and partially radical changes in transmission rate and transmit power. For a homogeneous transmit power, PULSAR is further able to maintain local and global fairness. However, future work is required to optimize fairness if Transmit Power Control (TPC) is applied.

7.3.4 Impact on the Observed Inter-Reception Distance

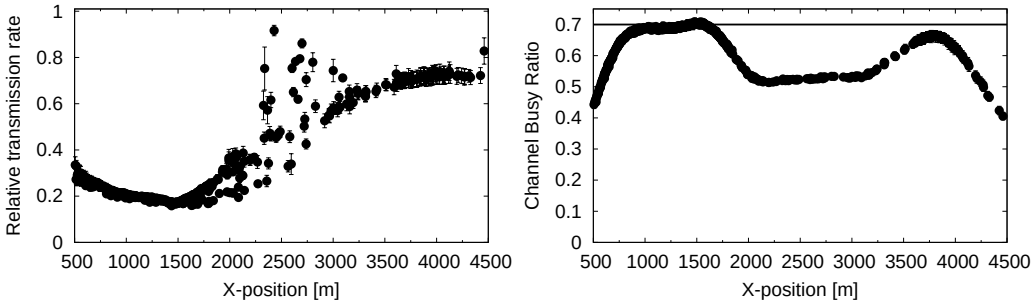
We conclude this section by presenting the distribution of the observed IRDs, i.e., the changes in the relative distance between ego vehicle and trailing vehicle between the reception of two successive messages. We emphasize that it is not the intention



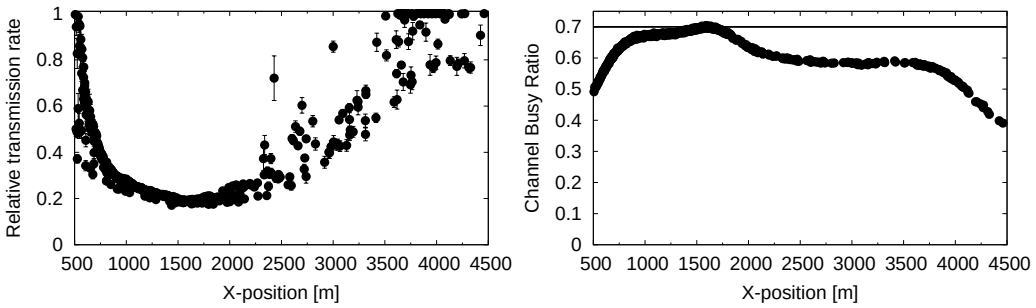
(a) Case 1: *Fixed* transmit power and *fixed* rate boundaries



(b) Case 2: *Fixed* transmit power and *dynamic* rate boundaries



(c) Case 3: *Dynamic* transmit power and *fixed* rate boundaries



(d) Case 4: *Dynamic* transmit power and *dynamic* rate boundaries

Figure 7.21: Relative transmission rate and resulting CBR at simulation time 50 s in the traffic jam scenario

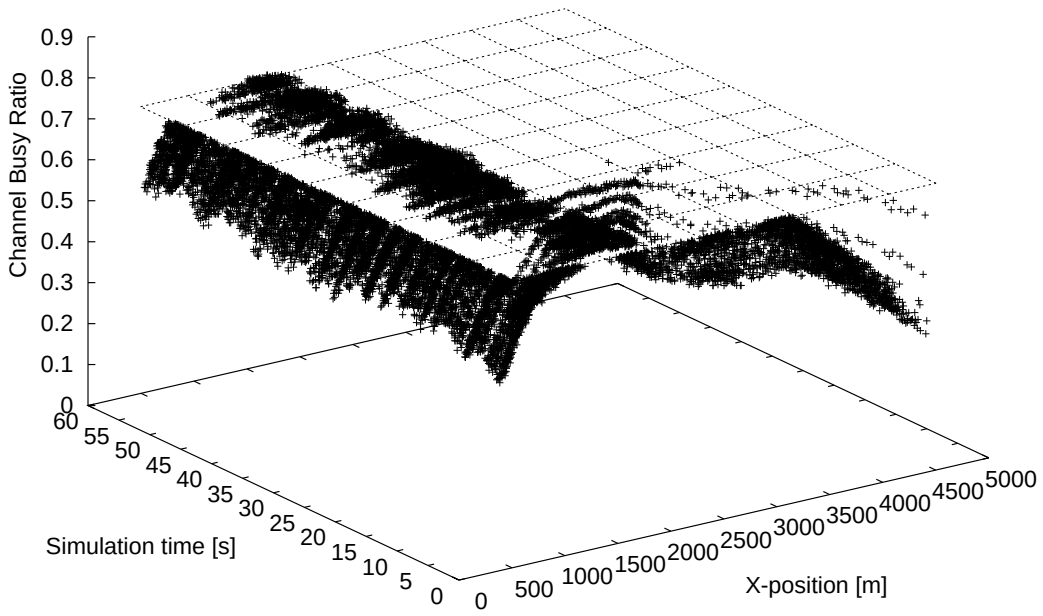
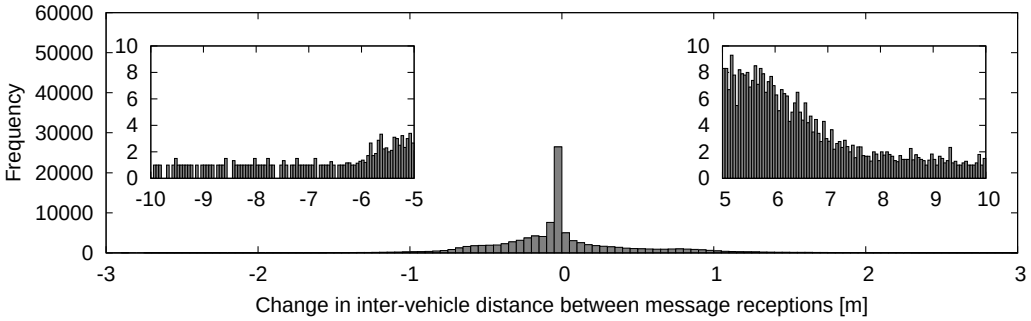


Figure 7.22: Observed CBR over the course of a simulation using dynamic rate boundaries and transmit power values in the traffic jam scenario

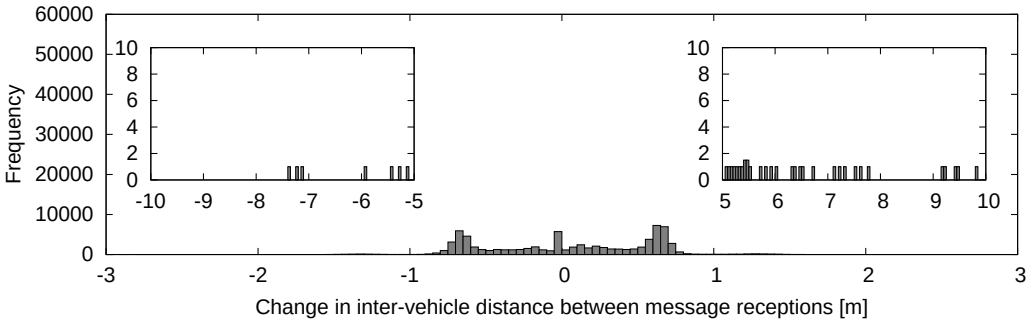
of this evaluation to draw any conclusions as far as the functionality of an actual safety application is concerned. Rather, we would like to demonstrate that the relative transmission rate approach can *potentially* make a difference from a safety applications' perspective. Further work is required to evaluate the impact of the relative transmission rate approach on an *actual* safety application.

Figure 7.23 on the facing page illustrates histograms of the observed IRDs for each of the four previously discussed combinations of fixed and dynamic transmission rate boundaries and transmit power levels. The values shown in the figure correspond to the average of ten independent simulation runs. In Figure 7.23a, we can see that the conventional rate control approach of using a fixed transmit power and adapting transmission rate based on channel load results in significant number of IRDs beyond ± 5 m. Using the terminology and line of reasoning of An et al. [24], we interpret large positive values as potential sources for false negative warnings, i.e., no warning despite an imminent danger, while large negative values might result in a false positive warning, i.e., a warning even though no danger exists. Also note that a large fraction of the transmitted messages resulted in an IRD near 0 m.

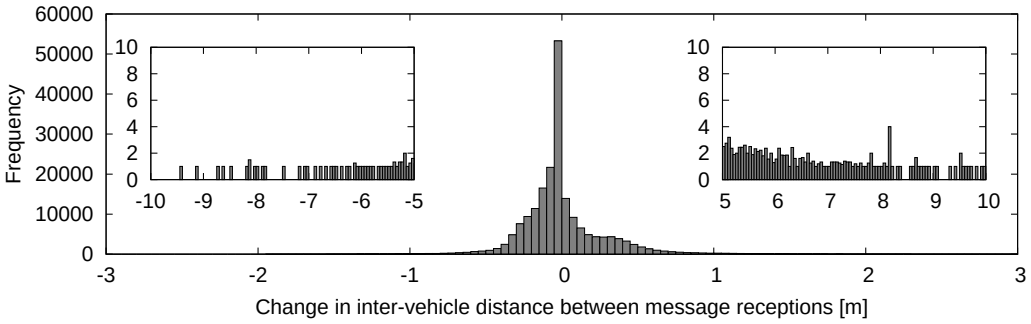
In Figure 7.23c, we can see that the approach using a dynamic transmit power and fixed rate boundaries results in a reduced number of extreme values at both ends. We account this observation to a higher overall transmission rate. Note, however, that still, a high number of messages is transmitted resulting in an IRD near 0 m.



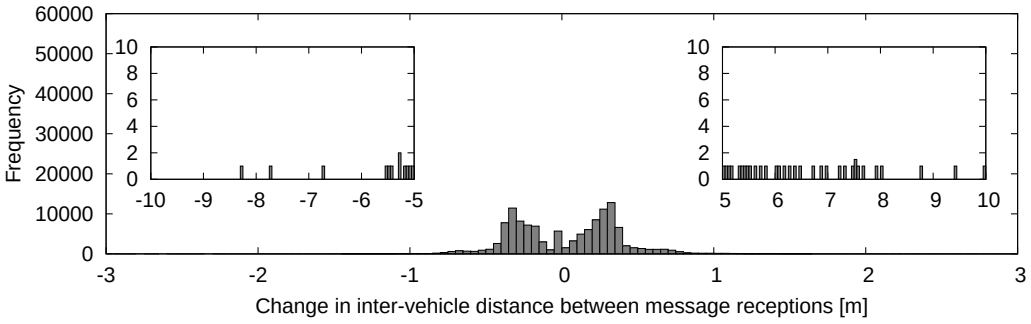
(a) Case 1: *Fixed* transmit power and *fixed* rate boundaries



(b) Case 2: *Fixed* transmit power and *dynamic* rate boundaries



(c) Case 3: *Dynamic* transmit power and *fixed* rate boundaries



(d) Case 4: *Dynamic* transmit power and *dynamic* rate boundaries

Figure 7.23: Distribution of observed IRDs in the traffic jam scenario

Figures 7.23b and 7.23d illustrate that the combination of a fixed transmit power and dynamic rate boundaries results in an overall reduced number of transmissions. This can be mostly attributed to the fact that significantly less messages are transmitted which result in an IRD near 0 m[4]. In addition, we observe that the number of extreme values is reduced at both ends of the scale compared to the combinations using fixed transmission rate boundaries. Between the two combinations with dynamic rate boundaries, we can see that an additional adaptation of the transmit power results in less samples beyond the configured minimum requirement of 0.5 m IRD.

Table 7.2 on the next page summarizes the discussion above by presenting the 0.1th, 1st, 99th and 99.9th percentiles of the IRDs observed in each of the four considered configurations. We can see that using the relative transmission rate concept, the 99.9th percentile can be reduced by a factor of approximately 4, while the 0.1th percentile can be reduced by a factor of 3.3 compared to the conventional congestion control approach using fixed transmit power and fixed transmission rate boundaries.

To conclude, the evaluation of the observed IRDs indicates that the suggested dynamic adaptation of transmit power and transmission rate boundaries depending on the driving context has the *potential* to make congestion control more compatible with the requirements of safety applications. Note, however, that this approach requires a careful choice of the input parameters. For example, if an ego vehicle optimizes its transmission parameters for the wrong following vehicle, the true following vehicle might detect the ego vehicle too late. We excluded this case in our study by applying global knowledge from the simulator. In addition, if the minimum transmission rates are chosen too high, the rate control mechanism can become ineffective. In future work, the interrelationship of congestion control and different safety applications could be further studied, especially taking into account random influences like GPS errors and the resulting tracking error of neighboring vehicles.

7.4 Comparison With Other Approaches

In this section, we compare PULSAR's performance (using uniform transmission rate boundaries and transmit power) against other rate control approaches which share similar objectives and protocol elements. Please note that the results presented here are based on our interpretation of the algorithm descriptions in the respective papers. While we have compared our results to the original authors' results and found a reasonable match, we cannot rule out the possibility that we did not implement all details as intended by the original authors.

[4] This result illustrates the similarity of the proposed relative rate adaptation based on IRD and the approach of Huang et al. [23] whose objective it is to reduce transmissions which lead to a *tracking error* close to zero.

	Tx power	Rate boundaries	$P_{0.1}$ [m]	P_1 [m]	P_{99} [m]	$P_{99.9}$ [m]
1	Fixed	Fixed	-4.20	-1.30	2.75	6.95
2	Fixed	Dynamic	-1.50	-0.95	1.30	2.00
3	Dynamic	Fixed	-2.05	-0.75	1.35	3.45
4	Dynamic	Dynamic	-1.25	-0.70	0.90	1.70

Table 7.2: k^{th} percentiles P_k of Inter-Reception Distance (IRD) observed in the traffic jam scenario

7.4.1 Khorakhun et al.

In [191], Khorakhun et al. introduce a congestion control algorithm which compares a node's CBR measurement against a threshold and adapts the transmission *interval*, i.e., the inverse of the transmission rate, using Additive Increase Additive Decrease (AIAD), cf. Section 3.1 on page 46. That is, when a node transmits a packet it calculates its new transmission interval t_{Tx} at iteration k as

$$t_{Tx,k} = \begin{cases} t_{Tx,k-1} - \beta & \text{if } U < U_{th} \\ t_{Tx,k-1} + \beta & \text{if } U > U_{th} \end{cases} \tag{7.5}$$

where U is the node's current CBR measurement, evaluated over the time since its last transmission, and U_{th} is a CBR threshold. For fairness, a node is only allowed to decrease its transmission interval if it is above the average transmission interval of its neighbors.

The approach by Khorakhun et al. shares some common aspects with PULSAR, since it is based on a binary control using CBR measurements as a feedback. In contrast to PULSAR, however, it adapts the transmission *interval* rather than the transmission *rate* and it uses AIAD rather than AIMD. In addition, the adaptation is performed asynchronously after each transmission rather than in fixed time intervals, which also affects the way the CBR is calculated. While the exchange of transmission interval information and the calculation of an average transmission interval is similar to the target rate mechanism in PULSAR, the obtained average is used differently in the adaptation. Finally, PULSAR shares and aggregates CBR measurements over two hops, while the approach by Khorakhun et al. is based on local measurements.

Figure 7.24 on the next page illustrates the convergence of the protocol by Khorakhun et al. on a linear road with static vehicles after 180 s simulation time in the same configuration as PULSAR in Figure 6.13 on page 155, using β = 20 ms in Equation (7.5). Note that while the protocol adapts the transmission interval, we use the transmission rate in the visualization to facilitate a comparison with PULSAR. Comparing Figures 7.24 and 6.13d, we observe that the transmission rate allocation of the approach by Khorakhun et al. resembles closely the result obtained by PULSAR using local CBR measurements in combination with the target rate mechanism. Like PULSAR

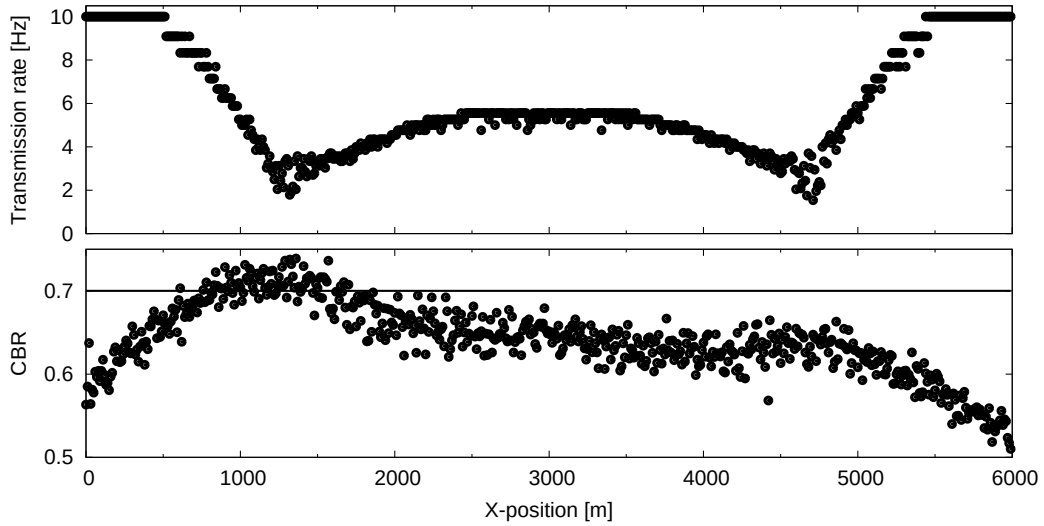


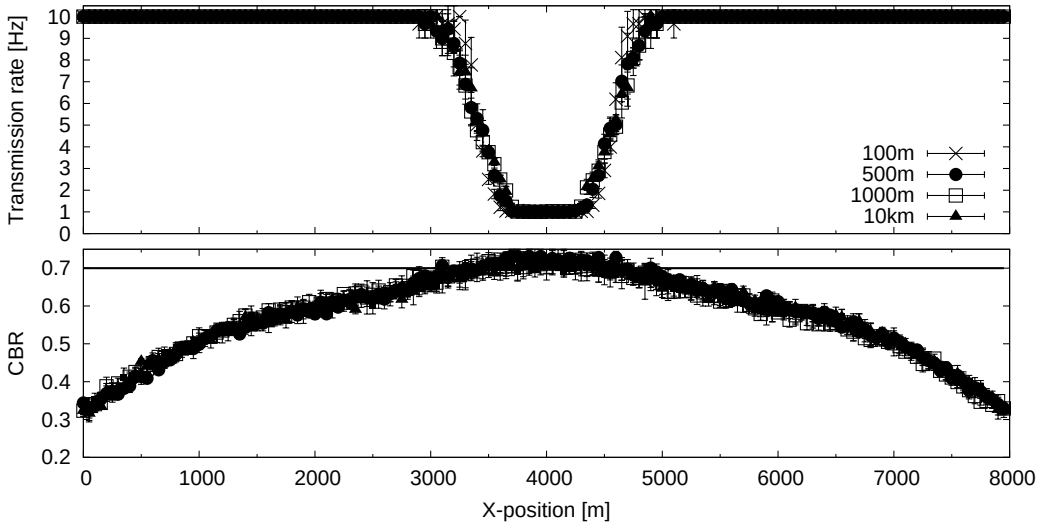
Figure 7.24: Convergence of the approach by Khorakhun et al. [191] after 180 s in the linear road scenario

without the exchange of CBR information, the approach by Khorakhun et al. is not able to converge to global fairness.

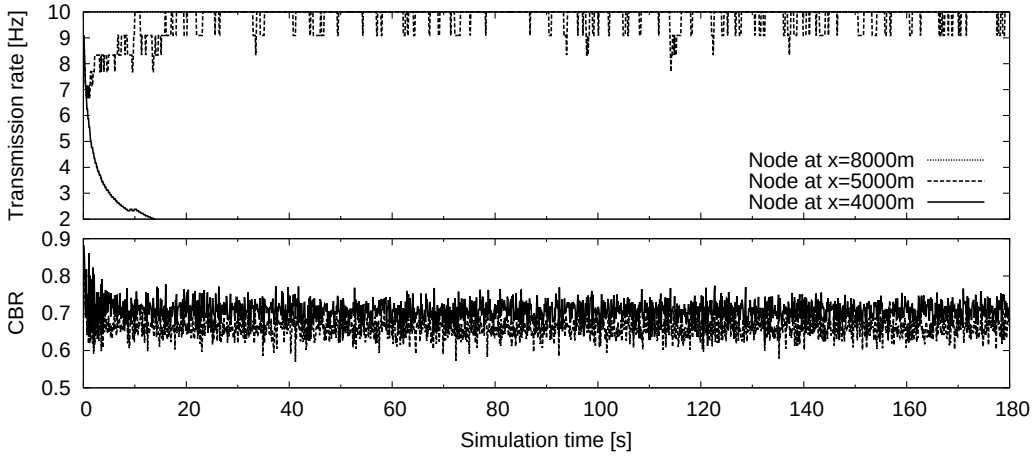
Figure 7.25 on the next page illustrates the spatial and temporal convergence of the considered protocol in the static crossing scenario used earlier in this chapter. We observe that the nodes near the center of the intersection, i.e., the most congested location, reduce their transmission rates to the minimum rate of 1 Hz. In other words, the global fairness principle is not fulfilled. Since in [191], the distance d_{neigh} from which neighborhood information is taken into account is not explicitly defined, we evaluated the impact of different values. However, as shown in Figure 7.25a, there is no significant difference in terms of global fairness. Note that while not explicitly shown here, the convergence of the protocol by Khorakhun et al. again resembles the one of PULSAR with local CBR measurements and target rate mechanism.

7.4.2 SOURC

In [202], Busche et al. introduce SOURC, an extended version of the approach by Khorakhun et al. [191], which includes a two-hop information sharing mechanism for CBR measurements as well as two CBR thresholds U_{min} and U_{max}. Each node calculates the maximum CBR received from its neighbors within its communication range, denoted as *local CBR* U_{local} as well as the *global CBR* U_{global}, which denotes the maximum received U_{local}. Based on this information as well as the node's own CBR measurement U_{self} and the average transmit interval $t_{Tx,avg}$ of the node's neighbors, SOURC adapts the transmission interval as illustrated by Algorithm 3 on page 208. In a nutshell, all nodes detecting congestion at their own location increase their transmission intervals, while only those with the highest transmission interval may



(a) Spatial convergence after 300 s



(b) Temporal convergence (d_{neigh} = 500 m)

Figure 7.25: Convergence of the approach by Khorakhun et al. [191] in the static crossing scenario

decrease it, given that no two-hop neighbor has reported congestion. All other nodes adjust their transmission intervals implicitly by adapting to the neighbor average.

Like PULSAR, SOURC is a binary scheme controlling the number of generated messages by using CBR measurements as a feedback. Since SOURC is an extended version of the approach by Khorakhun et al., the similarities and differences discussed in Section 7.4.1 apply to SOURC as well. In addition, SOURC employs a two-hop piggybacking scheme for the dissemination of CBR measurements which is very similar to PULSAR's. Like PULSAR, SOURC uses the maximum norm when aggregating CBR information. However, instead of aggregating the node's own measurement and CBR information received from neighbors into one value like PULSAR, SOURC uses a different CBR threshold for each type of feedback.

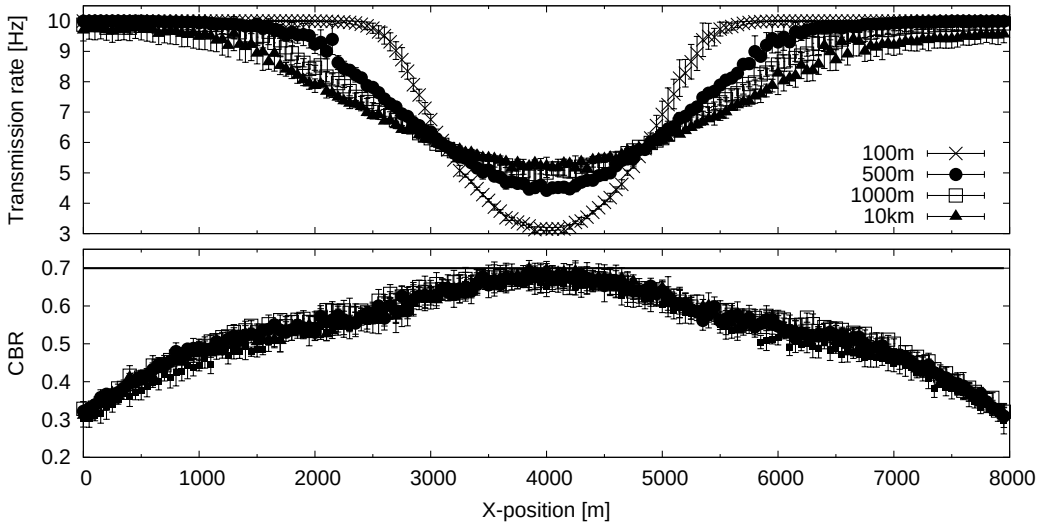
Algorithm 3: SOURC [202]

Data: U_{self}, U_{global}, $t_{Tx,k-1}$
Result: $t_{Tx,k}$
if $U_{self} > U_{max}$ **then**
  $t_{Tx,k} \leftarrow t_{Tx,k-1} + \beta$
else
  if $U_{global} < U_{min}$ **then**
    if $t_{Tx,k-1} \geq t_{Tx,avg}$ **then**
      $t_{Tx,k} \leftarrow t_{Tx,k-1} - \beta$
    else
      $t_{Tx,k} \leftarrow t_{Tx,avg}$
    end
  else
    $t_{Tx,k} \leftarrow t_{Tx,avg}$
  end
end

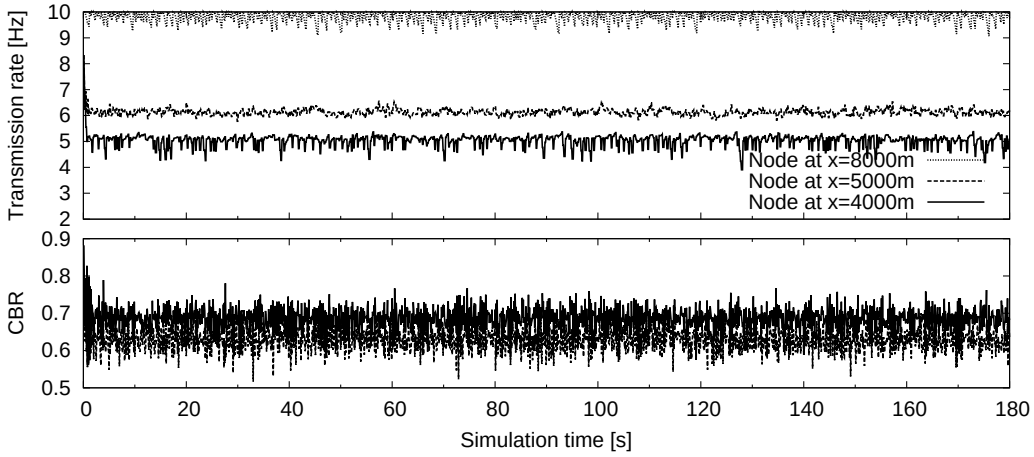
Figure 7.26 on the next page illustrates the convergence behavior of SOURC in the static crossing scenario discussed earlier. For the evaluation, we set β = 20 ms, U_{min} = 0.65 and U_{max} = 0.7. Another important parameter is the *neighbor distance* d_{neigh}, i.e., the distance from within which the protocol takes into account neighborhood information in terms of CBR values and transmission intervals. In [202], the authors state that SOURC aggregates information from all vehicles "inside communication range", but do not define this distance explicitly.

Figure 7.26a illustrates that the choice of d_{neigh} significantly influences SOURC's convergence behavior. A higher value of d_{neigh} benefits global fairness, as we can tell from the transmission rates of the nodes near the center of the intersection at x-position 4000 m. However, with increased d_{neigh} the distance at which nodes decrease their transmission rates due to congestion increases as well. In fact, since in many cases, nodes set their transmission interval to the average value of their neighbors, there is no actual boundary for the participation in congestion control.

Figure 7.26b illustrates the unbounded participation in congestion control using the transmission rates of three exemplary nodes with respect to simulation time. We can see that the node at x-position 8000 m adapts its transmission rate even though its contribution to the congestion in the intersection center is negligible if measurable at all, cf. the carrier sense probability in Figure 7.10 on page 183. While the figure is based on d_{neigh} = 1000 m, the same effect can be observed for the other evaluated neighbor distances as well, albeit with a different amplitude of the oscillations. In contrast, PULSAR has a clear boundary for the participation in congestion control, as illustrated by Figure 7.6b on page 177. As a result, nodes located near the border of the considered scenario do not adapt their transmission rates at all once a steady state has been reached.



(a) Spatial convergence after 300 s



(b) Temporal convergence (d_{neigh} = 1000 m)

Figure 7.26: Convergence of SOURC [202] in the static crossing scenario

Figure 7.27 on the following page illustrates the transmission rate and measured CBR of an exemplary vehicle traversing the previously considered highway scenario in which two groups of vehicles pass each other. Note that since SOURC is executed asynchronously, the figure shows snapshots of the current transmission rate taken every 100 ms. The corresponding CBR values were calculated over 100 ms intervals and are not equivalent to the ones used in the adaptation. However, they reveal oscillations in the system, as we discuss in the following.

By comparing Figure 7.27 with Figure 7.16 on page 192, we observe that SOURC results in less oscillations in the transmission rate than PULSAR, while achieving a similar degree of global fairness. On the other hand, SOURC results in more irregular oscillations in the observed CBR than PULSAR. Unlike PULSAR, SOURC does not

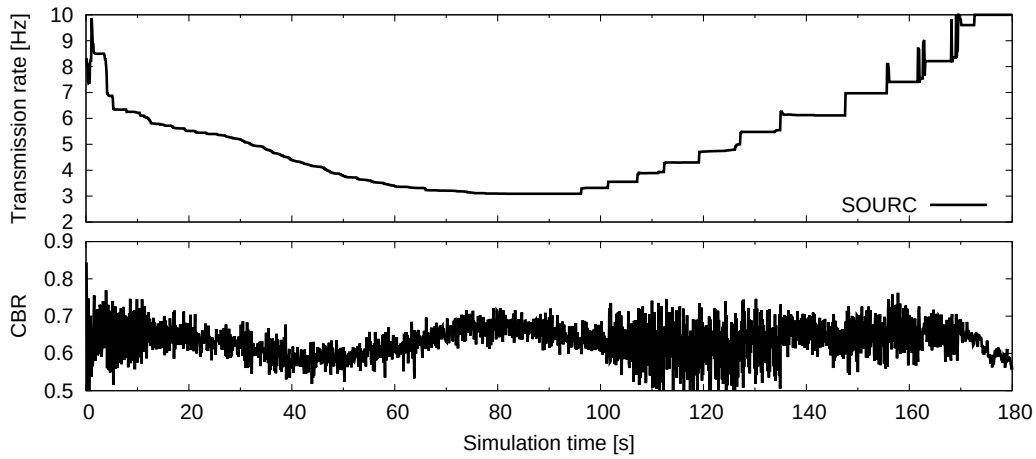


Figure 7.27: Adaptation by an exemplary node running SOURC [202] in the scenario where two groups of vehicles pass each other

contain a message rescheduling mechanism, whose lack of may lead to instabilities in the system as discussed in Section 6.3.1 on page 142.

To further investigate this issue, Figure 7.28 on the next page illustrates the convergence of SOURC in the previously considered scenario where 200 nodes share the same location. Like before, the figure shows snapshots of 100 ms of an individual node's transmission rate and the CBR evaluated since the last snapshot. We observe that while the node increases its transmission interval, the oscillations in the CBR increase as well. The node increases its transmission rate up to the maximum of 10 Hz. At this point, however, the CBR is not increased any further, which indicates the occurrence of synchronized packet collisions, similarly to Figure 6.7a on page 146. After a few seconds, however, SOURC stabilizes itself and reduces its transmission rate to approximately 6.5 Hz. At this point, the oscillations in the CBR become much more regular than before.

Another aspect that can be observed in Figure 7.28 is that it takes SOURC a very long time to converge to efficiency, i.e., to fill up the channel up to the desired CBR target of 0.7. For the first 20 s of simulation time, there is hardly any increase in transmission rate despite an underutilization of the channel. After that, the transmission rate increases exponentially. We attribute this behavior to two factors. First, there is an inversely proportional relationship between transmission interval and CBR, cf. Figure 5.13a on page 131. Second, SOURC couples transmission interval and adaptation interval. Thus, a long transmission interval additionally slows down the adaptation process, which contributes to the long convergence time in Figure 7.28.

Despite their apparent similarity, the evaluation shows that there are significant differences in the convergence behavior of SOURC and PULSAR. While SOURC has the advantage that it can be executed asynchronously, it appears to have difficulties in converging to efficiency and may result in an instability of the system due to the lack of a message rescheduling mechanism. In addition, it does not fulfill the participation

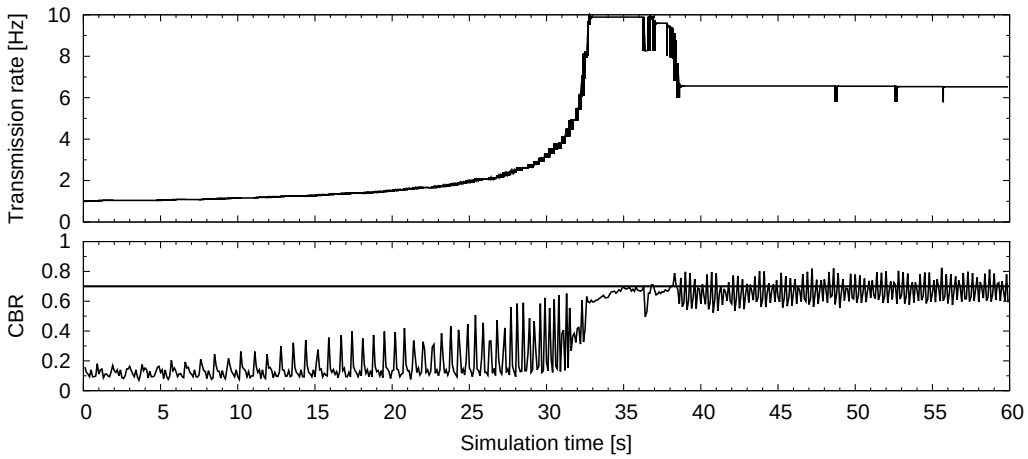


Figure 7.28: Convergence of SOURC [202] in scenario where 200 nodes share the same location (exemplary node)

fairness principle, since congestion information theoretically propagates indefinitely when nodes set their transmission interval to the neighbor average.

In future work, a hybrid protocol could be developed which combines the strengths of both approaches. Next to employing two-hop piggybacking of CBR information as found in both approaches, we envision such an approach to use the concept of two CBR thresholds and the asynchronous execution from SOURC and to combine it with the concept of participation distance, the homogeneous adaptation interval, the rescheduling mechanism and the adaptation of the transmission rate (rather than the transmission interval) from PULSAR.

7.4.3 LIMERIC

In [71], Kenney et al. introduce LIMERIC, a linear rate control algorithm based on CBR measurements. In contrast to the binary rate control employed in PULSAR and SOURC, LIMERIC uses the *difference* between the measured CBR U and the targeted CBR U_t to adapt a node's transmission rate r as

$$r_k = (1 - \alpha) r_{k-1} + \beta (U_t - U) \tag{7.6}$$

where α and β are parameters whose given default values guarantee stability for up to 285 vehicles. To prevent instability if more vehicles are present, LIMERIC limits the size of an increment or decrement to a fixed value with a default of 1 Hz.

Since LIMERIC was designed to improve the convergence of binary control schemes such as AIMD, it is of interest in the context of this thesis to evaluate how LIMERIC performs in the PULSAR framework and how its convergence compares to AIMD. Note that we thereby replace AIMD by LIMERIC without making any further changes

to the PULSAR protocol. We evaluate two configurations of LIMERIC corresponding to a different maximum number of vehicles sharing the wireless medium:

(1) $\alpha = 0.1$ and $\beta = 1/150 \cdot 1712 = 11.413$, which is the default configuration of LIMERIC according to [53], adapted to the message size used in this thesis. With a duration of 584 μs per message, 1712 messages correspond to a CBR of 100%, cf. Figure 5.4a on page 113. This configuration guarantees stability for up to 285 vehicles [53], which is also true for a delay of one adaptation interval according to [219].

(2) $\alpha = 0.1$ and $\beta = 1/240 \cdot 1712 = 7.133$, which according to Inequality (6) in [53] and Equation (43) in [219] guarantees stability for up to 455 vehicles.

To facilitate a comparison with the results for AIMD, we apply the same parameters as before, i.e., a transmit power of 20 dBm and an adaptation interval of 200 ms. In other words, the protocol as described in Chapter 6 remains unchanged except for the increase/decrease algorithm, including the synchronization of the execution.

Figure 7.29 on the next page illustrates the convergence of LIMERIC using configuration (1) and local CBR measurements as a feedback in a scenario where 200 nodes share the same location. Compared to the performance of AIMD in Figure 7.2 on page 171, we can see that LIMERIC converges significantly faster and shows a smaller oscillation size. Note that LIMERIC does not converge to the target CBR of 0.7, but stays slightly below it. The gap between the target CBR and the convergence result depends on the number of vehicles sharing the channel and the number of vehicles for which α and β were chosen, which is why these parameters can't simply be selected to accommodate, e.g., 1000 vehicles. In this case, LIMERIC would converge to a CBR of approximately 0.55 in the 200 node scenario.

Figure 7.30a on page 214 depicts the convergence of LIMERIC using configuration (1) after 300 s simulation time in the static crossing scenario considered earlier. We can see that with global knowledge, LIMERIC converges to a similar shape as water filling. The transmission rate in the center of the intersection is 5.38 Hz, which is approximately 10% lower than the water filling result of 5.9 Hz. With two-hop piggybacking, LIMERIC converges to 4.84 Hz at the center of the intersection, which is slightly higher than the result of 4.54 Hz observed using AIMD (with target rate mechanism) at the same adaptation interval of 200 ms. Finally, with one-hop piggybacking and local measurements, the minimum transmission rate in the scenario is 3.68 Hz and 2.56 Hz, respectively, compared to 1 Hz using AIMD. The example illustrates that the linear control of LIMERIC makes more use of the available CBR information than the binary control of AIMD. However, our results indicate that LIMERIC, like AIMD, needs two-hop feedback in order to converge toward global fairness under the assumed conditions.

Figure 7.30b shows that with two-hop piggybacking, the three exemplary nodes located at x-positions 4000 m, 4500 m and 5000 m converge to a similar rate near

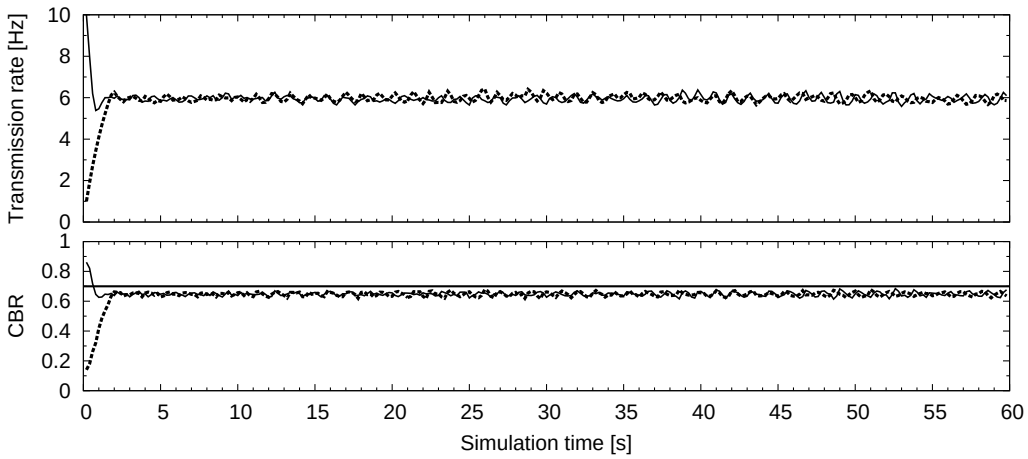
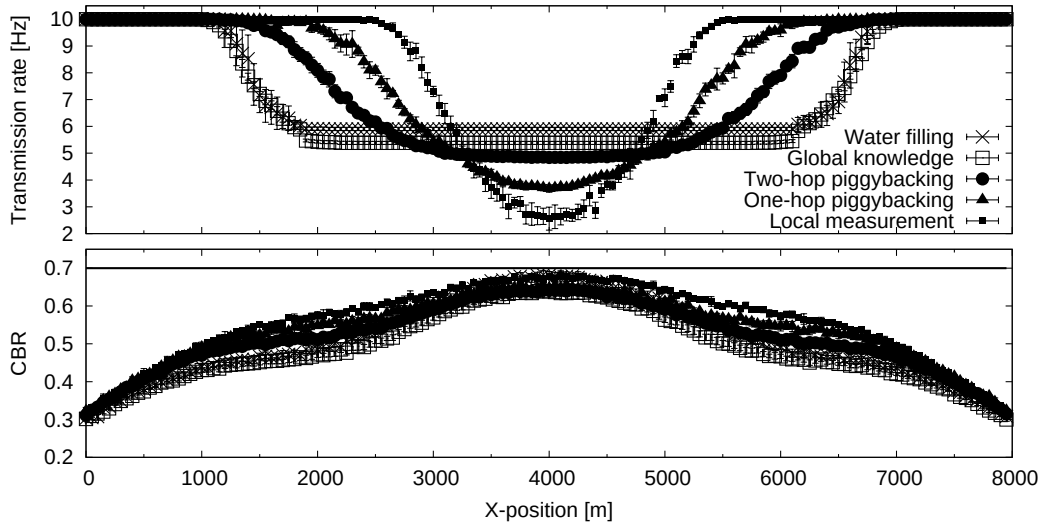


Figure 7.29: Convergence of LIMERIC [53] in a scenario where 200 nodes share the same location (exemplary node, using configuration (1))

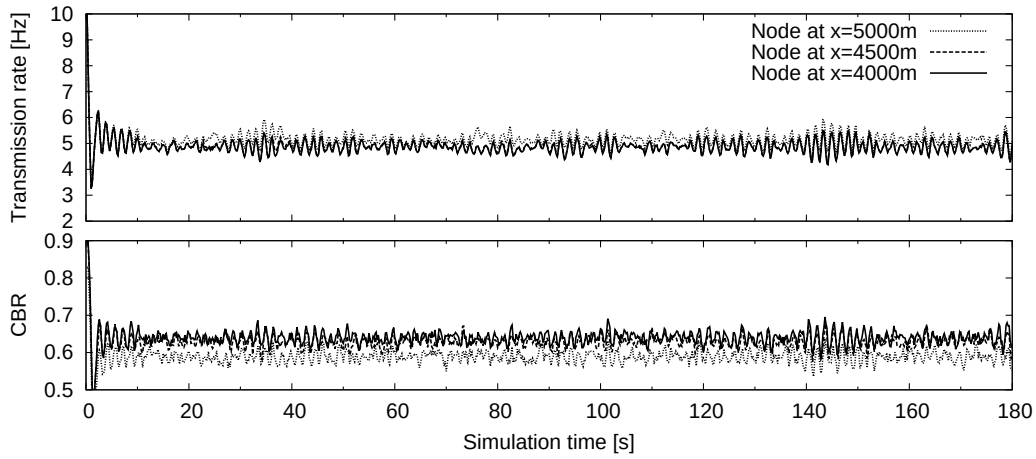
5 Hz. In general, the observed oscillations are lower than with AIMD, but they occur more in clusters.

So far, we focused on static scenarios only which excluded vehicle movement. To study its convergence in a mobile scenario, we additionally evaluated LIMERIC in the highway scenario where two groups of vehicles pass each other. Figure 7.31a on page 215 illustrates the transmission rate and measured CBR of our previously considered example vehicle traversing the scenario, now running LIMERIC in configuration (1). Compared to AIMD with target rate in Figure 7.16 on page 192, we observe significantly higher oscillations of both the transmission rate as well as of the observed CBR when the two groups start influencing each other. Near simulation time 80 s, the transmission rate of our sample vehicle oscillates between approximately 1 Hz and 5 Hz, while the observed CBR oscillates between approximately 0.4 and 0.8. In contrast, LIMERIC shows minimal oscillations when using global knowledge from the simulator.

According to our understanding, the observed oscillations result from the following factors. First, the total number of 451 vehicles exceeds the stability criterion of configuration (1). Figure 7.31c illustrates that with configuration (2), the oscillations are significantly reduced. Second, the oscillation size can be reduced if the local measurement of the node is *not* delayed by one adaptation interval in the CBR aggregation process, despite the bounce-back effect described in Section 6.4.4 on page 160. The result can be observed in Figure 7.31c, where the combination of configuration (2) and the immediate reaction to the local CBR measurement reduces oscillation size significantly. However, the bounce-back effect may impact LIMERIC's ability to converge to global fairness, as shown in Figure 7.31b, which compares the convergence using configuration (1) without the delay of local information to the outcome using global knowledge. We can see that while the oscillation size is reduced compared to Figure 7.31a, LIMERIC now oscillates above or below the outcome of global knowl-



(a) Spatial convergence after 300 s

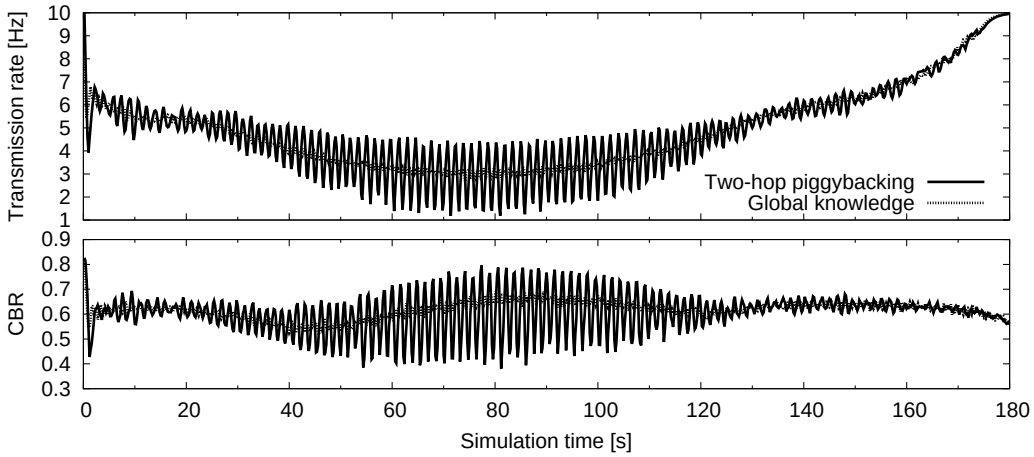


(b) Temporal convergence (with two-hop piggybacking)

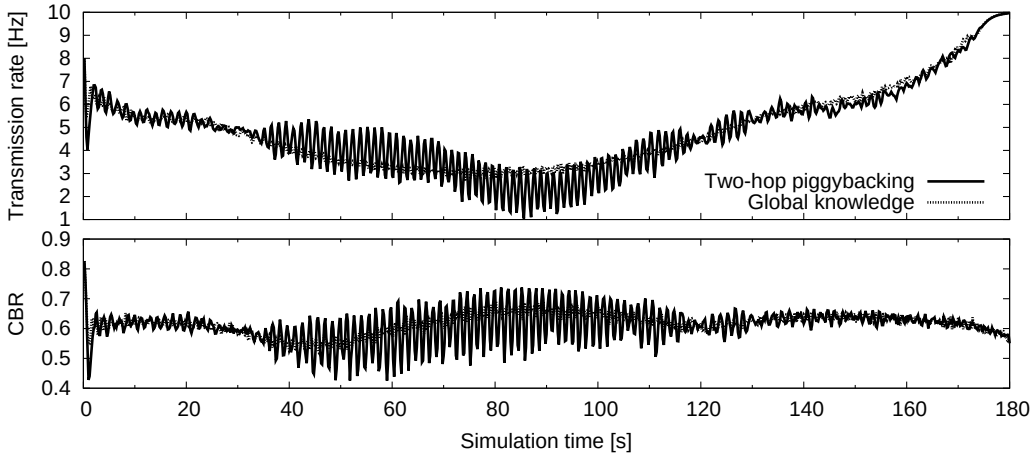
Figure 7.30: Convergence of LIMERIC [53] in the static crossing scenario

edge which approximates max-min fairness. Third, a smoother convergence may be achieved with asynchronous adaptation, which we leave for future work.

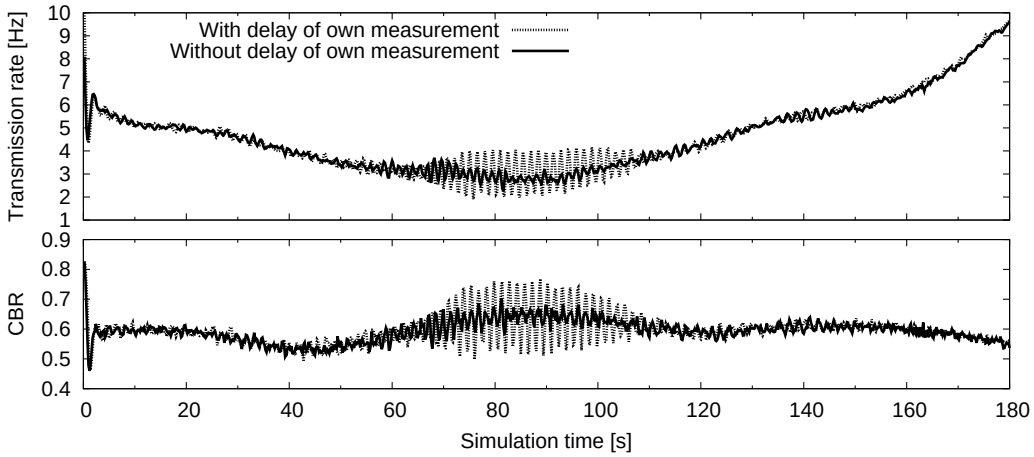
In summary, our results indicate that LIMERIC has the potential to improve the oscillation size resulting from AIMD in certain situations. However, the actual outcome appears to depend much on parameter choice and to be sensitive to the aggregation mechanism of the CBR feedback, while the outcome of AIMD in combination with the target rate mechanism seems to be more predictable and more robust to changes in vehicle density. In future work, the impact of different CBR filtering and aggregation techniques on the convergence of LIMERIC in mobile scenarios could be further investigated.



(a) Two-hop piggybacking using configuration (1)



(b) Two-hop piggybacking using configuration (1) and no delay of own measurement



(c) Two-hop piggybacking using configuration (2)

Figure 7.31: Adaptation by an exemplary node running LIMERIC [53] in the scenario where two groups of vehicles pass each other

7.5 Discussion

We conclude this chapter by discussing two known limitations of PULSAR and of the results in general presented in this thesis.

7.5.1 Piggybacking in Urban Scenarios

The two-hop piggybacking mechanism introduced in Section 6.4.4 on page 160 has been designed with a Line of Sight (LOS) environment in mind, e.g., in a highway situation. Under Non Line of sight (NLOS) conditions, however, the two-hop piggybacking approach might be too conservative.

Figure 7.32 on the facing page depicts four vehicles A, B, C and D in a crossing situation. Let us assume that d_2 corresponds to the configured hop distance and that the sum of distances d_1 and d_2 lies within the participation range of node A. Thus if under LOS conditions, C indicates a congested channel, A should participate in congestion control. On the other hand, if node D reports a congested channel, node A should participate in congestion control as well, since d_3, i.e., the line-of-sight distance between nodes A and D, lies within A's participation distance due to the triangle inequality.

However, if there is no line-of-sight between A and D, e.g., because of a building at the corner of the intersection, A does not necessarily have to participate in congestion control if D reports a congested channel. In this case, it might be useful to include a direction information with the relayed information. Alternatively, one-hop piggybacking could be used, potentially in combination with a higher carrier sense threshold to shorten the gap between carrier sense range and information dissemination range. We leave the exploration of how non-line-of-sight conditions impacts congestion control in general for future work.

7.5.2 Enhanced Distributed Channel Access (EDCA)

As discussed in Section 2.1.2 on page 13, IEEE 802.11p employs the Enhanced Distributed Channel Access (EDCA) mechanism of IEEE 802.11e which facilitates priority access for different kinds of traffic. A subtle difference between EDCA and the default Distributed Coordination Function (DCF) is the way the back-off counter is decremented. In DCF, the back-off counter is decremented *after* the medium was sensed idle for one slot time. In EDCA, it is always decremented at the *beginning* of each slot. As a consequence, EDCA can be expected to count down faster and to have lower channel access delays.

In this thesis, we use the network simulator ns-2.34 which includes an overhaul of the Physical layer (PHY) and Medium Access Control (MAC) layer according to the

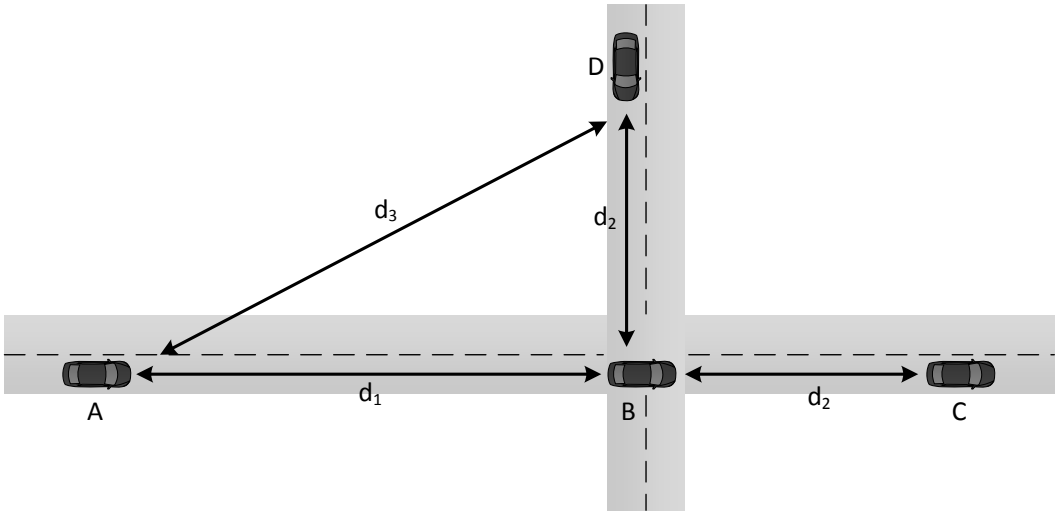


Figure 7.32: Schematic illustration of two-hop piggybacking in a crossing scenario

IEEE 802.11 standard but does not include EDCA [106]. At the time of conducting the simulation study in Chapter 4 on page 83, we were not aware of the difference in the countdown method and did not consider the lack of an EDCA implementation to be of consequence, since we did not intend to simulate different priorities of packets. To keep results consistent, we decided to use the DCF countdown method throughout the thesis. In addition, a comparison between both methods indicated that high channel access time values still occur in overload situations, since the busy time of the medium becomes the dominating factor. However, a further exploration of this topic is required, which we leave for future work.

8

Conclusions and Outlook

The next generation of driver assistance systems is envisioned to introduce cooperation between vehicles by means of communication technology. In this thesis, we have studied the scalability problem in Vehicle to Vehicle (V2V) safety communications resulting from the exchange of periodic status messages on top of which a cooperative awareness is established among vehicles. Without regulation, these messages can easily overload the channel. The underlying optimization problem has at least six degrees of freedom and the constraints of selecting parameters in a decentralized way and without message acknowledgments in a pervasive wireless broadcast environment while adhering to the real-time requirements of safety applications whose exact definition is still an ongoing field of research.

To make the problem tractable, we first studied the control implications and interrelationships of the identified degrees of freedom and select two control dimensions, i.e., transmit power and transmission rate. We then defined an optimization criterion based on the state of the art regarding the requirements of safety applications. In the next step, we evaluated a large parameter space to identify the optimal parameter combinations based on the selected optimization criterion. We identified a commonality in these optimized configurations which can help to significantly simplify the design of congestion control. We further identified a suitable system model, convergence target and fairness criteria for a congestion control solution. In addition, we proposed a novel method how congestion control can work under the guidance of safety applications and operate within limits derived from the driving context using a function to map transmission rate to safety benefit. Next, we introduced a concrete congestion control protocol to implement all of the aforementioned aspects and discussed several

design options. Finally, we demonstrated the general feasibility, fairness and stability of our concept and compared our protocol against other state of the art approaches, showing that "details matter" in the protocol design process.

With declarations of intent issued on both sides of the Atlantic to make V2V safety communications a reality in the near future, the topic of scalability and thus of channel congestion control has gained momentum in the community. Corresponding standards are on the way in Europe as well as in the United States. The topic can be expected to further gain in importance considering two recent developments. First, there are ongoing considerations to integrate IEEE 802.11p into mobile devices in order to enable pedestrians to participate in Cooperative Intelligent Transport Systems (C-ITS)[1]. This would expand the scalability problem to new dimensions. Second, it is under consideration in the United States to allow Wi-Fi usage in the 5.9 GHz band[2]. If the available spectrum has to be shared with non-safety transmissions, some aspects of congestion control may have to be revised, e.g., the selection of a carrier sense threshold.

We believe that the consolidation of the state of the art and as well as the insights presented in this thesis can contribute to the understanding and theory of congestion control for Vehicle Safety Communications (VSC) in general and provide a potential starting point for the development of the next generation of corresponding protocols. In particular, the results of the presented optimization study can help to simplify the design of congestion control without sacrificing the operability of safety applications. Combined with the concept of adapting the *relative* transmission rate of a vehicle based on its driving context, we have pointed out a way how congestion control can be designed to operate under the guidance of safety applications while providing a global fairness in terms of safety benefit. However, we consider these results early steps towards an optimal solution in this multidimensional optimization problem. In our opinion, more research is required especially with respect to the precise spatio-temporal requirements of safety applications as well as to the distribution of Packet Inter-Reception Time (IRT) in different environments.

In particular, we suggest the following research paths for future work:

- Confirmation of the results of the presented optimization study using an analytical approach, especially regarding the potential existence of an optimal probability of reception, cf. Section 4.2.2 on page 89.

[1] "Cooperative ITS for all: Enabling DSRC in mobile devices", presentation by Sundar Subramanian at the ETSI ITS Workshop 2013, available at *http://docbox.etsi.org/Workshop/2013/201302_ITSWORKSHOP/S05_COOPERATIVEITSandFUTUREASPECTS/QUALCOMM_SUBRAMANIAN.pdf*

[2] "ITS Communication Technologies: Disaster is Looming", presentation by Russell Shields at the ETSI ITS Workshop 2013, available at *http://docbox.etsi.org/Workshop/2013/201302_ITSWORKSHOP/S01_KEYNOTES/YGOMI_SHIELDS.pdf*

- Exploration of how the inversely proportional relationship of IRT-based awareness metrics and channel load can be addressed by congestion control, cf. Section 5.4.2 on page 126.
- Evaluation of the potential of the relative transmission rate concept to support an actual implementation of a safety application rather than a hypothetical one based on global knowledge, cf. Section 7.3 on page 191.
- Exploration of potential mapping functions between transmission rate and safety benefit, cf. Section 5.4.1 on page 124.
- Exploration of the feasibility of an open-loop Transmit Power Control (TPC) approach in a real-world environment, cf. Section 5.1.4 on page 109.
- Analysis of the implications of Non Line of sight (NLOS) conditions on congestion control, cf. Section 7.5.1 on page 216.
- Development of sanity checks for Channel Busy Ratio (CBR) measurements, cf. Section 6.4.2 on page 154.
- Integration of elements of SOURC [202] into PULSAR, cf. Sections 7.4.2 on page 206.

A

Simulation Environment and Configuration

A.1 Propagation model

In this thesis, we model radio wave propagation according to the following power-law model with log-normal fading. The resulting Received Signal Strength (RSS) in dBm is calculated as

$$P_{Rx}(d) = P_{Tx} - L - 10\gamma log_{10}(\frac{d}{d_o}) + X_\sigma \tag{A.1}$$

where P_{Tx} is the transmit power in dBm, L is the reference loss in dB, γ is the path loss exponent, d_o is the reference distance and X_σ is a Gaussian random variable with standard deviation σ. We thereby follow the recommendations of Mittag in [48] who performed a detailed analysis of the results of different measurement campaigns. Mittag found that for highway environments, the path loss results reported by different authors are very similar, i.e., Kunisch et al. [220], Karedal et al. [221] and Paier et al. [222]. The corresponding RSS per distance is illustrated by Figure A.1 on the next page. In this thesis, we apply the parameters of Kunisch et al. [220], i.e., L = 59.7 dBm, γ = 1.85 and σ = 3.2.

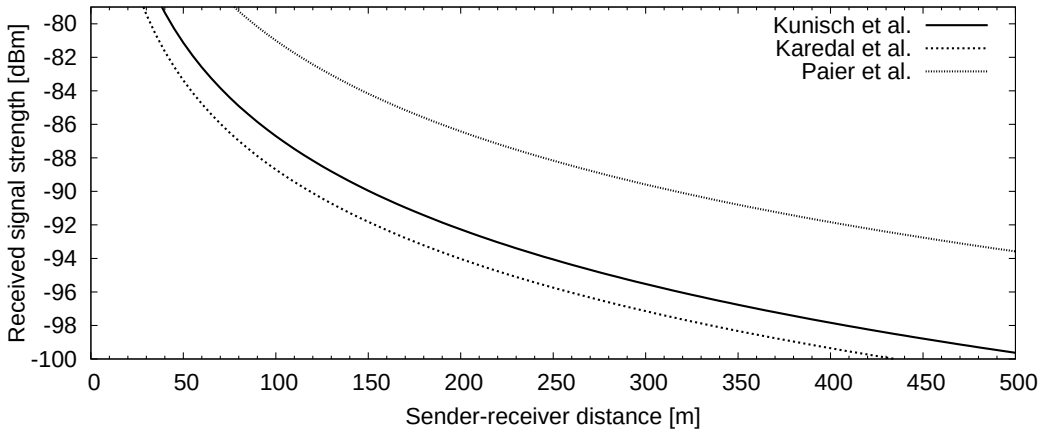


Figure A.1: RSS per distance according to the results of [220][221][222]

A.2 Network Simulator

The results presented in this thesis are based on the network simulator NS-2[1] in version 34, which is a discrete event simulator for network research. Due to its open source availability as well as the number of available models, NS-2 is a commonly used tool in the V2X community. The applied version includes the overhaul by Chen et al. [106] regarding the Physical layer (PHY) and Medium Access Control (MAC) layer according to the IEEE 802.11 standard. The simulator has also been validated against measurements by Bansal et al. [213]. Our default simulation parameters are summarized in Table A.1 on the facing page.

[1] *http://www.isi.edu/nsnam/ns/*

Parameter	Default value
Packet size including MAC overhead	400 Bytes
Data rate	6 Mbps
Noise floor	-99 dBm
Power monitor threshold	-102 dBm
Carrier sense threshold	-95 dBm
Minimum contention window	15
Preamble/header capture SINR	5 dB
Path loss exponent γ	1.85
Reference loss L	59.7 dBm
Fading	Lognormal (σ=3.2)
Slot time	13 μs
Header duration	40 μs
Symbol duration	8 μs

Table A.1: Default simulation parameters

Bibliography

[1] Tao Zhang and Luca Delgrossi. *Vehicle Safety Communications: Protocols, Security, and Privacy*. Wiley & Sons, 1st edition, 2012.

[2] Car2Car Communication Consortium. European Vehicle Manufacturers Working Hand in Hand on Deployment of Cooperative Intelligent Transport Systems and Services (C-ITS), 2012.

[3] Hannes Hartenstein and Kenneth P Laberteaux, editors. *VANET - Vehicular Applications and Inter-Networking Technologies*. Wiley & Sons, 1st edition, 2010.

[4] Electronic Communications Committee. ECC Decision of 14 March 2008 on the Harmonised Use of the 5875-5925 MHz Frequency Band for Intelligent Transport Systems (ITS), 2008.

[5] Association of Radio Industries and Businesses. 700 MHz Band Intelligent Transport Systems (ARIB STD-T109 V 1.0), 2012.

[6] United States Department of Transportation. Vehicle Safety Communications - Applications (VSC-A) Final Report, 2011.

[7] Fan Bai, Tamer Elbatt, Gavin Holland, Hariharan Krishnan, and Varsha Sadekar. Towards Characterizing and Classifying Communication-Based Automotive Applications from a Wireless Networking Perspective. In *Proceedings of the IEEE Workshop of Automotive Networking and Applications (AutoNet)*, 2006.

[8] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Definitions (ETSI TR 102 638 V1.1.1), 2009.

[9] Tessa Tielert, Moritz Killat, Hannes Hartenstein, Raphael Luz, Stefan Hausberger, and Thomas Benz. The Impact of Traffic-Light-to-Vehicle Communication on Fuel Consumption and Emissions. In *Proceedings of the Internet of Things Conference*, 2010.

[10] Tessa Tielert, David Rieger, Hannes Hartenstein, Raphael Luz, and Stefan Hausberger. Can V2X Communication Help Electric Vehicles Save Energy? In *Proceedings of the International Conference on ITS Telecommunications (ITST)*, 2012.

[11] Murat Caliskan, Daniel Graupner, and Martin Mauve. Decentralized Discovery of Free Parking Places. In *Proceedings of the ACM International Workshop on Vehicular Ad hoc Networks (VANET)*, 2006.

[12] Institute of Electical and Electronics Engineers. 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 802.11p-2010), 2010.

[13] United States Department of Transportation. Vehicle Safety Communications Project Task 3 Final Report - Identify Intelligent Vehicle Safety Applications Enabled by DSRC. Technical report, National Highway Traffic Safety Association, 2005.

[14] Thomas Mangel, Timo Kosch, and Hannes Hartenstein. A Comparison of UMTS and LTE for Vehicular Safety Communication at Intersections. In *Proceedings of the IEEE Vehicular Networking Conference*, 2010.

[15] Andreas Festag, Panos Papadimitratos, and Tessa Tielert. Design and Performance of Secure Geocast for Vehicular Communication. *IEEE Transactions on Vehicular Technology*, 59(5):2456–2471, 2010.

[16] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service (ETSI TS 102 637-2 V1.2.1), 2011.

[17] Society of Automotive Engineers. Dedicated Short Range Communications (DSRC) Message Set Dictionary (SAE J2735), 2009.

[18] Steven Shladover and Swe-Kuang Tan. Analysis of Vehicle Positioning Accuracy Requirements for Communication-Based Cooperative Collision Warning. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 10(3):131–140, September 2006.

[19] Miguel Sepulcre, Jens Mittag, Paolo Santi, Hannes Hartenstein, and Javier Gozalvez. Congestion and Awareness Control in Cooperative Vehicular Systems. *Proceedings of the IEEE*, 99(7):1260–1279, July 2011.

[20] Marc Torrent-Moreno, Jens Mittag, Paolo Santi, and Hannes Hartenstein. Vehicle-to-Vehicle Communication: Fair Transmit Power Control for Safety-Critical Information. *IEEE Transactions on Vehicular Technology*, 58(7):3684–3703, 2009.

[21] Miguel Sepulcre, Javier Gozalvez, and J. Hernandez. Cooperative Vehicle-to-Vehicle Active Safety Testing Under Challenging Conditions. *Transportation Research Part C: Emerging Technologies*, 26:233–255, January 2013.

[22] Van Jacobson and Michael Karels. Congestion a Avoidance and Control. *ACM SIGCOMM Computer Communication Review*, 18(4):314–329, August 1988.

[23] Ching-Ling Huang, Yaser Fallah, Raja Sengupta, and Hariharan Krishnan. Adaptive Intervehicle Communication Control for Cooperative Safety Systems. *IEEE Network*, 24(1):6–13, 2010.

[24] Natalya An, Michael Maile, Daniel Jiang, Jens Mittag, and Hannes Hartenstein. Balancing the Requirements for a Zero False Positive/Negative Forward Collision Warning. In *Proceedings of the 10th Annual Conference on Wireless On-Demand Network Systems and Services (WONS)*, 2013.

[25] Tessa Tielert, Daniel Jiang, Qi Chen, Luca Delgrossi, and Hannes Hartenstein. Design Methodology and Evaluation of Rate Adaptation Based Congestion Control for Vehicle Safety Communications. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2011.

[26] Tessa Tielert, Daniel Jiang, Hannes Hartenstein, and Luca Delgrossi. Joint Power/Rate Congestion Control Optimizing Packet Reception in Vehicle Safety Communications. In *Proceedings of the 10th ACM International Workshop on VehiculAr Inter-NETworking, Systems, and Applications (VANET)*, 2013.

[27] Qi Chen, Daniel Jiang, Tessa Tielert, and Luca Delgrossi. Mathematical Modeling of Channel Load in Vehicle Safety Communications. In *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, 2011.

[28] Jérôme Härri, Moritz Killat, Tessa Tielert, Jens Mittag, and Hannes Hartenstein. DEMO: Simulation-as-a-Service for ITS Applications. In *Proceedings of the 71st IEEE Vehicular Technology Conference*, 2010.

[29] John Kenney. Dedicated Short-Range Communications (DSRC) Standards in the United States. *Proceedings of the IEEE*, 99(7):1162 –1182, July 2011.

[30] Long Le, Andreas Festag, Roberto Baldessari, and Wenhui Zhang. CAR-2-X Communication in Europe. In Stephan Olariu and Michele C. Weigle, editors, *Vehicular Networks: From Theory to Practice*, chapter 4. Chapman & Hall/CRC Computer & Information Science Series, 2009.

[31] Christian Weiß. V2X Communication in Europe - From Research Projects Towards Standardization and Field Testing of Vehicle Communication Technology. *Computer Networks*, 55(14):3103–3119, October 2011.

[32] Georgios Karagiannis, Onur Altintas, Eylem Ekici, Geert Heijenk, Boangoat Jarupan, Kenneth Lin, and Timothy Weil. Vehicular Networking: A Survey and Tutorial on Requirements, Architectures, Challenges, Standards and Solutions. *IEEE Communications Surveys & Tutorials*, 13(4):584–616, 2011.

[33] Hannes Hartenstein and Kenneth Laberteaux. A Tutorial Survey on Vehicular Ad Hoc Networks. *IEEE Communications Magazine*, 46(June):164–171, 2008.

[34] Katrin Sjöberg. *Medium Access Control for Vehicular Ad Hoc Networks*. PhD thesis, Chalmers University of Technology, Sweden, 2013.

[35] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Access Layer Specification for Intelligent Transport Systems Operating in the 5 GHz Frequency Band (Draft ETSI EN 302 663 V1.2.0), 2012.

[36] Benjamin Henty, Reginald Cooper, and Daniel Stancil. A Measurement Study of Time-Scaled 802.11a Waveforms Over The Mobile-to-Mobile Vehicular Channel at 5.9 GHz. *IEEE Communications Magazine*, 46(5):84–91, May 2008.

[37] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Harmonized Channel Specifications for Intelligent Transport Systems Operating in the 5 GHz Frequency Band (ETSI TS 102 724 V1.1.1), 2012.

[38] Raffi Sevlian, Carl Chun, Ian Tan, Ahmad Bahai, and Kenneth Laberteaux. Channel Characterization for 700 MHz DSRC Vehicular Communication. *Journal of Electrical and Computer Engineering*, 2010:1–9, January 2010.

[39] Theodore Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 2nd edition, 2002.

[40] Francesca Martelli, M. Elena Renda, and Paolo Santi. Measuring IEEE 802.11p Performance for Active Safety Applications in Cooperative Vehicular Systems. In *Proceedings of the 73rd IEEE Vehicular Technology Conference (VTC Spring)*, 2011.

[41] Derek Caveney. Cooperative Vehicular Safety Applications. *IEEE Control Systems Magazine*, 30(4):38–53, August 2010.

[42] Institute of Electical and Electronics Engineers. 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 802.11-2012), 2012.

[43] Daniel Jiang, Qi Chen, and Luca Delgrossi. Optimal Data Rate Selection for Vehicle Safety Communications. In *Proceedings of the 5th ACM International Workshop on VehiculAr Inter-NETworking (VANET)*, 2008.

[44] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Properties of the MAC Layer in Safety Vehicular Ad Hoc Networks. *IEEE Communications Magazine*, 50(5):192 –200, May 2012.

[45] Robert Schmidt, T. Köllmer, Tim Leinmüller, Bert Böddeker, and Günter Schäfer. Degradation of Transmission Range in VANETs Caused by Interference. *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, 32(4):224–234, 2009.

[46] Marc Torrent-Moreno, Steven Corroy, Felix Schmidt-Eisenlohr, and Hannes Hartenstein. IEEE 802.11-Based One-Hop Broadcast Communications. In *Proceedings of the 9th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, 2006.

[47] Jens Mittag and Hannes Hartenstein. Is CSMA Able to Coordinate Multiple Access in Vehicular Radio Channels Effectively? In *Proceedings of the 12th International Conference on ITS Telecommunications (ITST)*, 2012.

[48] Jens Mittag. *Characterization, Avoidance and Repair of Packet Collisions in Inter-Vehicle Communication Networks*. PhD thesis, Karlsruhe Institute of Technology (KIT), Germany, 2012.

[49] James Kurose and Keith Ross. *Computer Networking: A Top-Down Approach*. Pearson Education, 6th edition, 2012.

[50] Sarah Sharafkandi, Gaurav Bansal, John Kenney, and David Du. A Novel Use of EDCA to Improve Vehicle Safety Communication. In *Proceedings of the 9th ACM International Workshop on Vehicular Inter-Networking, Systems and Applications (VANET)*, 2012.

[51] Sarah Sharafkandi, Gaurav Bansal, John Kenney, and David Du. Using EDCA to improve Vehicle Safety Messaging. In *Proceedings of the Vehicular Networking Conference (VNC)*, 2012.

[52] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Communications Architecture (ETSI EN 302 665 V1.1.1), 2010.

[53] John Kenney. Standards and Regulations. In Hannes Hartenstein and Kenneth P Laberteaux, editors, *VANET - Vehicular Applications and Inter-Networking Technologies*, chapter 10, pages 365—-428. Wiley & Sons, 2010.

[54] Institute of Electical and Electronics Engineers. Information Technology - Telecommunications and Information Local and Metropolitan Area Networks - Specific Requirements; Part 2: Logical Link Control (ANSI/IEEE Std 802.2, 1998 edition), 1998.

[55] Institute of Electical and Electronics Engineers. Local and Metropolitan Area Network Standards (IEEE 802-2001), 2001.

[56] Institute of Electical and Electronics Engineers. IEEE Standard for Wireless Access in Vehicular Environments (WAVE)- Multi-channel Operation (IEEE Std 1609.4-2010), 2011.

[57] Andreas Festag, Long Le, and Maria Goleva. Field Operational Tests for Cooperative Systems. In *Proceedings of the 8th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2011.

[58] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 4: Geographical Addressing and Forwarding for Point-to-Point and Point-to-Multipoint Communications; Sub-Part 1: Media-Independent Functionality (ETSI TS 102 636-4-1 V1.1.1), 2011.

[59] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; GeoNetworking; Part 5: Transport Protocols; Sub-Part 1: Basic Transport Protocol (ETSI TS 102 636-5-1 V1.1.1), 2011.

[60] Institute of Electical and Electronics Engineers. IEEE Standard for Wireless Access in Vehicular Environments (WAVE) - Networking Services (IEEE Std 1609.3-2010), 2010.

[61] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Users and Applications Requirements; Part 1: Facility Layer Structure, Functional Requirements and Specifications (ETSI TS 102 894-1 V1.1.1), 2013.

[62] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2 : Specification of Cooperative Awareness Basic Service (Draft ETSI EN 302 637-2 V1.3.0), 2013.

[63] European Telecommunications Standards Institute. Vehicular Communications; Part 3 : Specifications of Decentralized Environmental Notification Basic Service (Draft ETSI EN 302 637-3 V1.2.0), 2013.

[64] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); V2X Applications; Part 1: Road Hazard Signalling (RHS) Application Requirements Specification (TS 101 539-1 V1.1.1), 2013.

[65] Institute of Electical and Electronics Engineers. IEEE Standard for Wireless Access in Vehicular Environments - Security Services for Applications and Management Messages (IEEE Std 1609.2-2013), 2013.

[66] André Weimerskirch, Jason Haas, Yih-Chun Hu, and Kenneth Laberteaux. Data Security in Vehicular Communication Networks. In Hannes Hartenstein and Kenneth P Laberteaux, editors, *VANET - Vehicular Applications and Inter-Networking Technologies*, chapter 9, pages 299–363. Wiley & Sons, Chichester, UK, January 2010.

[67] Elmar Schoch and Frank Kargl. On the Efficiency of Secure Beaconing in VANETs. In *Proceedings of the 3rd ACM Conference on Wireless Network Security (WiSec)*, 2010.

[68] Michael Feiri, Jonathan Petit, and Frank Kargl. Congestion-Based Certificate Omission in VANETs. In *Proceedings of the 9th ACM International Workshop on Vehicular Inter-Networking, Systems and Applications (VANET)*, 2012.

[69] Panos Papadimitratos, Levente Buttyan, Tamas Holczer, Elmar Schoch, Julien Freudiger, Maxim Raya, Zhendong Ma, Frank Kargl, Antonio Kung, and Jean-Pierre Hubaux. Secure Vehicular Communication Systems: Design and Architecture. *IEEE Communications Magazine*, 46(11):100–109, November 2008.

[70] Maxim Raya and Jean-Pierre Hubaux. The Security of Vehicular Ad Hoc Networks. In *Proceedings of the 3rd ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN)*, 2005.

[71] John Kenney, Gaurav Bansal, and Charles Rohrs. LIMERIC: A Linear Message Rate Control Algorithm for Vehicular DSRC Systems. In *Proceedings of the 8th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2011.

[72] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Decentralized Congestion Control Mechanisms for Intelligent Transport Systems Operating in the 5 GHz Range; Access Layer Part (ETSI TS 102 687 V1.1.1), 2011.

[73] Sundar Subramanian, Marc Werner, Shihuan Liu, Jubin Jose, Radu Lupoaie, and Xinzhou Wu. Congestion Control for Vehicular Safety: Synchronous and Asynchronous MAC Algorithms. In *Proceedings of the 9th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2012.

[74] David Eckhoff, Nikoletta Sofra, and Reinhard German. A Performance Study of Cooperative Awareness in ETSI ITS G5 and IEEE WAVE. In *Proceedings of the 10th Annual Conference on Wireless On-Demand Network Systems and Services (WONS)*, 2013.

[75] Han-Shue Tan and Jihua Huang. DGPS-Based Vehicle-to-Vehicle Cooperative Collision Warning: Engineering Feasibility Viewpoints. *IEEE Transactions on Intelligent Transportation Systems*, 7(4):415–428, December 2006.

[76] United States Departement of Defense. Global Positioning System - Standard Positioning Service Performance Standard, 4th Edition, 2008.

[77] Martin Mauve and Björn Scheuermann. VANET Convenience and Efficiency Applications. In Hannes Hartenstein and Kenneth P Laberteaux, editors, *VANET - Vehicular Applications and Inter-Networking Technologies*, chapter 4, pages 81–105. Wiley & Sons, 2010.

[78] USDOT (United States Department of Transportation). Vehicle Safety Communications - Applications (VSC-A) Final Report: Appendix Volume 1 - System Design and Objective Test, 2011.

[79] Use Cases, Functional Specifications and Safety Margin Applications for the SAFESPOT Project (Deliverable N. D8.4.4), 2008.

[80] CAR 2 CAR Communication Consortium Manifesto: Overview of the C2C-CC System v. 1.1, 2007.

[81] Huiqin Chen, Libo Cao, and David Logan. Investigation Into the Effect of an Intersection Crash Warning System on Driving Performance in a Simulator. *Traffic injury prevention*, 12(5):529–37, October 2011.

[82] Beshr Sultan and Mike Mcdonald. Assessing The Safety Benefit of Automatic Collision Avoidance Systems (During Emergency Braking Situations). In *Proceedings of the 18th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, 2003.

[83] Jason Haas and Yih-Chun Hu. Communication Requirements for Crash Avoidance. In *Proceedings of the 7th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2010.

[84] Michele Segata and Renato Lo Cigno. Emergency Braking: A Study of Network and Application Performance. In *Proceedings of the 8th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2011.

[85] Stefan Joerer, Michele Segata, Bastian Bloessl, Renato Lo Cigno, Christoph Sommer, and Falko Dressler. To Crash or Not to Crash: Estimating its Likelihood and Potentials of Beacon-Based IVC Systems. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2012.

[86] Moritz Killat and Hannes Hartenstein. Vehicular Ad Hoc Networks: How to Show the Impact on Traffic Safety? In *Proceedings of the 65th IEEE Vehicular Technology Conference (VTC Spring)*, 2007.

[87] Moritz Killat, Tristan Gaugel, and Hannes Hartenstein. Enabling Traffic Safety Assessment of VANETs by Means of Accident Simulations. In *Proceedings of the 19th IEEE Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2008.

[88] Wenhui Zhang, Andreas Festag, Roberto Baldessari, and Long Le. Congestion Control for Safety Messages in VANETs: Concepts and Frameworks. In *Proceedings of the 8th Conference on ITS Telecommunications (ITST)*, 2008.

[89] Xue Yang, Jie Liu, Feng Zhao, and Nitin Vaidya. A Vehicle-to-Vehicle Communication Protocol for Cooperative Collision Warning. In *Proceedings of the 1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous)*, 2004.

[90] Tamer ElBatt, Siddhartha K Goel, Gavin Holland, Hariharan Krishnan, and Jayendra Parikh. Cooperative Collision Warning Using Dedicated Short Range Wireless Communications. In *Proceedings of the 3rd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, 2006.

[91] Fan Bai and Hariharan Krishnan. Reliability Analysis of DSRC Wireless Communication for Vehicle Safety Applications. In *Proceedings of the IEEE Intelligent Transportation Systems Conference (ITSC)*, 2006.

[92] Jens Mittag, Florian Thomas, Jérôme Härri, and Hannes Hartenstein. A Comparison of Single- and Multi-Hop Beaconing in VANETs. In *Proceedings of the 6th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2009.

[93] Miguel Sepulcre, Javier Gozalvez, Jérôme Härri, and Hannes Hartenstein. Application-Based Congestion Control Policy for the Communication Channel in VANETs. *IEEE Communications Letters*, 14(10):951–953, 2010.

[94] Natalya An, Tristan Gaugel, and Hannes Hartenstein. VANET: Is 95% Probability of Packet Reception Safe? In *Proceedings of the 11th International Conference on ITS Telecommunications (ITST)*, 2011.

[95] Robert Schmidt and Tim Leinmüller. A Spatio-Temporal Metric for the Evaluation of Cooperative Awareness. In *Proceedings of the 18th World Congress on Intelligent Transport Systems*, 2011.

[96] Hongsheng Lu and Christian Poellabauer. Analysis of Application-Specific Broadcast Reliability for Vehicle Safety Communications. In *Proceedings of the 8th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2011.

[97] Francesca Martelli, M. Elena Renda, Giovanni Resta, and Paolo Santi. A Measurement-Based Study of Beaconing Performance in IEEE 802.11p Vehicular Networks. In *Proceedings of the IEEE INFOCOM*, 2012.

[98] Bernhard Kloiber, Cristina Rico-Garcia, Jérôme Härri, and Thomas Strang. Update Delay: A New Information-Centric Metric for a Combined Communication and Application Level Reliability Evaluation of CAM Based Safety Applications. In *Proceedigs of the 19th ITS World Congress*, 2012.

[99] Sanjit Kaul, Marco Gruteser, Vinuth Rai, and John Kenney. Minimizing Age of Information in Vehicular Networks. In *Proceedings of the 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2011.

[100] Transportation Research Board. Highway Capacity Manual, 2000.

[101] Yu-Chee Tseng, Sze-Yao Ni, Yuh-Shyan Chen, and Jang-Ping Sheu. The Broadcast Storm Problem in a Mobile Ad Hoc Network. *Wireless Networks*, 8(2/3):153–167, March 2002.

[102] Ozan K. Tonguz, Nawaporn Wisitpongphan, and Fan Bai. DV-CAST: A Distributed Vehicular Broadcast Protocol for Vehicular Ad Hoc Networks. *IEEE Wireless Communications*, 17(2):47–57, April 2010.

[103] Sangoh Oh, Marco Gruteser, and Dario Pompili. Coordination-Free Safety Messages Dissemination Protocol for Vehicular Networks. *IEEE Transactions on Vehicular Technology*, PP(99):1–1, 2012.

[104] Jing Deng, Ben Liang, and Pramod K. Varshney. Tuning the Carrier Sensing Range of IEEE 802.11 MAC. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, 2004.

[105] Piyush Gupta and P.R. Kumar. The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[106] Qi Chen, Felix Schmidt-Eisenlohr, Daniel Jiang, Marc Torrent-Moreno, Luca Delgrossi, and Hannes Hartenstein. Overhaul of IEEE 802.11 Modeling and Simulation in NS-2. In *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems (MSWiM)*, 2007.

[107] Jing Zhu, Xingang Guo, Lily Yang, Steven Conner, Sumit Roy, and Mousumi Hazra. Adapting Physical Carrier Sensing to Maximize Spatial Reuse in 802.11 Mesh Networks. *Wireless Communications and Mobile Computing*, 4(8):933–946, December 2004.

[108] Jing Zhu, Benjamin Metzler, Xingang Guo, and York Liu. Adaptive CSMA for Scalable Network Capacity in High-Density WLAN: A Hardware Prototyping Approach. In *Proceedings of IEEE INFOCOM*, 2006.

[109] Yong Yang, Jennifer Hou, and Lu-Chuang Kung. Modeling the Effect of Transmit Power and Physical Carrier Sense in Multi-Hop Wireless Networks. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2007.

[110] Hui Ma, Rajiv Vijayakumar, Sumit Roy, and Jing Zhu. Optimizing 802.11 Wireless Mesh Networks Based on Physical Carrier Sensing. *IEEE/ACM Transactions on Networking*, 17(5):1550–1563, October 2009.

[111] Xue Yang and Nitin Vaidya. On physical carrier sensing in wireless ad hoc networks. In *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, 2005.

[112] Daniel Jiang, Qi Chen, and Luca Delgrossi. Communication Density: A Channel Load Metric for Vehicular Communications Research. In *Proceedings of the*

IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS), 2007.

[113] Marc Torrent-Moreno. *Inter-Vehicle Communications: Achieving Safety in a Distributed Wireless Environment*. Dissertation, Karlsruhe Institute of Technology (KIT), Germany, 2007.

[114] William Stallings. *Data And Computer Communications*. Pearson Education, 8th edition, 2009.

[115] Raj Jain and K. K. Ramakrishnan. Congestion Avoidance in Computer Networks with a Connectionless Network Layer: Concepts, Goals and Methodology. In *Proceedings of the Computer Networking Symposium*, 1988.

[116] Raj Jain, K. K. Ramakrishnan, and Dah-Ming Chiu. Congestion Avoidance in Computer Networks with a Connectionless Network Layer. Technical report, Digital Equipment Corporation, December 1997.

[117] Dimitri Bertsekas and Robert Gallager. Flow Control. In *Data Networks*, chapter 6, pages 493—-536. Prentice Hall, 2nd edition, 1992.

[118] Frank Kelly, A.K. Maulloo, and D.K.H. Tan. Rate Control for Communication Networks: Shadow Prices, Proportional Fairness and Stability. *Journal of the Operational Research Society*, 49(3):237–252, April 1998.

[119] Jean-Yves Le Boudec. Rate Adaptation, Congestion Control and Fairness: A Tutorial, 2012.

[120] Dah-Ming Chiu and Raj Jain. Analysis of the Increase and Decrease algorithms for Congestion Avoidance in Computer Networks. *Computer Networks and ISDN Systems*, 17(1):1–14, June 1989.

[121] Sergey Gorinsky and Harrick Vin. Additive Increase Appears Inferior. Technical Report October 1999, Dept. of Computer Sciences, University of Texas at Austin, 2000.

[122] Sergey Gorinsky and Harrick Vin. Extended Analysis of Binary Adjustment Algorithms. Technical report, Department of Computer Sciences, University of Texas at Austin, 2002.

[123] Cui-Qing Yang and Alapati Reddy. A Taxonomy for Congestion Control Algorithms in Packet Switching Networks. *IEEE Network*, 9(4):34–45, 1995.

[124] K. K. Ramakrishnan and Raj Jain. A Binary Feedback Scheme for Congestion Avoidance in Computer Networks. *ACM Transactions on Computer Systems*, 8(2):158–181, May 1990.

[125] Ruttikorn Varakulsiripunth, Norio Shiratori, and Shoichi Noguchi. A Congestion-Control Policy on the Internetwork Gateway. *Journal of Computer Networks and ISDN Systems*, 11(1):43–58, 1986.

[126] Larry Peterson and Bruce Davie. *Computer Networks: A Systems Approach*. Morgan Kaufmann, 4th edition, 2008.

[127] Kevin Fall and W. Richard Stevens. *TCP/IP Illustrated, Volume 1: The Protocols*, volume 1. Addison Wesley, 2nd edition, 1995.

[128] Phil Karn and Craig Partridge. Improving Round-Trip Time Estimates in Reliable Transport Protocols. In *Proceedings of the ACM Workshop on Frontiers in Computer Communications Technology (SIGCOMM)*, 1987.

[129] George Xylomenos, George Polyzos, Petri Mähönen, and Mika Saaranen. TCP Performance Issues over Wireless Links. *IEEE Communications Magazine*, 39(4):52–58, April 2001.

[130] Hari Balakrishnan, Venkata Padmanabhan, Srinivasan Seshan, and Randy Katz. A Comparison of Mechanisms for Improving TCP Performance Over Wireless Links. In *Proceedings of the Annual Conference of the ACM Special Interest Group on Data Communication (SIGCOMM) on the Applications, Technologies, Architectures, and Protocols for Computer Communication*, 1996.

[131] Ye Tian, Kai Xu, and Nirwan Ansari. TCP in Wireless Environments: Problems and Solutions. *IEEE Communications Magazine*, 43(3):S27–S32, March 2005.

[132] Roy Yates. A Framework for Uplink Power Control in Cellular Radio Systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, 1995.

[133] Vaggelis Douros and George Polyzos. Review of Some Fundamental Approaches for Power Control in Wireless Networks. *Computer Communications*, 34(13):1580–1592, 2011.

[134] Mohammed Elmusrati and Heikki Koivo. Multi-Objective Totally Distributed Power and Rate Control for Wireless Communications. In *Proceedings of the 57th Vehicular Technology Conference (VTC-Spring)*, 2003.

[135] Mung Chiang, Prashanth Hande, Tian Lan, and Chee Wei Tan. Power Control in Wireless Cellular Networks. *Foundations and Trends in Networking*, 2(4):381–533, April 2007.

[136] Ad Kamerman and Leo Monteban. WaveLAN-II: A High-Performance Wireless LAN for the Unlicensed Band. *Bell Labs Technical Journal*, Summer 1997:118–133, 1997.

[137] Jongseok Kim, Seongkwan Kim, Sunghyun Choi, and Daji Qiao. CARA: Collision-Aware Rate Adaptation for IEEE 802.11 WLANs. In *Proceedings of the*

25th IEEE International Conference on Computer Communications (INFOCOM), 2006.

[138] Sumit Rangwala, Apoorva Jindal, Ki-Young Jang, Konstantinos Psounis, and Ramesh Govindan. Understanding Congestion Control in Multi-Hop Wireless Mesh Networks. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2008.

[139] Saad Biaz and Shaoen Wu. Rate Adaptation Algorithms for IEEE 802.11 Networks: A Survey and Comparison. In *Proceedings of the IEEE Symposium on Computers and Communications (ISCC)*, 2008.

[140] Daji Qiao, Sunghyun Choi, Amit Jain, and Kang G. Shin. MiSer: An Optimal Low-Energy Transmission Strategy for IEEE 802.11a/h. In *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2003.

[141] Vivek Mhatre, Konstantia Papagiannaki, and Francois Baccelli. Interference Mitigation Through Power Control in High Density 802.11 WLANs. In *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM)*, 2007.

[142] Kishore Ramachandran, Ravi Kokku, Honghai Zhang, and Marco Gruteser. Symphony: Synchronous Two-Phase Rate and Power Control in 802.11 WLANs. In *Proceeding of the 6th International Conference on Mobile Systems, Applications, and Services (MobiSys)*, 2008.

[143] Chieh-Yih Wan, Shane Eisenman, and Andrew Campbell. CODA: Congestion Detection and Avoidance in Sensor Networks. In *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys)*, 2003.

[144] Sumit Rangwala, Ramakrishna Gummadi, Ramesh Govindan, and Konstantinos Psounis. Interference-Aware Fair Rate Control in Wireless Sensor Networks. *ACM SIGCOMM Computer Communication Review*, 36:63–74, August 2006.

[145] Jeongyeup Paek and Ramesh Govindan. RCRT: Rate-Controlled Reliable Transport Protocol for Wireless Sensor Networks. *ACM Transactions on Sensor Networks*, 7(3):1–45, September 2010.

[146] Bret Hull, Kyle Jamieson, and Hari Balakrishnan. Mitigating Congestion in Wireless Sensor Networks. In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys)*, 2004.

[147] Avinash Sridharan and Bhaskar Krishnamachari. Explicit and Precise Rate Control for Wireless Sensor Networks. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys)*, 2009.

[148] Paolo Santi. Topology Control in Wireless Ad Hoc and Sensor Networks. *ACM Computing Surveys*, 37(2):164–194, June 2005.

[149] Shan Lin, Jingbin Zhang, Gang Zhou, Lin Gu, John Stankovic, and Tian He. ATPC: Adaptive Transmission Power Control for Wireless Sensor Networks. In *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys)*, 2006.

[150] Jaein Jeong, David Culler, and Jae-Hyuk Oh. Empirical Analysis of Transmission Power Control Algorithms for Wireless Sensor Networks. In *Proceedings of the 4th International Conference on Networked Sensing Systems*, 2007.

[151] Swetha Narayanaswamy, Vikas Kawadia, R.S. Sreenivas, and P.R. Kumar. Power Control in Ad-Hoc Networks: Theory, Architecture, Algorithm and Implementation of the COMPOW Protocol. In *Proceedings of the European Wireless Conference*, 2002.

[152] Mihail Sichitiu and Maria Kihl. Inter-Vehicle Communication Systems: A Survey. *IEEE Communications Surveys & Tutorials*, 10(2):88–105, 2008.

[153] Christian Lochert, Björn Scheuermann, and Martin Mauve. A Survey on Congestion Control for Mobile Ad Hoc Networks. *Wireless Communications and Mobile Computing*, 7(5):655–676, June 2007.

[154] Marwan Krunz, Alaa Muqattash, and Sung-ju Lee. Transmission Power Control in Wireless Ad Hoc Networks: Challenges, Solutions, and Open Issues. *IEEE Network*, 18(5):8–14, 2004.

[155] Jun Yin, Xiadong Wang, and Dharma Argawal. Optimal Packet Size in Error-Prone Channel for IEEE 802.11 Distributed Coordination Function. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, 2004.

[156] Craig Robinson, Lorenzo Caminiti, Derek Caveney, and Kenneth Laberteaux. Efficient Coordination and Transmission of Data for Cooperative Vehicular Safety Applications. In *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks (VANET)*, 2006.

[157] Craig Robinson, Derek Caveney, Lorenzo Caminiti, Grirish Baliga, Kenneth Laberteaux, and P.R. Kumar. Efficient Message Composition and Coding for Cooperative Vehicular Safety Applications. *IEEE Transactions on Vehicular Technology*, 56(6):3244–3255, November 2007.

[158] Michael Feiri, Jonathan Petit, and Frank Kargl. Evaluation of Congestion-Based Certificate Omission in VANETs. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2012.

[159] Giorgio Calandriello, Panos Papadimitratos, Jean-Pierre Hubaux, and Antonio Lioy. On the Performance of Secure Vehicular Communication Systems. *IEEE Transactions on Dependable and Secure Computing*, 8(6):898–912, November 2011.

[160] Joseph Camp and Edward Knightly. Modulation Rate Adaptation in Urban and Vehicular Environments. In *Proceedings of the 14th ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2008.

[161] Pravin Shankar, Tamer Nadeem, Justinian Rosca, and Liviu Iftode. CARS: Context-Aware Rate Selection for Vehicular Networks. In *Proceedings of the IEEE International Conference on Network Protocols*, 2008.

[162] Pralhad Deshpande and Samir Das. BRAVE: Bit-Rate Adaptation in Vehicular Environments. In *Proceedings of the 9th ACM International Workshop on Vehicular Inter-Networking, Systems, and Applications (VANET)*, 2012.

[163] Fan Bai, Daniel Stancil, and Hariharan Krishnan. Toward Understanding Characteristics of Dedicated Short Range Communications (DSRC) from a Perspective of Vehicular Network Engineers. In *Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2010.

[164] Joseph Fernandez, Kevin Borries, Lin Cheng, B. V. K. Vijaya Kumar, Daniel Stancil, and Fan Bai. Performance of the 802.11p Physical Layer in Vehicle-to-Vehicle Environments. *IEEE Transactions on Vehicular Technology*, 61(1):3–14, 2012.

[165] Robert Schmidt, Tim Leinmüller, Elmar Schoch, Frank Kargl, and Günter Schäfer. Exploration of Adaptive Beaconing for Efficient Intervehicle Safety Communication. *IEEE Network*, 24(1):14–19, January 2010.

[166] Ching-Ling Huang, Hariharan Krishnan, Raja Sengupta, and Yaser Fallah. Implementation and Evaluation of Scalable Vehicle-to-Vehicle Transmission Control Protocol. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2010.

[167] Giuseppe Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, 2000.

[168] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Why VANET Beaconing Is More Than Simple Broadcast. In *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, 2011.

[169] Martijn van Eenennaam, Luuk Hendriks, Georgios Karagiannis, and Geert Heijenk. Oldest Packet Drop (OPD): A Buffering Mechanism for Beaconing

in IEEE 802.11p VANETs (Poster). In *Proceedings of the Vehicular Networking Conference (VNC)*, 2011.

[170] René Reinders, Martijn van Eenennaam, Georgios Karagiannis, and Geert Heijenk. Contention Window Analysis for Beaconing in VANETs. In *Proceedings of the 7th International Wireless Communications and Mobile Computing Conference*, 2011.

[171] Nathan Balon and Jinhua Guo. Increasing Broadcast Reliability in Vehicular Ad Hoc Networks. In *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks (VANET)*, 2006.

[172] Yvonne Mertens, Matthias Wellens, and Petri Mähönen. Simulation-Based Performance Evaluation of Enhanced Broadcast Schemes for IEEE 802.11-Based Vehicular Networks. In *Proceedings of the IEEE Vehicular Technology Conference (VTC)*, 2008.

[173] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Enhancements of IEEE 802.11p Protocol for Access Control on a VANET Control Channel. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2011.

[174] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Broadcast Communication in Vehicular Ad-Hoc Network Safety Applications. In *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC)*, 2011.

[175] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Local Density Estimation for Contention Window Adaptation in Vehicular Networks. In *Proceedings of the 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011.

[176] Stephan Eichler and Christoph Schroth. A Multi-Layer Approach for Improving the Scalability of Vehicular Ad-Hoc Networks. In *Proceedings of the ITG-GI Conference on Communication in Distributed Systems (KiVS)*, 2007.

[177] Tarik Taleb, Keisuke Ooi, and Kazuo Hashimoto. An Efficient Collision Avoidance Strategy for ITS Systems. In *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2008.

[178] Yunpeng Zang, Lothar Stibor, Xi Cheng, Hans-Jürgen Reumerman, Arthur Paruzel, and Andre Barroso. Congestion Control in Wireless Networks for Vehicular Safety Applications. In *Proceedings of the 8th European Wireless Conference*, 2007.

[179] Xiaomin Ma, Jinsong Zhang, and Tong Wu. Reliability Analysis of One-Hop Safety-Critical Broadcast Services in VANETs. *IEEE Transactions on Vehicular Technology*, 60(8):3933–3946, October 2011.

[180] Jason Fuemmeler, Nitin Vaidya, and Venugopal Veeravalli. Selecting Transmit Powers and Carrier Sense Thresholds in CSMA Protocols for Wireless Ad Hoc Networks. In *Proceedings of the 2nd Annual International Workshop on Wireless Internet (WICON)*, 2006.

[181] Robert Schmidt, Tim Leinmüller, Bert Böddeker, and Günter Schäfer. Adapting the Wireless Carrier Sensing for VANETs. In *Proceedings of the 6th International Workshop on Intelligent Transportation (WIT)*, 2010.

[182] Robert Schmidt, Achim Brakemeier, Tim Leinmüller, Frank Kargl, and Günter Schäfer. Advanced Carrier Sensing to Resolve Local Channel Congestion. In *Proceedings of the 8th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2011.

[183] Razvan Stanica, Emmanuel Chaput, and André-Luc Beylot. Physical Carrier Sense in Vehicular Ad-Hoc Networks. In *Proceedings of the 8th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems*, 2011.

[184] Maen Artimy. Local Density Estimation and Dynamic Transmission-Range Assignment in Vehicular Ad Hoc Networks. *IEEE Transactions on Intelligent Transportation Systems*, 8(3):400–412, September 2007.

[185] Guiseppe Caizzone, Paolo Giacomazzi, Luigi Musumeci, and Giacomo Verticale. A Power Control Algorithm With High Channel Availability for Vehicular Ad Hoc Networks. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2005.

[186] Xu Guan, Raja Sengupta, Hariharan Krishnan, and Fan Bai. A Feedback-Based Power Control Algorithm Design for VANET. In *Proceedings of Mobile Networking for Vehicular Environments*, 2007.

[187] Lin Cheng and Rahul Shakya. VANET Adaptive Power Control from Realistic Propagation and Traffic Modeling. In *Proceedings of the IEEE Radio and Wireless Symposium (RWS)*, 2010.

[188] Marc Torrent-Moreno, Paolo Santi, and Hannes Hartenstein. Distributed Fair Transmit Power Adjustment for Vehicular Ad Hoc Networks. In *Proceedings of the 3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks*, 2006.

[189] Jens Mittag, Felix Schmidt-Eisenlohr, Moritz Killat, Jérôme Härri, and Hannes Hartenstein. Analysis and Design of Effective and Low-Overhead Transmission Power Control for VANETs. In *Proceedings of the 5th ACM International Workshop on Vehicular Inter-Networking (VANET)*, 2008.

[190] Hongsheng Lu and Christian Poellabauer. Balancing Broadcast Reliability and Transmission Range in VANETs. In *Proceedings of the IEEE Vehicular Networking Conference*, 2010.

[191] Chonlatee Khorakhun, Holger Busche, and Hermann Rohling. Congestion Control for VANETs Based on Power or Rate Adaptation. In *Proceedings of the 5th International Workshop on Intelligent Transportation (WIT)*, 2008.

[192] Neda Nasiriani, Yaser Fallah, and Hariharan Krishnan. Fairness and Stability Analysis of Congestion Control Schemes in Vehicular Ad-hoc Networks. In *Proceedings of the 10th IEEE Consumer Communications and Networking Conference (CCNC)*, 2012.

[193] Yaser Fallah, Ching-Ling Huang, Raja Sengupta, and Hariharan Krishnan. Analysis of Information Dissemination in Vehicular Ad-Hoc Networks With Application to Cooperative Vehicle Safety Systems. *IEEE Transactions on Vehicular Technology*, 60(1):233–247, January 2011.

[194] Javier Gozalvez and Miguel Sepulcre. Opportunistic Technique for Efficient Wireless Vehicular Communications. *IEEE Vehicular Technology Magazine*, 2(4):33–39, December 2007.

[195] Miguel Sepulcre, Javier Gozalvez, Jérôme Härri, and Hannes Hartenstein. Contextual Communications Congestion Control for Cooperative Vehicular Networks. *IEEE Transactions on Wireless Communications*, 10(2):385–389, February 2011.

[196] Bernhard Kloiber, Jérôme Härri, and Thomas Strang. Dice the TX power - Improving Awareness Quality in VANETs by Random Transmit Power Selection. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2012.

[197] Lars Wischhof and Hermann Rohling. Congestion control in Vehicular Ad Hoc Networks. In *Proceedings of the IEEE International Conference on Vehicular Electronics and Safety*, 2005.

[198] Timo Kosch, Christian Adler, Stephan Eichler, Christoph Schroth, and Markus Strassberger. The Scalability Problem of Vehicular Ad Hoc Networks and How to Solve It. *IEEE Wireless Communications*, 13(5):22–28, October 2006.

[199] Christoph Sommer, Ozan K. Tonguz, and Falko Dressler. Adaptive Beaconing for Delay-Sensitive and Congaestion-Aware Traffic Information Systems. In *Proceedings of the IEEE Vehicular Networking Conference (VNC)*, 2010.

[200] Mohamed Salah Bouassida and Mohamed Shawky. A Cooperative and Fully-Distributed Congestion Control Approach within VANETs. In *Proceedings of the 9th International Conference on Intelligent Transport Systems Telecommunications (ITST)*, 2009.

[201] Jianhua He, Hsiao-Hwa Chen, Thomas Chen, and Wenqing Cheng. Adaptive Congestion Control for DSRC Vehicle Networks. *IEEE Communications Letters*, 14(2):127–129, 2010.

[202] Holger Busche, Chonlatee Khorakhun, and Hermann Rohling. Self-Organized Update Rate Control for Inter-Vehicle Networks. In *Proceedings of the 7th International Workshop on Intelligent Transportation (WIT)*, 2010.

[203] Gaurav Bansal and John Kenney. Achieving Weighted-Fairness in Message Rate-Based Congestion Control for DSRC Systems. In *Proceedings of the 5th IEEE International Symposium on Wireless Vehicular Communications (WiVeC)*, 2013.

[204] Gaurav Bansal, Hongsheng Lu, John Kenney, and Christian Poellabauer. EMBARC : Error Model based Adaptive Rate Control for Vehicle-to-Vehicle Communications. In *Proceedings of the 10th ACM International Workshop on Vehicular Inter-Networking, Systems, and Applications (VANET)*, 2013.

[205] Ryusuke Fukui, Hiroyuki Koike, and Hiromi Okada. Dynamic Integrated Transmission Control (DITRAC) over Inter-Vehicle Communications in ITS. In *Proceedings of the 55th IEEE Vehicular Technology Conference (VTC-Spring)*, 2002.

[206] Martijn van Eenennaam, Georgios Karagiannis, and Geert Heijenk. Towards Scalable Beaconing in VANETs. In *Proceedings of the 4th ERCIM workshop on eMobility*, 2010.

[207] Roberto Baldessari, Damiano Scanferla, and Long Le. Joining Forces for VANETs: A Combined Transmit Power and Rate Control Algorithm. In *Proceedings of the 7th International Workshop on Intelligent Transportation (WIT)*, 2010.

[208] Miguel Sepulcre and Javier Gozalvez. Wireless Vehicular Adaptive Radio Resource Management Policies in Congested Channels. In *Proceedings of the 4th International Symposium on Wireless Communication Systems (ISWCS)*, 2007.

[209] Joseph Camp and Edward Knightly. Modulation Rate Adaptation in Urban and Vehicular Environments: Cross-Layer Implementation and Experimental Evaluation. *IEEE/ACM Transactions on Networking*, 18(6):1949–1962, December 2010.

[210] Miguel Sepulcre and Javier Gozalvez. On the Importance of Application Requirements in Cooperative Vehicular Communications. In *Proceedings of the 8th International Conference on Wireless On-Demand Network Systems and Services (WONS)*, 2011.

[211] Xiaomin Ma, Xianbo Chen, and Hazem Refai. On the Broadcast Packet Reception Rates in One-Dimensional MANETs. In *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, 2008.

[212] Miguel Sepulcre and Javier Gozalvez. Experimental Evaluation of Cooperative Active Safety Applications Based on V2V Communications. In *Proceedings of the 9th ACM International Workshop on Vehicular Inter-Networking, Systems, and Applications (VANET)*, 2012.

[213] Gaurav Bansal, John Kenney, and Aaron Weinfield. Cross-Validation of DSRC Radio Testbed and NS-2 Simulation Platform for Vehicular Safety Communications. In *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, 2011.

[214] Marc Torrent-Moreno, Paolo Santi, and Hannes Hartenstein. Fair Sharing of Bandwidth in VANETs. In *Proceedings of the 2nd ACM International Workshop on Vehicular Ad Hoc Networks (VANET)*, 2005.

[215] Steven W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, 1 edition, January 1997.

[216] David W. Allan, Neil Ashby, and Clifford C. Hodge. The Science of Timekeeping, 1997.

[217] Adrian Lahanas and Vassilis Tsaoussidis. Exploiting the Efficiency and Fairness Potential of AIMD-Based Congestion Avoidance and Control. *Computer Networks*, 43(2):227–245, October 2003.

[218] Lin Cheng, Benjamin Henty, Daniel Stancil, Fan Bai, and Priyantha Mudalige. Mobile Vehicle-to-Vehicle Narrow-Band Channel Measurement and Characterization of the 5.9 GHz Dedicated Short Range Communication (DSRC) Frequency Band. *IEEE Journal on Selected Areas in Communications*, 25(8):1501–1516, October 2007.

[219] Gaurav Bansal, John Kenney, and Charles Rohrs. LIMERIC: A Linear Adaptive Message Rate Algorithm for DSRC Congestion Control. *IEEE Transactions on Vehicular Technology*, 62(9):4182 – 4197, 2013.

[220] Jürgen Kunisch and Jörg Pamp. Wideband Car-to-Car Radio Channel Measurements and Model at 5.9 GHz. In *Proceedings of the 68th IEEE Vehicular Technology Conference*, 2008.

[221] Johan Karedal, Nicolai Czink, Alexander Paier, Fredrik Tufvesson, and Andreas Molisch. Path Loss Modeling for Vehicle-to-Vehicle Communications. *IEEE Transactions on Vehicular Technology*, 60(1):323–328, January 2011.

[222] Alexander Paier, Johan Karedal, Nicolai Czink, Helmut Hofstetter, Charlotte Dumard, Thomas Zemen, Fredrik Tufvesson, Andreas Molisch, and Christoph Mecklenbräuker. Car-to-Car Radio Channel Measurements at 5 GHz: Pathloss, Power-Delay Profile, and Delay-Doppler Spectrum. In *Proceedings of the 4th International Symposium on Wireless Communication Systems*, 2007.