Quality-Driven Cross-Layer Protocols for Video Streaming over Vehicular Ad-Hoc Networks

by

Mahdi Asefi

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2011

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The emerging vehicular ad-hoc networks (VANETs) offer a variety of applications and new potential markets related to safety, convenience and entertainment, however, they suffer from a number of challenges not shared so deeply by other types of existing networks, particularly, in terms of mobility of nodes, and end-to-end quality of service (QoS) provision. Although several existing works in the literature have attempted to provide efficient protocols at different layers targeted mostly for safety applications, there remain many barriers to be overcome in order to constrain the widespread use of such networks for non-safety applications, specifically, for video streaming: 1) impact of high speed mobility of nodes on end-to-end QoS provision; 2) cross-layer protocol design while keeping low computational complexity; 3) considering customer-oriented QoS metrics in the design of protocols; and 4) maintaining seamless single-hop and multi-hop connection between the destination vehicle and the road side unit (RSU) while network is moving.

This thesis addresses each of the above limitations in design of cross-layer protocols for video streaming application. 1) An adaptive MAC retransmission limit selection scheme is proposed to improve the performance of IEEE 802.11p standard MAC protocol for video streaming applications over VANETs. A multi-objective optimization framework, which jointly minimizes the probability of playback freezes and start-up delay of the streamed video at the destination vehicle by tuning the MAC retransmission limit with respect to channel statistics as well as packet transmission rate, is applied at road side unit (RSU). Two-hop transmission is applied in zones in which the destination vehicle is not within the transmission range of any RSU. In the multi-hop scenario, we discuss the computation of access probability used in the MAC adaptation scheme and propose a cross-layer path selection scheme; 2) We take advantage of similarity between multi-hop urban VANETs in dense traffic conditions and mesh connected networks. First, we investigate an application-centric routing scheme for video streaming over mesh connected overlays. Next, we introduce the challenges of urban VANETs compared to mesh

networks and extend the proposed scheme in mesh network into a protocol for urban VANETs. A classification-based method is proposed to select an optimal path for video streaming over multi-hop mesh networks. The novelty is to translate the path selection over multi-hop networks to a standard classification problem. The classification is based on minimizing average video packet distortion at the receiving nodes. The classifiers are trained offline using a vast collection of video sequences and wireless channel conditions in order to yield optimal performance during real time path selection. Our method substantially reduces the complexity of conventional exhaustive optimization methods and results in high quality (low distortion). Next, we propose an application-centric routing scheme for real-time video transmission over urban multi-hop vehicular ad-hoc network (VANET) scenarios. Queuing based mobility model, spatial traffic distribution and probability of connectivity for sparse and dense VANET scenarios are taken into consideration in designing the routing protocol. Numerical results demonstrate the gain achieved by the proposed routing scheme versus geographic greedy forwarding in terms of video frame distortion and streaming start-up delay in several urban communication scenarios for various vehicle entrance rate and traffic densities; and 3) finally, the proposed quality-driven routing scheme for delivering video streams is combined with a novel IP management scheme. The routing scheme aims to optimize the visual quality of the transmitted video frames by minimizing the distortion, the start-up delay, and the frequency of the streaming freezes. As the destination vehicle is in motion, it is unrealistic to assume that the vehicle will remain connected to the same access router (AR) for the whole trip. Mobile IP management schemes can benefit from the proposed multi-hop routing protocol in order to adapt proxy mobile IPv6 (PMIPv6) for multi-hop VANET for video streaming applications. The proposed cross-layer protocols can significantly improve the video streaming quality in terms of the number of streaming freezes and start-up delay over VANETs while achieving low computational complexity by using pattern classification methods for optimization.

Acknowledgments

First and foremost, I would like to express my utmost admiration and deepest gratitude to my Ph.D supervisors, Professor Xuemin (Sherman) Shen and Professor Jon W. Mark, for their continuous guidance, kind support, and great enthusiasm for high quality research. They have provided me with useful weekly guidance, constant encouragement, constructive criticism, and financial support over the course of my Ph.D. It has been a tremendous pleasure to work with them and I wish to take this opportunity to thank them for their constant encouragement and patience.

I would like to thank the members of my dissertation committee: Professors Xianbin Wang, Guang Gong, Zhou Wang , and Xinzhi Liu for having accepted to take the time out of their busy schedules to read my thesis and to provide me with their insightful comments and suggestions.

I have been fortunate to work among members of the Broadband Communications Research (BBCR) Group. I would like to thank former and current BBCR members as they created a very pleasant, intellectually stimulating and friendly environment. I remain grateful to Professor Weihua Zhuang. The knowledge of stochastic processes and communications over fading dispersive channels I learned from her courses, E&CE 604 and E&CE 614, have been the foundation of the research work in this dissertation. Special thanks go to Dr. Ping Wang, Dr. Kuang-Hao (Stanley) Liu, Dr. Humphrey Rutagemwa, Dr. Xinhua Ling and Dr. Lin Cai for their assistance and guidance, many fruitful and valuable discussions especially at the initial stages of my Ph.D studies. I would like to thank my colleagues and friends at VANET/DTN discussion subgroup of BBCR. My deep appreciation goes to Dr. Yanfei Fan, Dr. Bong Jun (David) Choi, Hao Liang, Tom H. Luan, Mianxiong Dong, Khadige Abboud, and other colleagues at BBCR group. Especially, I would like to thank my dear friend Sandra Céspedes for many lively discussions and research collaborations.

I would like to express my gratitude towards Professor Sagar Naik, and Professor Zhou

Wang whose courses in Protocols, Software, and Issues in Mobile System (EC&E 750-T4) and Image Processing and Visual Communications (EC&E 710-T13), respectively, have broadened my mind and helped me to accomplish this dissertation.

I gratefully acknowledge the financial support from Natural Science and Engineering Research Council (NSERC) of Canada, Ontario Research & Development Challenge Fund Bell scholarship, and University of Waterloo Graduate Scholarships. Thanks to Wendy Boles, Annette Dietrich, Karen Schooley, Lisa Hendel, Susan Widdifield, Betty Slowinski and other staffs in Electrical and Computer Engineering Department at University of Waterloo for their friendly attitude and administrative support.

I am forever indebted to my parents, Hossein Asefi and Soheila Hekmat for their endless love, encouragement and support without which I would not have got this far. I dedicate this dissertation to them.

Contents

List of Tables

List of Figures

List of Notations

B	bidirectional coded frame
C_{ij}	channel maximum capacity of link {ij}
$D(S(c,v))$	transmission delay under strategy $S(c,v)$
D_i	transmission delay of segment i of video clip
D_s	playback start-up delay
$D^m_{av,succ}(L,M)$	average time of packet successful transmission
$D^m_{av,unsucc}(L,M)$	average time of packet unsuccessful transmission
$D^m_{av}(L,M)$	average packet transmission time
D_{max}	maximum tolerable delay at application layer
$E(D_s)$	average playback start-up delay
$E(F)$	average frequency of playback freezes
I	intra-coded frame
L	packet payload length
$M_{ij}(s)$	moment generating function of t_{ij}
M_{ij}	retransmission limit of link $(i \rightarrow j)$
M_i	MAC retransmission value for all the packet of the ith video segment
M	set of all admissible MAC retransmission values
N_{APP}	total APP later strategies
N_{MAC}	total MAC later strategies
N_{NET}	total NET later strategies

N_{PHY}	total PHY later strategies
N	set of nodes in network
O	packet overhead length
$P_{ij}^{D}(M_{ij}, p_{ij}(x))$	link packet drop probability
$P_{ij}^{T}(M_{ij}, p_{ij}(x))$	overall packet loss probability
$P_a(x)$	probability that the vehicle is connected either to RSU1 or RSU2
$P_{e,ack}$	ACK packet error probability
$P_{e,data}$	data packet error probability
$P_{ij}^{B}(M_{ij}, p_{ij}(x))$	link blocking probability
$P_{succ}^{m}(L, M)$	probability of successful packet transmission within the M retries
P	inter-coded frame
$Q(x)$	tail probability of the standard normal distribution
Q	multimedia quality
$R(S(c, v))$	Rate of cross-layer strategy $S(c, v)$
R_{max}	maximum channel rate
R	transmission range of the RSU
$S(c, v)$	cross-layer strategy
$S^{OPT}(c, v)$	optimal cross-layer strategy
S	Set of video packets in the network
T_c	channel coherence time
T_{CCH}	length of CCH interval
T_{SCH}	length of SCH interval
T_{SYNC}	length of SYNC interval
T_{bad}^{m}	time of transmission with error
T_{good}^{m}	time of transmission without error

U	hue component
V_r	vectorial velocity of the receiver
V_s	vectorial velocity of the sender
$Var(D)$	variance of the playback start-up delay
$Var(F)$	variance of frequency of playback freezes
V	intensity component
$Y(t)$	RSS at time t
Y_{SNR}	average SNR of the received signal
Y	luminance component
b	playback threshold
c	channel conditions
d_0	reference distance
f_0	center frequency of the signal
f_m	maximum Doppler frequency shift
$f_{ij}(x)$	probability density function of t_{ij}
l	Euclidean distance between RSU and the destination vehicle
m	PHY mode
$n(M_{ij}, p_{ij}(x))$	mean packet retransmissions to be successfully received by node j
$p_1(x)$	probability of vehicle being directly connected to either RSU1 or RSU2
$p_2(x)$	probability of vehicle being connected to at least one connected vehicle
$p_{ij}(x)$	packet error probability for link $(i \rightarrow j)$
p_{ij}	packet error probability of the link $\{ij\}$
p_{rx}	received signal power at the reference distance
p_{th}	received signal power threshold
p_{tx}	received signal power at the destination

t time instant

v_a variance of inter-arrival rate of packets at the destination buffer

v_s variance of service interval at the destination buffer

v video codec and content characteristics

x distance of the vehicle from the reference

List of Abbreviations

AP	access probability
AQPS	asynchronous quorum-based power-saving
AR	access router
ARQ	automatic retransmission request
BCMP	Baskett, Chandy, Muntz, Palacios
CBR	constant bit rate
CCH	control channel
CCW	cooperative collision warning
CH	cluster head
CIF	common intermediate format
CM	cluster member
CMDD	Content, map or database download
CVIA	controlled vehicular Internet access
CVW	cooperative violation warning
CW	contention window
DCT	discrete cosine transform
DORA	dynamic optimal fragmentation with rate adaptation algorithm

DSRC	dedicated short range communications
E2EJ	end-to-end jitter
EDCA	enhanced distributed channel access
EEBL	emergency electronic brake light
ETSI	European Telecommunications Standards Institute
FBM	fractional Brownian motion
FCFS	first-come-first-serve
FCR	fast collision resolution
FEC	forward error correction
FMIPv6	fast handover for mobile IPv6
GOP	group of pictures
GPS	global positioning system
HMM	hidden Markov model
ITS	intelligent transportation system
LMA	local mobility anchor
LRD	long range dependence
MAC	medium access control
MAG	mobile access gateway

MANET	mobile ad-hoc network
MIH	Media Independent Handover
MN	mobile node
MPEG	moving picture experts group
MR	mobile router
MSE	mean square error
NEMO	network mobility
OBU	on board unit
OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiple access
OSI	open system interconnection
PAN	parking availability notification
PCN	V2V post crash notification
PMIPv6	proxy mobile IPv6
PRMA	packet reservation multiple access
PSL	parking spot locator
PSNR	peak signal-to-noise ratio
QCIF	quarter common intermediate format
QoS	quality of service
RBVT	road-based vehicular traffic

RFN	road feature notification
RHCN	road hazard condition notification
RS	Reed-Solomon
RSS	received signal strength
RSU	road side unit
RTT	round-trip time
RTVR	Real-time video relay
RVP/D	remote vehicle personalization diagnostics
SA	service announcement
SCH	service channel
SNR	signal-to-noise ratio
SVA	stopped or slow vehicle advisor
SVM	support vector machine
TCP	transport control protocol
TDMA	time division multiple access
UDP	user datagram protocol
V2I	vehicle-to-infrastructure
V2V	vehicle-to-vehicle
VANET	vehicular ad-hoc network
VBR	variable bit rate
VCF	video content features

VEP video encoding parameters

VoD video on demand

VoIP voice over Internet protocol

WAVE wireless access in vehicular environments

WCC wireless channel conditions

WLAN wireless local area network

WSA WAVE service advertisement

Chapter 1

Introduction

Vehicular transportation is one of the crucial means of transportation. To enhance the transportation safety, communication-based safety applications empowered by vehicular ad-hoc networks (VANETs) have recently attracted attention from industry and academia. VANETs belong to a general class of mobile ad-hoc communication networks with vehicles acting as fast moving mobile nodes. More specifially, a VANET consists of 1) on board units (OBUs) installed on the vehicles 2) roadside units (RSUs) deployed along sides of the urban roads/highways which facilitate both vehicle-to-vehicle (V2V) communications between vehicles and vehicle-to-infrastructure (V2I) communications between vehicles and RSUs. For example, traffic information can be exchanged via V2V communications through periodic beaconing to allow drivers to be better aware of surrounding traffic conditions. Peer-to-peer applications such as information sharing and gaming can also be supported through V2V communications [2]. In the presence of RSUs, not only safety messages can be broadcast to the driver, but drive-thru Internet can be delivered to a vehicle via Internet backbone. Through V2I communications, infotainment services such as advertisements, parking lot availability, and automatic tolling can be also provided

with ease.

The Federal Communications Commission (FCC) in the United States has allocated 75 MHz of licensed spectrum at 5.9 GHz as the dedicated short range communications (DSRC) band for intelligent transportation systems (ITSs) to enable communication-based safety and infotainment services, . In Europe, different frequency bands are used for vehicular communications, for instance, unlicensed frequency band at $2010-2020$ MHz is used in Fleetnet. Recently, the European Telecommunications Standards Institute (ETSI) has also allocated a radio spectrum of 30 MHz at 5.9 GHz for ITSs. To increase spectrum utilization and improve QoS (e.g., network throughput, packet dropping rate, end-to-end delay, fairness, etc.), multiple channels are expected to be employed in vehicular communications. Location information of vehicles is generally available, thanks to the global positioning system (GPS). End-to-end paths for information delivery can then be established via location-aware V2V and/or V2I transmission. In short, this emerging vehicular networking paradigm is expected to enable a plethora of communication-based automotive applications in the near future, ranging from seamless inter-vehicle video streaming to road traffic monitoring to collision warning/avoidance.

1.1 Research Challenges in VANETs

Despite the fact that VANETs are organized in an ad-hoc manner, they are very different from traditional mobile ad-hoc networks (MANETs) in terms of network architecture, mobility pattern, energy constraint, and application scenarios. It has been demonstrated in the literature that directly applying approaches designed for MANETs, does not result in effective and/or efficient performance. To succeed in vehicular environment, it is imperative to devise new approaches directly tailored for VANETs. Some of the main challenges for communication over VANETs can be cited as follows.

- Frequent link disconnection: The topology of the network is changing rapidly in VANET which can result in disconnection in communication links. Specially, vehicles move with high speed on highways may render the communication link between two vehicles to last for only few minutes. If the vehicles move in opposite direction, the connection time can be reduced to only few seconds. Thus, connectivity analysis and mobility-aware resource management are important.

- Highly dynamic spatial-temporal traffic conditions: The spatial traffic density can vary from very sparse (e.g., rural areas) to very dense (e.g., traffic jams). The temporal traffic density in one area can be variant depending on the time of the day. Penetration rate, which is defined as the percentage of the vehicle on the road equipped with an OBU is another factor which has impact on the vehicular density, especially in the early stages of VANET deployment. Another concern is the impact of the dynamic traffic variations on the wireless channel impairments such as slow and fast fading. Channel conditions can vary greatly in temporal and spatial domains and therefore, adaptive cross-layer channel protocols with resistance to channel impairments are imperative.

- Heterogeneity of data dissemination: VANETs are supposed to support a wide variety of applications ranging from safety to infotainment to entertainment applications. Road safety applications need low latency and high reliability whereas resource utilization, packet loss and fairness are the common performance measures for infotainment applications. Internet applications such as video streaming require high throughput and short delay in order to satisfy the users. In light of heterogeneous information services, channel access protocols and network resource allocation strategies should be adaptive.

It should be noted that link-layer protocols and techniques can not completely solve the aforementioned VANET challenges. A holistic solution should be formulated by considering the features of different layers in VANET protocol stack [2].

1.2 Research Challenges and Motivations for Video Streaming over VANETs

Video streaming to vehicles is no longer a luxury application due to existence of wireless networks, such as wireless local area network (WLAN), WiMax, and 3G cellular networks which enable high-rate video access from anywhere. The vehicle engine provides sufficient power for intensive data processing and communications. The on-board buffer storage, positioning system, and intelligent antenna further facilitate the efficient video forwarding and collaborative downloading among vehicles or from/to RSUs. ITSs for VANET have stimulated the development of several interesting applications such as vehicle collision warning, security distance warning, driver assistance, cooperative driving, and cooperative cruise control [3]. The inherent characteristics of vehicular networks, such as dynamic topology and high mobility, pose challenging conditions for the deployment of delay-sensitive video streaming applications.

While connectivity to the Internet has become an essential part of our daily life, access to high-rate Internet from vehicles is a luxury application in most areas. In cellular or satellite communications, the available rate for each user is far from enough to deliver the media-rich Internet contents. Nearly 10% of the waking time of the American people is spent in vehicles in which people have very limited or no access to Internet [4].

Technical challenges arise when the unpredictable nature of the wireless vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) radio channels meet the requirements

of the high data rate and low latency for video transport. As the wireless link quality varies, video transmission rate needs to be adapted accordingly. The video streams containing different contents also derive different utility from a change in allocated rate. As a consequence, the same allocated rate over a fast link would require lower fraction of channel time than over a slow link; the same increase in allocated rate may benefit a hard stream containing complex motion more than another easy stream with little or no motion. Hence, under adverse conditions, wireless stations need to optimally adapt their multimedia compression and routing strategies jointly across the protocol stack in order to guarantee a predetermined quality at the receiver by cross-layer exchanging of information between different open system interconnection (OSI) protocol layers. In this thesis, we study and propose different cross-layer protocols for seamless delivery of video packets in urban VANET scenarios.

1.3 Research Contributions

- MAC-centric cross-layer design: an adaptive medium access control (MAC) retransmission limit selection scheme is proposed to improve the performance of IEEE 802.11p standard MAC protocol for video streaming applications over VANET. A multi-objective optimization framework, which jointly minimizes the probability of playback freezes and start-up delay of the streamed video at the destination vehicle by tuning the MAC retransmission limit with respect to channel statistics as well as packet transmission rate, is applied at RSU. Periodic channel state estimation is performed at the RSU using the information derived from the received signal strength (RSS) and Doppler shift effect. Estimates of access probability between the RSU and the destination vehicle is incorporated in the design of the adaptive MAC scheme. The adaptation parameters are embedded in the user datagram

protocol (UDP) packet header. Two-hop transmission is applied in zones in which the destination vehicle is not within the transmission range of any RSU. In the multi-hop scenario, we discuss the computation of access probability used in the MAC adaptation scheme and propose cross-layer path selection scheme followed by a discussion on mobile IP management to maintain continuous video streaming. Compared with the non-adaptive IEEE 802.11p standard MAC, numerical results show that the proposed adaptive MAC protocol exhibits significantly fewer playback freezes while introduces only a slight increase in start-up delay.

- Network-centric cross-layer design: service-oriented vehicular networks face challenge to deliver delay-sensitive data such as video packets. Most research works on video streaming consider network-centric QoS metrics rather than the user perceived quality. Multi-hop urban vehicular networks in dense traffic conditions have similarities with mesh connected networks plus new challenges arising from highly dynamic nature of the network. First, we investigate an application-centric routing scheme for video streaming over mesh connected overlays. Next, we introduce the challenges of urban VANETs compared to mesh networks and extend the proposed scheme in mesh network to a protocol for urban VANETs. A classification-based method is proposed to select an optimal path for video streaming over multi-hop mesh networks. Our main contribution is the translation of path selection over multi-hop networks to a standard classification problem. The classification is based on minimizing average video packet distortion at the receiving nodes. We also consider long range dependence (LRD) characteristics of variable bit rate (VBR) video encoders by modeling the link as a fractional Brownian motion (FBM) queue. In our method, a support vector machine is applied at each hop to optimally assign the next hop (partial path) for each video packet received at that hop. Wire-

less channel conditions (WCC), including packet loss probability of the channel and maximum achievable rate, are used as class prototypes. Sample feature vectors include video content features (VCF) and video encoding parameters (VEP) extracted from video sequences. The classifiers are trained offline using a vast collection of video sequences and wireless channel conditions in order to yield optimal performance during real time path selection. Our method substantially reduces the complexity of conventional exhaustive optimization methods and results in high quality (low distortion). Simulations are conducted over an elementary multi-hop structure and cascades of such structure which not only proves the superiority of our method in terms of low complexity and high peak signal-to-noise ratio (PSNR) performance, but also provides important insights that can guide the design of network infrastructures and streaming protocols for video streaming. Next, we will extend the mentioned routing protocol to an application-centric routing framework for real-time video transmission over urban multi-hop VANET scenarios. Queueing based mobility model, spatial traffic distribution and probability of connectivity for sparse and dense VANET scenarios are taken into consideration in designing the routing protocol. Numerical results demonstrate the gain achieved by the proposed routing protocol versus geographic greedy forwarding in terms of video frame distortion and streaming start-up delay in several urban communication scenarios for various vehicle entrance rate and traffic densities.

- Routing+IP management: The inherent characteristics of vehicular networks, such as dynamic topology and high mobility, pose challenging conditions for the deployment of delay-sensitive applications, e.g., video streaming. We first propose a new quality-driven routing scheme for delivering video streams from a fixed network to a destination vehicle via multi-hop communications in urban VANET scenarios. The

routing scheme aims to optimize the visual quality of the transmitted video frames by minimizing the distortion, the start-up delay, and the frequency of the streaming freezes. As the destination vehicle is in motion, it is unrealistic to assume that the vehicle will remain connected to the same AR for the whole trip, thus we further propose an efficient mobility management protocol. The proposed network mobility management scheme incorporates a handover prediction mechanism, and is a novel adaptation of PMIPv6 for multi-hop VANET. Numerical results show that our routing scheme achieves less frame distortion, lower start-up delay and fewer streaming freezes compared to the greedy geographical routing protocol. Also, the proposed multi-hop PMIPv6 achieves lower handover delay with less signalling cost when compared to the MANET-centric network mobility (NEMO) protocol.

1.4 Outline of the Thesis

Chapter 2 provides some background on video quality measurement at different levels (trace level, frame level, networking and streaming level), cross layer optimization, and challenges for quality-driven cross-layer protocols for video streaming over VANETs. This is followed by literature review of different cross-layer protocols at diffract layers and their advantages and drawbacks.

In **chapter 3**, an adaptive MAC retransmission limit in which the adaptation is based on an optimization of playback streaming quality is proposed. A multi-objective optimization framework is applied at the RSU, which jointly minimizes the probability of playback freezes and start-up delay of the streaming at the destination vehicle by tuning the MAC retransmission limit with respect to channel statistics (packet error rate and packet transmission rate).

In **chapter 4**, a novel application-centric multi-hop routing protocol for video transmis-

sion in urban VANET scenarios is proposed. We have translated the design of routing protocol into an optimization problem using a probabilistic model to formulate average end-to-end distortion of the delivered video frames over the entire video session. Performance of the protocol in different vehicular scenarios versus increment in coverage range through more communication hops has been evaluated. In addition, analytical and simulation results are provided which show that the proposed application-centric routing protocol achieves lower start-up delay and frame distortion compared to the traditional greedy forwarding protocol. Decrement in the size of search space at each hop, translated the NP-hard optimization problem to a polynomial time which can be implemented.

In **chapter 5**, we have proposed a novel network layer cross-layer protocol that integrates a quality-driven geo-routing protocol with a PMIPv6-based multi-hop scheme for seamless video streaming in urban vehicular scenarios. The integrated scheme takes advantage of the geographic features in VANET for data delivery. On one hand, the geo-routing protocol considers important metrics for video quality such as distortion, start-up delay and frequency of freezes, so that the multi-hop path is selected accordingly. On the other hand, the multi-hop PMIPv6 scheme provides seamless delivery of data packets during the handovers, and incorporates a prediction mechanism based on the vehicle location. Numerical results have demonstrated that the proposed quality-driven routing protocol achieves a lower start-up delay, number of streaming freezes and frame distortion, when compared to the traditional greedy protocol. Similarly, the multi-hop PMIPv6 scheme achieved a reduced handover latency and signalling cost, due to the handover prediction mechanism.

Chapter 6 presents a summary of the thesis contributions and discusses several future research directions.

Chapter 2

Background and Related Work

2.1 Network Performance Evaluation of Single Layer and Multi-Layer Video

Video consists of a sequence of pictures or images called a *frame*. The reciprocal of the frame rate is the display time of a frame on the screen, known as *frame period*. Each frame consists of elements called *pixels* or *pels*. The *frame format* specifies the size of the frame in terms of pixels. For example, common intermediate format (CIF) has 352×288 pixels, and the quarter common intermediate format (QCIF) has 176×144 pixels. CIF and QCIF are usually considered in network research [5]. Each pixel is represented by three components: the luminance component (Y), and the two chrominance components, hue (U) and intensity (V). Since human visual system is more sensitive to luminance information than color information, the chrominance components are typically sub-sampled to one set of U and V samples per four Y samples. Thus, with chroma sub-sampling there are 352×288 Y samples, 176×144 U samples and 176×144 V samples in each CIF Video frame. Each sample is typically quantized into 8 bits, resulting in frame size

of 152,064 bytes per uncompressed CIF frame.

2.1.1 Principles of Non-Scalable Video Encoding

We give a brief discussion on the *non-scalable* or single layer video compression which is employed in moving picture experts group (MPEG) and H.26x standards. The two main principles of these standards are *intra-frame* coding using discrete cosine transform (DCT) and *inter-frame* coding using motion estimation and compensation between successive video frames. In intra-frame coding each frame is divided into blocks of Y samples, U samples and V samples. Each block is transformed into 8×8 block of DCT coefficients. A larger quantization scale means coarser quantization which results in fewer bits per sample. Coefficients less than a predefined threshold are set to zero and removed. The quantized coefficients are then zigzag scanned, run-level coded, and variable-length coded to achieve further compression.

In inter-frame coding, MPEG introduces three types of frame: intra-coded (I), inter-coded (P) and bidirectional coded (B). Similar frame types exist in H.26x video coding. They are organized into groups called group of pictures (GOP). By definition, the sequence of frames from given I frame up to and including the next I frame is referred to as one GOP. The pattern of I, P, and B frames that make up a GOP is called *GOP structure*. A typical GOP pattern is shown in Figure 2.1.

In an I frame, all blocks are intra-coded as outlined above. In a P frame the macro-blocks are inter-coded using motion compensation with reference to the preceding I or P frame. In a B frame the macro-blocks are inter-coded with reference to the preceding I or P frame, which serves as forward reference, and the succeeding I or P frame, which serves as backward reference, as illustrated by the dashed arrows in Figure 2.1.

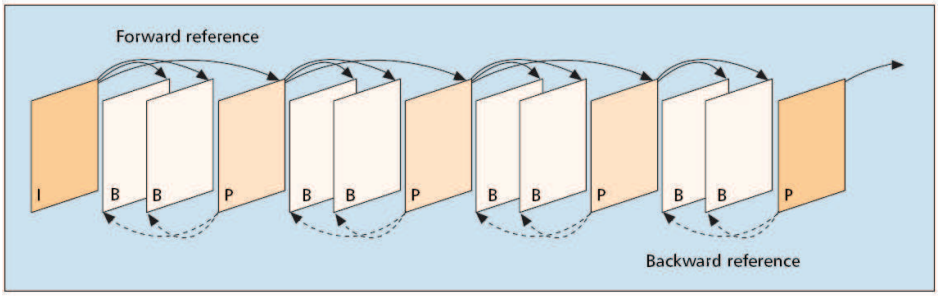


Figure 2.1: Typical MPEG group of pictures (GOP) pattern with reference used for predictive coding of P and B frames.

2.1.2 Principles of Scalable Video Encoding

With conventional layered encoding, the video is encoded hierarchically into a base layer and one or more enhancement layers. Decoding of base layer gives basic video quality while decoding of base layer together with one or more enhancement layers gives enhanced video quality. MPEG has introduced different types of scalability: spatial, temporal, signal-to-noise ratio (SNR) and data partitioning.

With temporal scalable encoding, the enhancement layer frames are interleaved between base layer frames. Each enhancement layer frame is inter-coded with reference to the immediately preceding base layer frame and the immediately succeeding base layer frame (as illustrated in Figure 2.2(a)) for a scenario where I and P frames form the base layer and B frames form the enhancement layer. The base layer of the temporal scalable encoding provides a basic video quality with a low frame rate. Adding the enhancement layer to the base layer increases the frame rate.

With spatial scalability the base layer provides a small video format (e.g., QCIF); adding the enhancement layer increases the video format (e.g., CIF). The base layer of the spatial scalable encoding can be up-sampled to give a coarse video at the larger format. To perform a spatial scalable encoding, the original (uncompressed) video is first

downsampled to the smaller base layer format and the downsampled video is encoded employing the intra and inter coding techniques described above. A base layer consisting of only I and P frames is illustrated in Figure 2.2(b).

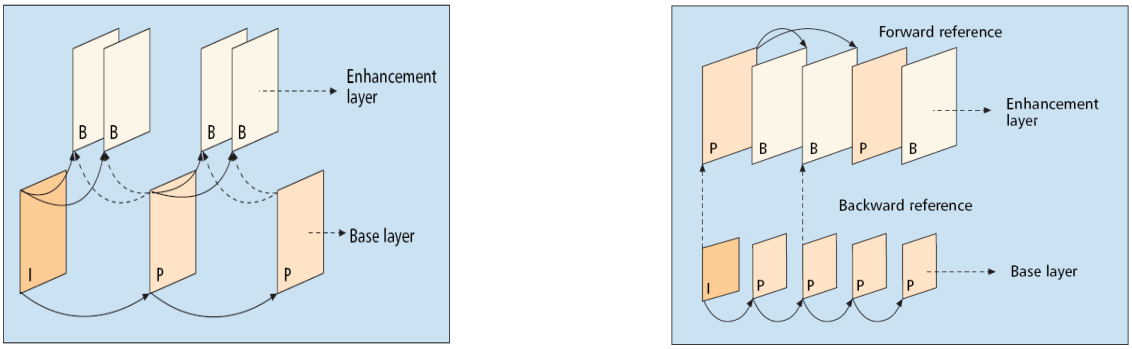


Figure 2.2: (a) Example for temporal scalable encoding: I and P frames form the base layer and B-frames form the enhancement layer. (b)Example for spatial scalable encoding. The downsampled video is encoded into a base layer stream consisting of I and P frames. The difference between the decoded and upsampled base layer and the original video is encoded into the P and B frames in the enhancement layer.

2.1.3 Performance Evaluation of Video in Networking Research

Generally there are three different methods to characterize encoded video for the purpose of networking research: video bit stream, video traffic trace, and video traffic model. The advantage of the bit stream is that it allows to conduct networking experiments where the quality of the video after suffering losses in the network is evaluated. One limitation of the bit stream is that it is very large in size (several gigabytes for one hour of compressed video or several tens of gigabytes for one hour of uncompressed video). Another limitation of bit streams is that they are usually proprietary and/or protected by copyright. This limits the access of networking researchers to bit streams, and also limits the exchange of bit streams among research groups [5]. Video traces are an alternative to

bit streams. While the bit streams give the actual bits carrying the video information, the traces only give the number of bits used for the encoding of the individual video frames. A traffic model is typically developed based on the statistical properties of a set of video trace samples of the real video traffic. The developed traffic model is verified by comparing the traffic it generates with the video traces. If the traffic model is deemed sufficiently accurate, it can be used for the mathematical analysis of networks, for model driven simulations, and also for generating so called virtual (synthetic) video traces.

There are three fundamental factors to be considered when using video traces in simulations [5]: 1) the definition of video related performance metrics; 2) the generation of traffic workload for system under study; and 3) statistically sound estimation of performance metrics of interest. The first step in evaluating the performance of the network is generating appropriate traffic (load) from traces. These issues range from picking and preparing the video streams (traces) to the packetization of the video frames. We first investigate issues at stream level and then turn to individual video packets and frames. At the stream level, we first need to select the videos used in evaluation. Generally it is advised to cover as many videos as possible in order to cover all the possible video features likely to be encountered in the considered networking scenarios [6]. [7] provides many examples for scalable and non-scalable encoders. The video-traffic trace is an abstraction of real-video stream. It typically gives the frame number, frame type (I,B,P), and frame size in a text file to describe the characteristics of real-video traffic. We generate sender trace file and receiver trace file to record the details of packets during the simulation. Finally, by using the trace files generated during simulation process and the encoded videos trace generation step, we can regenerate received video. After decoding the constructed video, the transmission can be evaluated by calculating quality metrics such as PSNR.

2.2 Cross-Layer Optimization

2.2.1 Cross Layer Problem Formulation

According to [8], cross layer design can be formulated as an optimization problem which has as objective the joint selection of strategies across multiple OSI layers. Assuming that the joint cross layer strategy is

$$\begin{aligned} S(c,v) = \{&PHY_1, ..., PHY_{N_{PHY}}, MAC_1, ..., MAC_{N_{MAC}}, \\ &NET_1, ..., N_{N_{NET}}, APP_1, ..., APP_{N_{APP}}\} \end{aligned} \tag{2.1}$$

where N_{PHY}, N_{MAC}, N_{NET}, N_{APP} are the total PHY, MAC, NET and APP layer strategies. The cross-layer optimization problem attempts to find the optimal composite strategy $S^{OPT}(c,v)$ which maximizes the multimedia quality Q:

$$S^{OPT}(c,v) = \arg\max Q(S(c,v)) \tag{2.2}$$

subject to the following constraints on the transmission rate and delay:

$$\begin{aligned} R(S(c,v)) &\leq R_{max} \\ D(S(c,v)) &\leq D_{max} \end{aligned} \tag{2.3}$$

In the above equations, c represents the underlying channel conditions, and v is derived based on video codec and content characteristics. $R(S(c,v))$ and $D(S(c,v))$ are rate and delay of the transmission under strategy $S(c,v)$. R_{max} is the maximum transmission bit-rate and D_{max} is the maximum tolerable delay at the application layer.

2.2.2 Challenges in Solving the Cross-Layer Optimization

WLANs have gained high attention because of the proliferation of technologies such as Bluetooth, IEEE 802.11, 3G, and WiMAX. Due to flexible and cost efficient infrastructure of WLANs, they are widely used to carry out a variety of delay sensitive multimedia streaming applications such as video-conferencing, emergency services, tele-medicine, and surveillance monitoring. In particular, IEEE 802.11 WLAN has emerged as a prevailing technology for (indoor) broadband wireless access because it supports real-time conversational multimedia applications like voice over Internet protocol (VoIP) and video streaming [9]. The IEEE 802.11p standard for communication over VANET closely follows the previously proposed IEEE 802.11e attributes.

VANETs provide dynamically varying medium with limited support for QoS required by delay sensitive, bandwidth intense and loss tolerant multimedia applications such as video streaming. Hence, applications must be adaptive to resources available in order to present the best service quality possible to the users. Adaptation represents the ability of network protocols and applications to observe and respond to the channel variation. Central to the adaptation is the concept of cross-layer design which entails exchanging of the parameters between different OSI layers [10]. Cross-layer optimization for video streaming over VANETs is a challenging task for mainly the following reasons [11].

- Analytical derivation of relation between delay, rate and quality is difficult and sometimes only statistical relations can be established.
- Correlation exists between strategies applied at each layer and other strategies applied at the same or different layers.
- Channel characteristics and multimedia traffic vary dynamically.

- Different applications have very diverse QoS requirements in terms of data rates, delay bounds, and packet loss probabilities.
- Heterogeneity among end users exists in terms of latency, video visual quality, processing capabilities, power, and bandwidth. It is thus a challenge to design a delivery mechanism that not only achieves efficiency in network bandwidth, but also meets the heterogeneous requirements of the end users.
- Mobility of vehicles which impose challenge on having continuous connection to the RSUs.
- Market penetration rate is another important issue. Most of the protocols designed specifically for VANET scenarios need information about the location of the neighboring nodes, speed and direction of movement, etc., which will not be accessible without OBUs installed on the vehicles.

To understand the possible impact of cross-layer optimization, let us first analyze the impact of different layers on throughput efficiency and delay performance presented in [12]. Assume that polling-based mode of 802.11a MAC standard is used for video transmission. To protect video, adaptive retransmission at MAC and Reed-Solomon (RS) codes at the application layer is considered in addition to the PHY layer modulation and coding strategies provided by 802.11a. The packet's length is L bytes and overhead is O bytes. We first analyze average transmission duration of a MAC frame under different conditions. This will be used later to compute the application layer throughput efficiency. Assume that a packet with L bytes is transmitted using PHY mode m. The probability of successful transmission is given by

$$P^m_{good\ cycle}(L) = (1 - P^m_{e,ack})(1 - P^m_{e,data}) \tag{2.4}$$

where $P_{e,ack}$ is the ACK packet error probability and $P_{e,data}$ is the data packet error probability. These can be calculated from corresponding packet sizes. The average transmission duration for packet of size L-byte payload, given the transmission is successful with transmission limit of M, can be obtained as

$$D^m_{av,succ}(L,M) = \sum_{i=0}^{M} \frac{P^m_{succ}(i|L)}{P^m_{succ}(L,M)} [iT^m_{bad}(L) + T_{good^m}(L)] \tag{2.5}$$

where T^m_{good} is the time of transmission when neither data packet nor the ACK packet is in error and T^m_{bad} is the transmission time when either payload or ACK is in error. The probability that the packet with L bytes data payload is transmitted successfully after the ith retransmission, using PHY mode m, is given by

$$P^m_{succ}(i|L) = [1 - P^m_{good-cycle}(L)]^i P^m_{good-cycle}(L) \tag{2.6}$$

The probability that the packet with L byte payload is transmitted successfully within the M retransmission limit under PHY mode m is given by

$$P^m_{succ}(L,M) = 1 - [1 - P^m_{good-cycle}(L)]^{M+1} \tag{2.7}$$

The average transmission duration for a packet with L-byte payload, given that the transmission is not successful with the retransmission limit M, is

$$D^m_{av,unsucc}(L,M) = (M+1)T^m_{bad}(L) \tag{2.8}$$

The average transmission duration for a packet with L byte payload and with a retrans-

mission limit of M is

$$D_{av}^{m}(L,M) = D_{av,succ}^{m}(L,M)P_{succ}^{m}(L,M) + D_{av,unsucc}^{m}(L,M)(1-P_{succ}^{m}(L,M)) \tag{2.9}$$

The throughput efficiency of 802.11a with the use of the (N|K) RS code at the application layer can be computed based on average packet duration. The RS decoder can correct up to $N-K$ packet erasures. If there are more than $N-K$ packet erasures, then this results in a decoding failure. Therefore, the probability of error after RS decoding is

$$P_{RS}^{m} = 1 - \sum_{i=0}^{N-K} \binom{N}{i} (P_r^m)^i (1-P_r^m)^{N-i} \tag{2.10}$$

where the resulting residual error probability P_r^m of the data packet after Z retransmissions is

$$P_r^m = 1 - P_{succ}(L,Z) \tag{2.11}$$

When a decoding failure happens, there are $N-i(i<K)$ correctly received packets, including both video and parity packets. These video packets can be utilized for video decoding and, on average, $(K/N)(N-i)$ packets out of $N-i$ correctly received packets are video packets. Therefore, the throughput efficiency, taking into account the APP layer RS coding and header overheads of higher layers, is [13]

$$E_{RS}^{m}(L_a,Z,N,K) = \frac{8L_a(K(1-P_{RS}^{m}) + \sum_{i=N-K}^{N}(N-i)\frac{K}{N}\binom{N}{i}(P_r^m)^i(1-P_r^m)^{N-i})}{ND_{av}^{m}(L_a+O,Z)DZ(m)} \tag{2.12}$$

Now, we discuss the interaction between various transmission strategies deployed at different OSI layers and their impact on video quality. Consider determining the optimal

packet size that should be selected at the application layer to maximize the video quality Q for a given RS code and retransmission limit. Based on computed throughput efficiency E_{RS}^{m} that takes into account the application-layer RS coding and the header overheads of the higher layer protocols, the associated video quality $Q_{E_{RS}^{m}}$ can be computed using analytical rate distortion model. The optimal rate packet size L_a can be computed by:

$$L_a^* = arg \max_{L_a} Q_{E_{RS}^{m}} \tag{2.13}$$

2.2.3 Classification of Cross-Layer Solutions

To gain further insight into the cross-layer optimization problem and to compare various proposed cross layer approaches, a classification of cross layer solutions based on structural view is provided in [8] as follows:

1. **Top-down approach:** The higher-layer protocols optimize their parameters and the strategies at the next lower layer. This cross-layer solution has been deployed in most existing systems, wherein the APP identifies the MAC parameters and strategies, while the MAC selects the optimal PHY layer modulation scheme.

2. **Bottom-up approach:** The lower layers try to insulate the higher layers from losses and bandwidth variations. This cross-layer solution is not optimal for multimedia transmission, due to the incurred delays and unnecessary throughput reductions.

3. **Application-centric approach:** The APP layer optimizes the lower layer parameters one at a time in a bottom-up (starting from the PHY) or top-down manner, based on its requirements. However, this approach is not always efficient, as the APP operates at slower timescales and coarser data granularity (multimedia flows

or group of packets) than the lower layers (bits or packets), and hence is not able to instantaneously adapt their parameters to achieve an optimal performance.

4. **MAC-centric approach:** In this approach the APP layer passes its traffic information and requirements to the MAC, which decides which APP layer packets/flows should be transmitted and at what QoS level. The MAC also decides the PHY layer parameters based on the available channel information. The disadvantage of this approach resides in the inability of the MAC layer to perform adaptive source channel coding trade-offs given the time-varying channel conditions and multimedia requirements.

5. **Integrated approach:** In this approach, strategies are determined jointly. Unfortunately, exhaustively trying all the possible strategies and their parameters in order to choose the composite strategy leading to the best quality performance is impractical due to the associated complexity. A possible solution to solve this complex cross-layer optimization problem, in an integrated manner, is to use learning and classification techniques. To the best of our knowledge only the work in [11] has done preliminary work in this area which is the inspiration of our proposed research. Success in this area would lead very close to optimal performance with low complexity which would make it possible for deployment in commercial systems.

2.2.4 Commercial and Public Use of Vehicular Ad-Hoc Networks

From a value or customer benefit perspective, the communication-based automotive applications can be classified into three major categories: 1) safety-oriented, 2) comfort-oriented, and 3) commercial-oriented. These applications utilize both the vehicle-to-vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication modes. Among listed

categories of applications, safety-oriented applications received more attention due to their importance in reducing fatalities and economic losses caused by traffic crashes. Some of the applications of interest in this category include stopped or slow vehicle advisor (SVA), emergency electronic brake light (EEBL), V2V post crash notification (PCN), road hazard condition notification (RHCN), road feature notification (RFN), cooperative collision warning (CCW), cooperative violation warning (CVW) [14].

Convenience-oriented applications share traffic information among vehicles, RSUs and central traffic control systems in order to enable more efficient traffic flow. These applications improve the driving experience indirectly by maintaining roadways and organizations which maintain them more efficient. Some of the applications include free flow tolling scheme which enables roadways and congestion pricing via non-stop tolling (TOLL), parking availability notification (PAN) and parking spot locator (PSL). In Japan, it was found that 30% of congestion was due to toll gates. a reduction in roadway congestion has direct benefit to society in terms of fuel consumption and CO_2 emissions.

Commercial-oriented applications provide entertainment, higher productivity and satisfaction through various communication services such as web access and audio/video streaming. Commercial features may have the potential to create substantial demand from consumers since they span many aspects of a consumer's daily life. Some of the services include news, updating weather and traffic conditions, Internet and email access, and drive-through payments at fast food restaurants, car washes or parking garages which benefit the customers in terms of time saved. Some commercial applications are described as follows.

Remote vehicle personalization diagnostics (RVP/D) This allows drivers to personalize the vehicle settings remotely. For example, the driver may initiate a wireless connection of vehicle to the home network and download (upload) the latest

personalized vehicle settings.

Service announcement (SA) Businesses (e.g. restaurants) may use RSUs to announce services to vehicles as they pass within wireless communication range of RSUs.

Content, map or database download (CMDD) Within proximity of a hot-spot, the driver may initiate connection to the wireless connection of the hot-spot and download content (e.g. map, multimedia, web pages) to the hard drive of the radio/navigation system.

Real-time video relay (RTVR) A vehicle may initiate transmission of real-time video that may be useful to other drivers in the area (e.g. traffic jam scene). Other vehicles may display the information to their drivers or relay the real-time video via multi-hop broadcasts in order to extend the range of the communication to other vehicles.

2.2.5 Communication Patterns of VANETs

The enhanced version of IEEE 802.11 networks, which is IEEE 802.11p, forms the standards for wireless access for vehicular environments (WAVE). It operates at a frequency of 5.9 GHz, divided into 7 channels, each operating at a frequency of 10 MHz. It provides a high data rate, ranging from 6 Mbps to 27 Mbps and a short range radio communication of approximately 300 meters.

In [15], five distinct communication patterns are described that form the basis of almost all VANET applications. This classification is independent of the actual communication technology and assumes only the availability of a link-layer broadcast and unicast mechanism.

- **Beaconing**, which is continuous update of information among all neighboring

nodes to supply them with up-to-date status data such as the position, speed, and heading of a vehicle to allow for cooperative communication. Data packets are sent as link layer broadcasts to all neighbors in reception range. Communication is single-hop, and information in packets is typically not forwarded.

- **Geobroadcast**, which is essentially immediate broadcast of information in larger geographical area (e.g., for safety application to inform larger number of vehicles about occurrence of an accident). Then the message is sent via link layer broadcast to all immediate neighbors in transmission range. Every receiver located within the specified destination region forwards the unchanged message via broadcast in geographically restricted area.

- **Unicast routing**, which is transport of data through the ad hoc network to a certain destination (other vehicle or RSU). The goal is to use the network for unicast transport of messages, not distribution of messages. The communication can be either single-hop or multi-hop toward the destination.

- **Advanced information dissemination**, which is dissemination of information among vehicles enduring a certain time, capable of bridging network partitions and prioritizing information. It provides information to vehicles that arrive later in time and previously could not be reached due to network partitioning. Schemes for advanced message dissemination usually use single-hop broadcasts. They store messages and forward them multiple times, considering various parameters to determine when to resend a message.

- **Information aggregation**, in contrast to all other patterns, communicated data is processed and merged by network nodes and not simply forwarded. It has a

similar goal as advanced information dissemination (i.e., to spread information among vehicles). However, it can reduce communication overhead when events are detected by multiple vehicles as aggregation can also compress the data.

2.2.6 Packet Transmission Priority and the Hidden Terminal Problem

Video streaming considered to be secondary application in terms of priority when safety applications are present. For example, at the MAC sub-layer, packets containing safety messages are likely to have smaller contention window size compared to video packets. In [16], a spatial differentiation is introduced at the MAC layer where the potential relay nodes are assigned different access priorities through different contention window sizes to speed up dissemination of safety messages. Therefore, a relay node with a higher distance from the source node is assigned a shorter contention window size. In the network, interference may occur between two transmitting nodes because of the overlapping between their transmission ranges. If the transmitting nodes are not within the overlapping area, the resultant interference is traditionally known as hidden terminal problem. Most of the work in the literature address the collision and hidden terminal problem assuming single hop scenario and then extend the results to multi-hop communication assuming independence between the hops. This approach fails to model the dynamics of the network and hence the hidden terminal problem effectively considering that the hidden terminal, which is out of the transmission range of the sender, is assumed to be in isolation, however, in reality the hidden node itself may be exposed to hidden node activity in multi-hop communications. The probability of hidden terminal interference in multi-hop vehicular networks is derived in [17] with consideration of two priority classes of packets.

2.2.7 Performance Metrics for Streaming Applications

Performance measures of content downloading and streaming applications are mostly focused on network layer metrics (such as packet delivery ratio and end-to-end delay) and application layer metrics (such as throughput and delay jitter). File downloading applications are typically delay-tolerant and loss-sensitive. Therefore, packet level metric such as packet delivery ratio is the most important metric. On the other hand, video streaming applications are delay-sensitive, hence, end-to-end delay is the most important packet level metric. For example, media streaming applications use similar application-level performance measures developed for real-time media streaming over the Internet, including end-to-end throughput and end-to-end jitter while end-to-end jitter (E2EJ) refers to the variance of delays for several consecutive packets arriving at the destination [18]. In conclusion, end-to-end delay, jitter and end-to-end throughput are the major three performance metrics to illustrate quality of service, among other metrics.

2.3 Related Work

2.3.1 PHY-Centric Strategies

Adaptive Modulation rates for time varying channel conditions is critical for avoiding performance degradations. The inability to accurately choose the modulation rate for the current channel condition leads to loss or unnecessarily long packet transmission times and inefficient use of the channel [19]. Rate adaptation protocols address channel fading in one of two ways: 1) loss-triggered rate adaptation, the transmitter interprets channel state based upon failed or successful delivery (e.g., [20]); 2) SNR-triggered rate adaptation, in which the receiver uses the SNR to determine the modulation rate and informs the transmitter via the four-way handshake (e.g., [21]). Loss-triggered rate adaptation

methods require limited MAC and PHY interaction and are widely implemented and evaluated in indoor and outdoor settings.

2.3.2 Link Layer and MAC Sub-Layer Strategies

- **Reduction of contention periods:** several methods in literature focus on reducing the contention period time to ensure timely and reliable data delivery to mobile vehicles under heavy traffic where lack of radio resource may cause longer contention periods. In [22], dual channels are utilized to separate control packets from data packets which results in a reduction in the number of collisions.

  In [23], a contention-based MAC protocol for WLANs, namely, the fast collision resolution (FCR) algorithm, is proposed. To speed up the collision resolution, the backoff timers for all active nodes are actively redistributed and to reduce the average number of idle slots, smaller contention window sizes for nodes with successful packet transmissions are used. When a fixed number of consecutive idle slots are detected, the backoff timers are reduced exponentially fast.

  In [24], a region-based clustering mechanism is used to improve the performance of MAC operations for VANET. The service area is divided into a set of region units, and limit the number of vehicles in each region unit for the contentions of radio channels. Each region unit is then associated with a non-overlapping radio channel pool. Since the number of vehicles in each region unit is limited, the contention period is reduced and the throughput is increased.

  The controlled vehicular Internet access (CVIA) protocol [25] avoids contention and unnecessary RTS/CTS overhead by eliminating contention in relaying packets over long distances and provides fairness among segments of a street by using an analytical model to estimate the throughput of each segment and adjust the

protocol parameters so that it can equate the number of packets delivered from each segment.

- **Collision avoidance:** because VANETs do not have a fixed infrastructure, it is not easy for their nodes to know if the medium is in use or not [26]. Vehicles do not have shortage in terms of data storage and power. We assume that it is always possible for a vehicle to obtain its geographic position by using GPS, which provides good time synchronization through the network. Hence, VANET MAC protocols could take power constraints or time synchronization problems less into consideration, but they have to be concerned with the fast topology changes.

  ADHOC MAC [27] is based on a dynamic time division multiple access (TDMA) in which a broadcast signaling channel is set up in a completely distributed way to overcome hidden terminal problems and achieve reliable broadcasting. From this channel, all the nodes know the activity of their two-hop neighbors and can therefore avoid collisions with already set up connections.

  Directional antenna-based MAC protocols have been proposed in [28] and [29] for ad-hoc networks. If vehicles are equipped with a directional antenna, they could transmit in a specific direction. Therefore, the transmission space around a vehicular node is divided into N transmission angles of $(360/N)$ degrees which causes reduction of transmission collisions and increment in the channel reuse possibility.

- **Power efficiency:** joint power and sub-carrier allocation under delay constrained data delivery such as video streaming applications is investigated in [30] to improve power efficiency in V2I channels. This scheme is a cross-layer framework in which orthogonal frequency division multiplexing (OFDM) is employed at the Physical Layer and joint power and sub-carrier assignment works at the data link layer. The

contributions of this work includes formulating the cross-layer problem as a convex optimization problem with the help of a time-sharing factor. It has been shown that the power allocation policy converges to a classical water-filling policy if we remove delay constraints. DSRC-AA [31] is a generalization of the asynchronous quorum-based power-saving (AQPS) protocols which capitalizes on the clustering nature of moving vehicles and assigning different wake-up/sleep schedules to the cluster-head and the members of a cluster.

- **Delay-constrained optimization:** since the end-to-end delay performance has a complex dependence on the high-order statistics of cross-layer algorithms, optimization-based design methodologies that optimize the long term network utilization are not well-suited for delay-aware design. Hence, the authors in [32] exploit back pressure-type controllers to develop link rate allocation strategy that not only optimizes long-term network utilization, but also yields loop free multi-path routes between each source-destination pair. This optimization framework can be considered in design of multi-hop V2I and V2V communication protocols for delivering delay-constrained data (e.g., video packets).

- **Packet size adaptation:** in [33], an inter-vehicle information dissemination CSMA/CA based MAC protocol using dynamic optimal fragmentation with rate adaptation algorithm (DORA) has been proposed. DORA achieves maximum goodput in wireless mobile networks by computing a fragmentation threshold and transmitting optimal sized packets with maximum transfer rates. An adaptive on-demand UDP estimator is designed to reduce the SNR estimation overhead. Based on estimated SNR, the algorithm dynamically selects the optimal fragmentation and transmission rate in time varying channels with minimal overhead. This method can be

categorized as a cross-layer optimization algorithm between the PHY layer and the MAC sub-layer.

2.3.3 Network-Centric Strategies

In [34], an integration of VANET and 3G networks using mobile vehicles as gateways (dual-interfaced vehicle that relays data from other vehicle sources to the UMTS backhaul network) is introduced. Mobile gateways reduce the signaling overhead and required frequent changes in the pre-defined routing tables of vehicles compared to systems with static gateway deployment. The focus of this paper is on defining a mechanism that selects a minimum number of optimal VANET gateways at each instance and per moving direction. When the current serving gateway loses its optimality, a gateway handover mechanism is required for migration of the responsibilities of the existing gateway to a newly-elected optimal gateway.

The influence of cooperative transmission to the wireless link cost is first studied in [35], which shows that the quality of wireless links could be improved significantly. The objective functions for the end-to-end reliability and total power consumption is formulated using the improved link metrics.

- **Reactive vs. proactive protocols:** reactive routing protocols do not take into account mobility parameters during route discovery, resulting in paths which often break in highly mobile scenarios such as VANETs. Furthermore, the additional initial latency introduced by the route-discovery procedure poses serious challenges for reactive protocols. For this reason, reactive protocols, in their current format, are seen as inappropriate for time-critical applications such as video streaming for vehicular communications. Generally, it is not possible to come up with a routing protocol for VANETs which can satisfy QoS requirements for different applications,

hence attempts have been made to propose routing protocols which are tailored for different applications. Another critical issue is the consideration of the scalability in design of routing protocols which are robust to frequent path disruptions caused by vehicle mobility. For example, [36] uses movement information of vehicles (e.g. position, direction, speed and digital mapping of roads) to predict a possible link breakage. Proactive routing protocols are overwhelmed by rapid topology changes and may even fail to converge during the routing information exchange stage [37].

- **Position-based routing protocols:** Position-based routing schemes embed physical-position information of the vehicle in the packet header during the routing decision process. A road-based vehicular traffic (RBVT) is proposed in [38] which takes advantage of the layout of the roads to improve the performance of routing in VANETs. Traffic information combined with geographical forwarding is used to create road-based paths between transmitter and receiver nodes along the streets.

- **Flooding control protocols:** flooding control schemes are proposed in order to save system bandwidth and node buffer space, different flooding control. However, these control schemes all assume that nodes have prior information about the history of other nodes. Probabilistic metrics, such as delivery predictability and utility function, were proposed to select the better next-step candidates [37]. However, these flooding control schemes are not practical for partially connected VANETs. The low node density, combined with the difficulty of obtaining the information used in the routing determinations, limits the effectiveness of these schemes.

- **Clustering-based protocols:** vehicles are distinguished as either ordinary mobile nodes, gateway nodes or cluster-head nodes. Based on geographical locations, directions of movement, and other metrics, vehicles are grouped into different clusters.

Clustering enhances effective broadcasting and relaying of messages by reducing the overhead associated with signaling. The main challenge in clustering lies in the dynamic topology changes in VANET. In [36], vehicles are clustered based on their velocity vectors. This kind of clustering ensures that vehicles belonging to the same cluster are more likely to establish single and multi-hop paths as they are moving together which contribute to more stable communication in VANETs.

2.3.4 Mobile IP Management at Transport Layer

Basic extension of Mobile IPv6 for NEMO is used in vehicular networks. In [39], the handover procedure in the fast handover for mobile IPv6 (FMIPv6) protocol is optimized by using the IEEE 802.21 Media Independent Handover (MIH) services. Lower three layers information of the mobile node (MN), mobile router (MR) and the neighboring access networks are used for candidate AR discovery. Furthermore, a cross-layer mechanism is applied to make intelligent handover decisions in FMIPv6. MIH is a standard, developed by IEEE 802.21, to enable the handover of IP sessions from one layer 2 access technology to another, in order to support mobility of the end user devices. The lower layer information of the available links obtained by MIH services and the higher layer information such as QoS requirements of the applications are used by a policy engine to make intelligent handover decisions.

2.3.5 Summary

Due to the low resource utilization and lack of QoS guarantee in single layer protocol design, we anticipate that cross-layer protocols over MAC, Network and Application layers will be deployed in VANETs with infrastructure support for providing satisfactory QoS performance for delay-sensitive multimedia applications such as video-on-demand or video conferencing.

Chapter 3

Quality-Driven MAC Retransmission Limit Adaptation over VANETs

The vehicle engine provides sufficient power for intensive data processing and communications. The on-board buffer storage, positioning system, and intelligent antenna further facilitate efficient video forwarding and collaborative downloading among vehicles or from/to roadside receivers. Intelligent transportation systems (ITS) for VANETs have stimulated the development of several interesting applications such as vehicle collision warning, security distance warning, driver assistance, cooperative driving, cooperative cruise control, etc [3]. When a vehicle detects an emergency event such as road closure due to an accident, it is important to inform other vehicles and relevant rescue parties of the situation as early as possible. Video communication within a VANET has the potential to be of considerable benefit in an urban emergency, as it allows emergency vehicles approaching the scene to better understand the nature of the emergency. However, emergency events may not occur frequently, especially when the network size is small. Under normal operating conditions, it is important to maintain high service quality for applica-

tions such as video on demand (VoD) streaming when the network resources are used for other than transmission of safety messages [40]. A MAC protocol for such network needs to perform two kinds of coordination. For emergency events, the node with emergency message is given a higher transmission priority by broadcasting its traffic type to induce others to wait while the critical message is transmitted. When there is no emergency message, distributed channel coordination is performed.

Among various multimedia applications which can be utilized during non-emergency periods, our focus is on video streaming from a RSU to the destination vehicle. Streaming of high quality video to fast-moving vehicles is a promising application which faces fundamental challenges attributed to the high mobility and dynamic nature of the network. In order to have a smooth playout, it is necessary to have enough packets in the playback buffer at the destination [41]. Robust video streaming applications must be able to tolerate link failures or deep link fading, which normally occur due to node mobility or the unwillingness of the user to share node resources in a multi-hop network [42]. Hence, different protection schemes to be deployed in different layers of the protocol stack should be explored. An adaptive protocol with one time slot memory is proposed in [40], which allows transmission parameters to adjust to the traffic type (normal or critical). With an adaptive protocol, a user can change its mode of operation depending on its state. In [43], a modified version of packet reservation multiple access (PRMA) which can adapt to the traffic and data dynamics, is proposed. An access control scheme with QoS provision is proposed in [44], in which periodic admission control is utilized by RSUs to provide bounded throughput guarantees for soft real-time traffic over a multi-hop network.

Previous works, such as [4], focus on networking quality metrics like throughput and transmission delay. In this chapter, a multi-objective MAC retransmission limit adaptation scheme, which jointly minimizes the video streaming quality metrics, start-

up delay and the frequency of playback freezes, over urban VANET scenarios, is proposed. We map these quality metrics onto MAC retransmission parameter to formulate and solve the proposed multi-objective optimization. We study the variation of these two metrics as a function of MAC retransmission limit and found two separate regions for packet arrival rates, λ_{ij}, to be considered in our formulated optimization problem. Let μ_{ij} be the packet playback rate at the destination vehicle buffer. For $\lambda_{ij} \leq \mu_{ij}$, frequency of playback freezes increases with second degree function and the start-up delay decreases with exponential rate. For $\lambda_{ij} > \mu_{ij}$, although it is expected that the number of playback freezes be zero, there is finite probability for playback freezes to occur. This can be derived using diffusion approximation and the formula derived for frequency of playback freezes in the $\lambda_{ij} \leq \mu_{ij}$ region does not stand. In this region, the probability of playback freezes is incorporated in the objective function. We analyze the optimization problem in both regions and estimate the optimal MAC retransmission limit to maximize the video quality according to the metrics under consideration. In addition, wireless systems should work well in typical wireless environments, characterized by the path loss of the signals, multipath fading, interference to adjacent channels, and random errors. We specifically address the mobility impact on channel estimation, access connectivity probability and interference caused by neighboring vehicles in the design of the proposed adaptive quality-driven MAC protocol.

While connectivity to the Internet has become an essential part of our daily life, the access to high-rate Internet from vehicles is a luxury application in most areas. In cellular or satellite communications, the available rate for each user is far from enough to deliver the media-rich Internet contents. Nearly, 10% of the waking time of the American people is spent in vehicles in which people have very limited or no access to Internet [4]. The VANETs have been recently introduced to provide high data rate at low cost to vehi-

cles by utilizing DSRC which is based on IEEE 802.11. For transmission using DSRC, there are two different mechanisms that can be deployed at the link layer to combat the time-varying wireless channel conditions: 1) switching among different PHY modes, each with a different modulation scheme and data rate, and 2) performing link layer (or the MAC protocol layer) retry. In this chapter, we focus on the retry mechanism. According to the DSRC, when a transmitted packet is not acknowledged properly, retries can be performed and repeated until a certain limit is reached. Packets are dropped when they reach their retry limits. Retry is an efficient means to improve the reliability of the link. In the current wireless access in vehicular environments (WAVE) standard, this retry limit is normally configured statically, and there is no recommendation or guidline on how the retry limit can be adapted based on the channel conditions or the workload. However, the retry limit may affect not only the link packet erasure rate but also the playback buffer filling rate and hence the streaming quality of the video. The design of this scheme is based on a key observation, that is, for a given traffic characteristic and channel condition there exists an optimal retry limit setting for the wireless link under which the total losses due to both link erasure and buffer overflow will be minimum. It is observed that when traffic characteristics or channel conditions change, the optimal setting also changes, but always stay at a value that can balance the link erasure rate and the overflow rate. Adaptive protection strategies, like priority queueing, selective dropping, and unequal protection can be achieved through real-time MAC-layer retry-limit adaptation. This observation leads to the design of the real-time retry-limit adaptation scheme and the systematic method of retry-limit vector configuration for unequal protection in case of multi-layer video compression or existence of video applications with different priorities. A MAC retransmission limit adaptation has been proposed which is based on optimization of playback streaming quality. We formulated the design of the

adaptive MAC protocol as a multi-objective optimization framework at the RSU which jointly minimizes the probability of playback freezes and start-up delay of the streaming at the destination vehicle by adjusting the MAC retransmission limit with respect to channel statistics (packet error rate and packet transmission rate). The impact of mobility including access of vehicle to the RSU and Doppler shift effect are incorporated in design of the protocol. In the multi-hop scenario, we discuss the computation of access probability used in the MAC adaptation scheme and propose cross-layer path selection scheme followed by a discussion on mobile IP management to maintain continues video streaming. The proposed scheme can achieve significantly fewer playback freezes while introducing only a small increase in start-up delay.

It is highly desirable to always have successful direct transmission without retransmission in any system; however, such situation is not always available mainly because probing the system for prior knowledge of how to achieve successful direct transmissions can be a futile exercise. For this reason, in most systems including VANETs, retransmission will be necessary in order to allow the same transmission to be delivered without constant interruption, which is the reason why the destination vehicle has playback mechanism and is the motivation why in this work we consider frequency of playback freezes and start-up delay as criteria to optimize retransmission limit. The contributions of this chapter are as follows: 1) A MAC retransmission limit adaptation scheme proposal based on multi-objective optimization of playback streaming quality; 2) incorporation of the impact of mobility on vehicle access to the RSU and Doppler shift in the design of the adaptation scheme; and 3) extension of the proposed MAC retransmission scheme to multi-hop scenarios by proposing a cross-layer joint path selection and MAC retransmission adaptation.

The remainder of this chapter is organized as follows. The network topology and

system model are discussed in section 3.1. Section 3.2 describes the proposed cross-layer multi-objective MAC protocol design. Section 3.3 presents simulation results and validation of analytical results. Concluding remarks are summarized in section 3.4.

3.1 Network Topology and System Model

We consider a vehicular network based on DSRC in which vehicles are equipped with OBUs and broadcast the location, direction, speed, acceleration and traffic events to their neighbors [45]. We consider an infrastructure-based one-dimensional VANET as shown in Figure 3.1 wherein the RSUs are distributed uniformly along the road with intermediate distance of L and vehicles are randomly distributed according to a Poisson distribution. The RSUs are spaced by L meters, with an effective radio coverage of R meters each. The radio coverage of each vehicle is approximately 200 m. In this chapter, we focus on the wireless part of the network and assume required support is in place in order to transmit the video packets from video servers to RSUs along the road via wireline connection to the Internet backbone. For example the mobility management scheme proposed in [1] can be utilized in order to control the transmission of the video from an AR to the appropriate RSU along the road based on the location of the vehicle.

The video is streamed from an AR to the proper RSU, and from there to the destination vehicle. While the destination vehicle is in the transmission range of the RSU, they connect directly in a one-hop fashion. When the destination vehicle gets closer to the next RSU compared with its distance to the previous RSU, the AR switches the video streaming to the next RSU. In what follows, we first discuss the quality metrics to be optimized, issues like physical channel model, mobility modeling, access provision to the RSU, effect of Doppler shift on channel estimation, formulation of the optimization problem and our solutions on how to consider all the raised issues in the design of an

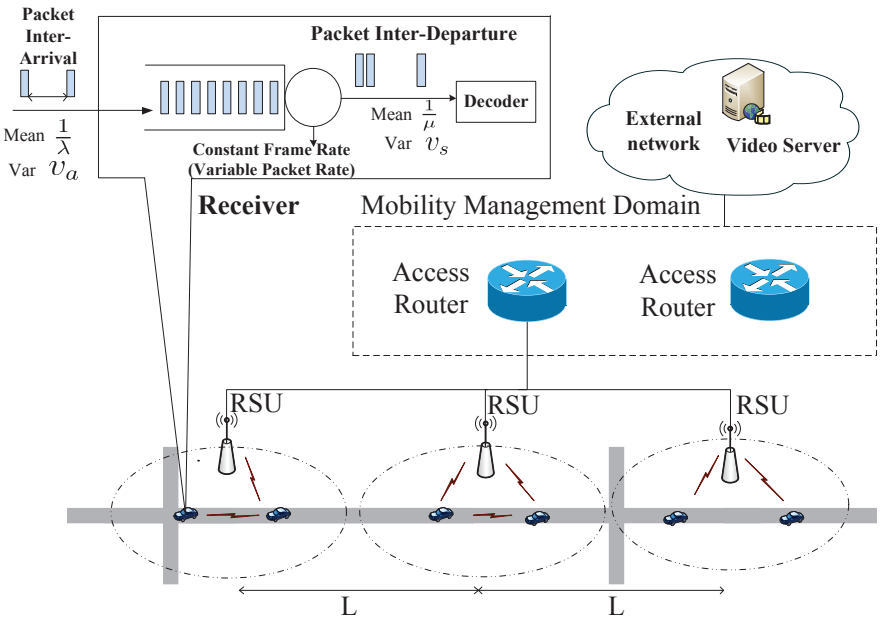


Figure 3.1: Network topology

effective MAC protocol in VANET scenarios.

3.1.1 Physical Channel

Orthogonal frequency division multiple access (OFDMA) is the deployed modulation scheme at the RSU in our system model. The principle of OFDMA technique is to split the available bandwidth into a number of orthogonal subbands, each supported by a subcarrier which allows the high rate stream to be divided into several lower bit rate streams. It can provide great flexibility for subcarrier assignment to maximize the system capacity and spectral efficiency [46]. The BS is connected to the RSUs which streams video to mobile vehicles through one-hop communications. The BS receives video streams from media server via a backbone network, which is assumed to have high bandwidth and lossless channel. IEEE standard for WAVE is the wireless medium access scheme used by the communicating entities. WAVE adopts enhanced distributed channel

access (EDCA) which is a contention-based channel access scheme with QoS provision, i.e., different packets are categorized based on their priority and different scheduling is applied for each category. The conventional first-come-first-serve (FCFS) algorithm is still recommended to be used in WAVE in coordination with IEEE 802.11e EDCA technology.

To ensure reliable system performance, it is necessary to evaluate it using realistic channel models before deployment. Many scatterers are present in the urban area. In addition to the scattering, there is a strongly dominant signal seen at the receiver, usually caused by line of sight which can be modeled by a log-normal shadowing model. A standard for vehicle-to-infrastructure (V2I) communication in the 5.9 GHz unlicensed national information infrastructure (UNII) has been developed [47]. The wireless link is assumed to be a memoryless channel. The packet error rate, p_{ij}, between nodes i and j for a packet of L bits can be approximated by the sigmoid function.

$$p_{ij} = \frac{1}{1 + e^{\xi(Y_{SNR} - \delta)}} \tag{3.1}$$

where Y_{SNR} is the SNR of the received signal; ξ and δ are the constants corresponding to the modulation and coding scheme for a given packet of length L. The goal is to support high mobility platforms using the IEEE 802.11 protocol. We assume that each RSU adopts some type of link adaptation scheme in order to maximize its outgoing link throughput. It can select adaptive modulation and coding schemes based on the detected SNR of the link.

3.1.2 Medium Access Control and Downlink Scheduling

Safety message delivery is categorized as an event-driven application with highest priority in which the safety messages will propagate from the source outwards as far as possible.

Video packets are associated with lower priority compared to safety messages and, furthermore, video packets are classified into three different priority classes as follows.

In common compression standards such as MPEG and H.264, encoded frames at the encoder side are arranged into structures called GOP. Each frame is composed of several packets. A GOP contains the following frame types : I-frame, P-frame and B-frame. I-frames are decoded independently while P-frames are decoded based on preceding I-frame or P-frame. B-frames are decoded based on preceding and succeeding I-frame or P-frame within the GOP. A GOP always begins with an I-frame. Afterwards, several P-frames follow. The B-frames are inserted between two consecutive anchor frames. Hence, a logical classification scheme is to consider I-frame packets as high priority packets while P-frames and B-frames are classified as low priority packets because a successfully received I-frame can correct any errors caused by any preceding frames. High and low priority packets are scheduled and delivered differently.

As mentioned before, WAVE is the deployed air interface control scheme in our system. The detailed channel operation can be found in [48]. The typical channel operations can be summarized as follows [49]. Both RSU and OBU support at least one control channel (CCH) and multiple service channels (SCHs). Figure 3.2 shows the scheduling architecture of packets in IEEE WAVE standard. First, upon power on, an OBU monitors the CCH until a WAVE service advertisement (WSA) sent by an RSU is received. A WSA carries the information of available SCHs and their access parameters, such as channel numbers. Based on the WSA information, the OBU then synchronizes with the RSU, and the OBU can exchange data with the RSU in SCHs. The CCH interval is used to send high-priority safety messages and WSAs, whereas the SCH interval is used for data services. The length of the CCH and SCH intervals, that is, T_{CCH} and T_{SCH}, can be adjusted adaptively. However, the total length of the sync interval, T_{SYNC}, is

fixed to $T_{SYNC} = T_{CCH} + T_{SCH} + 2T_{GUARD}$. The synchronization ensures the OBUs to monitor CCH during CCH intervals. Video packets are delivered in SCH intervals. The discussion above is based on single-channel WAVE devices. With a single channel, an OBU can work on either CCH or SCH at a time. If two or more channels are facilitated in a WAVE device, the operations in CCH interval and SCH interval can be conducted simultaneously. In this paper, we consider single-channel WAVE devices only as this is common for OBUs.

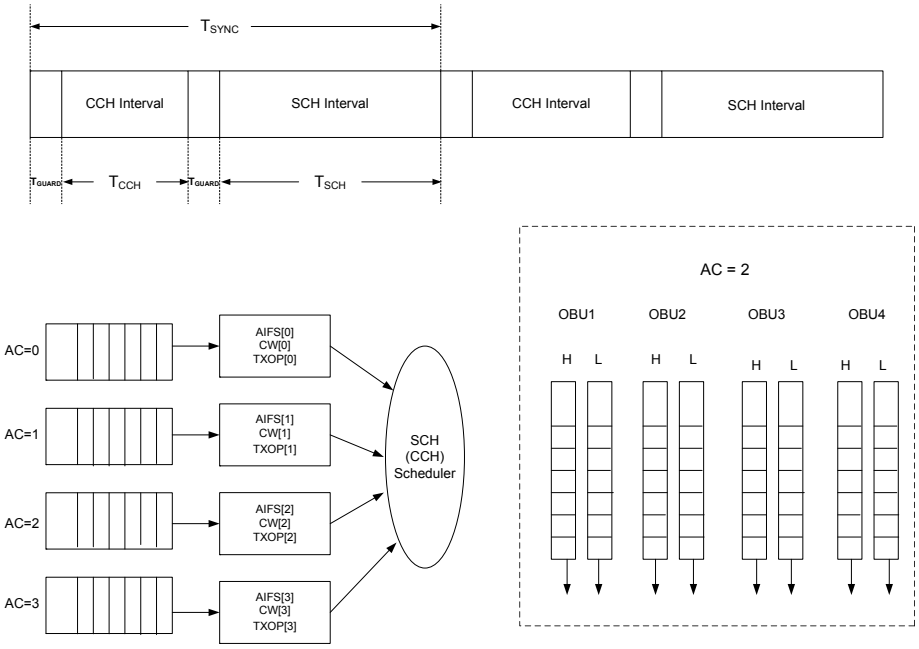


Figure 3.2: Scheduling architecture of video on demand in IEEE WAVE standard

3.1.3 Impact of Mobility

Related Works on Connectivity of Mobile Nodes

The transmission range (power) is an important design factor to keep the network connected while avoiding generation of interference to other receivers in the same chan-

nel. Asymptotic bounds for probability of connectivity (as the number of nodes in the network goes to infinity) are derived in [50], while nodes are uniformly distributed over a circular area. Connectivity of ad-hoc networks with finite number of nodes uniformly distributed over a one-dimensional network is analyzed in [51]. By projecting the nodes onto the x and y axes, they have extended their work to find an upper connectivity bound for a quadrant two-dimensional network of independent and identically distributed (i.i.d) nodes. Connectivity probability of one-dimensional ad-hoc networks in which location of nodes have non-identical distribution is investigated empirically in [52] which leads to an optimization of the number of nodes required to maintain the connectivity. Connectivity of one-dimensional stationary ad-hoc wireless networks are analyzed from another perspective in [53]. The authors find the probability that the network is composed of at most C clusters and the probability of network connectivity is presented as a special case when $C = 1$. Closed form analytical expression for the probability of connectivity of mobile ad hoc network of a finite number of nodes with random walk mobility model is presented in [54]. By Using the method of first and second moments in [55], a very strong zero-one law for the property of graph connectivity under the asymptotic regime is derived for i.i.d uniformly distributed nodes over the interval $[0, 1]$.

Connectivity of mobile nodes in a VANET over a single highway with multiple lanes, which allows the vehicles to pass each other, is studied in [56]. The vehicles enter the highway according to a Poisson process. The distribution of nodes over highway as well as distribution of nodes within the transmission range of an arbitrary node at steady state are presented. The latter distribution is an important factor to be considered in designing effective MAC protocols. In [57], connectivity requirements in terms of required penetration rate (number of nodes equipped with communication devices) and transmission power for dissemination of time-critical information in a one-way or two-way

VANET are derived while taking important physical-layer parameters, such as fading, propagation path loss, transmit power, and transmission data rate, into consideration. Another approach proposed in [58] applies the exponential inter-arrival time distribution between vehicles in order to obtain the inter-vehicle distance distribution and, accordingly, to derive explicit expression for the expected connectivity distance. To study the effect of speed on connectivity, they have provided bounds obtained using stochastic ordering techniques, based on the work of Miorandi and Altman in [59], that transform the problem of connectivity distance distribution into that of the distribution of the busy period of an equivalent infinite server queue.

A mobility model has been developed in [60] to represent the steady state distributions of both node population and node location based on the assumption that vehicles arrival rate to the highway follows a Poisson distribution. Furthermore, statistics of connectivity have been investigated using the proposed mobility model. The proposed traffic model in [61] is a combination of a deterministic fluid dynamic model and a stochastic model. The former model is used to characterize the general flow and evolution of traffic stream while the latter model is applied in order to address the random movement of an individual vehicle. They have considered traffic in a one-way, single-lane, semi-infinite signalized urban route segments separated by junctions.

In [62], two idealized mobility models for vehicular mobility were described, namely, Freeway model and Manhattan model. The freeway model is limited to movement of vehicles in one direction. The Manhattan model considers movement of vehicles in two opposite directions and the possibility of turning at intersections. In sparse situations, the mobility of the vehicles may be considered independent from each other; however, for dense situations, this assumption is not correct. According to DSRC, each vehicle is broadcasting information of position, current time, direction, speed, acceleration/deceleration, and

traffic events. The RSU can update information regarding location of the nodes in order to construct the path between the source (RSU) and the destination. In [63], a stochastic vehicular mobility model is proposed which is capable of including the behavior of vehicles in sparse and dense scenarios. The streets are divided into several sections and each section is associated with several speed categories. To model this system, a queueing network is considered which comprises queues of customers in cascade, each with several classes of customers. The customers are vehicles and the speed of vehicle determines the class of the customer. The vehicle with higher speed passes each section of street (get served at the corresponding queue) with higher rate. In [64] we model the system as having three classes of customers at the node which represent the middle part of the street and two classes of customers for the nodes that represent the front and end parts of the street. In a sparse scenario, we can consider different vehicles having independent mobility patterns. Therefore, we assume each node at the preceding queueing network as an $M/G/\infty$ node. In a dense scenario the mobility of vehicles can be modeled by a network of $M/G(n)/\infty$ queues where service rates are dependent on the number of vehicles in the street.

Mobility Model

Aside from the mobility model we select to use in this work, the goal is to estimate an important stochastic attribute of VANET, which is the probability of connectivity of vehicle to the RSU along the road. This probability will be used in the design of the proposed adaptive retransmission scheme. We rely on the results in several works (e.g., [65]) which prove that inter-vehicle distance along the highways follows exponential distribution. Based on this assumption, the connectivity probability bounds presented in [66] can be applied in the design of the MAC adaptation scheme.

Access Connectivity Probability

The access probability (AP) is the probability that an arbitrary vehicle is connected to an RSU, which is a function of key VANET parameters such as inter-RSU distance, vehicle density, and transmission ranges of RSUs. The RSUs are deployed uniformly along the road with Euclidean distance L between two adjacent RSUs and vehicles enter the street according to a Poisson distribution, i.e., the distance between vehicles follows an exponential distribution. Let the coordination reference be at the first RSU (RSU1) and the second RSU (RSU2) be located at distance L from RSU1. The distance of the vehicle from the reference is denoted by x. Let $g_R^\zeta(x)$ denote the probability that a vehicle and RSU separated by distance x are directly connected under channel model ζ which in our work, is the log-normal shadowing. The probability that the vehicle is connected either to RSU1 or RSU2 is

$$P_a(x) = 1 - (1 - g_R^\zeta(x))(1 - g_R^\zeta(L - x)) \tag{3.2}$$

In [66], the access probability in two-hop connection scenarios are computed under a unit disc model and log-normal shadowing fading model. We apply a similar approach to compute the access probability of vehicles in direct connection scenario under a log-normal shadowing fading model, which is commonly used to model the real world signal propagation where transmit power loss increase logarithmically with Euclidean distance between sender and receiver due to shadowing effect caused by the VANET's environment. The received signal power in dB is given by $p_{rx} = p_0 - 10\alpha \log_{10} \frac{1}{d_0} + N_\sigma$ where p_0 is the received signal power at the reference distance d_0, α is the path loss exponent, N_σ is the Gaussian random variable with zero mean and variance σ^2, and l is the Euclidean distance between RSU and the destination vehicle. The RSU and the destination vehicle

can establish a direct connection if the received signal power at the destination, p_{tx}, is greater than or equal to a certain threshold p_{th}. According to [66], by assuming symmetric wireless connection between RSU and vehicles and assigning $p_{th} = p_0 - 10\alpha \log_{10} \frac{x}{R}$, where R is the transmission range of RSUs, the result for g_R^ζ under the log-normal shadowing model can be greatly simplified as follows.

$$g_R^\zeta = Pr(p_{rx} \geq p_{th}) = Q(\frac{10\alpha}{\sigma} \log_{10} \frac{x}{R}) \tag{3.3}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-x^2}{2}} dx$ is the tail probability of the standard normal distribution. When $\sigma = 0$, the log-normal model will be simplified to a unit disc model with $g_R^\zeta = Pr(x \leq R)$. Substitution of (3.3) into (3.2) will result in direct access probability as

$$P_a(x) = 1 - (1 - Q(\frac{10\alpha}{\sigma} \log_{10} \frac{x}{R}))(1 - Q(\frac{10\alpha}{\sigma} \log_{10} \frac{L-x}{R})) \tag{3.4}$$

The RSUs in VANETs have lower coverage range compared with the base stations deployed in the LTE-Advanced networks which causes the nodes to stay for a shorter time in the coverage range of each RSU. Hence, the access probability of vehicles to the RSUs plays an important role in overall system performance and must be taken into consideration.

Impact of Doppler Shift on Channel Estimation

Due to the fast movement of vehicles and potential obstacles appearing along the road between RSU and the destination vehicle that cause shadowing effect, the level of SNR in the wireless channel, and hence packet error rate, is subject to change. Hence, it is imperative to address this issue in the design of an adaptive MAC protocol. The MAC needs to adapt itself to the wireless channel condition. We consider the effect of Doppler shift on the channel coherence time (i.e., the duration of time in which the

channel remains stationary). The estimator deployed at the RSU receives feedback on RSS, $Y(t)$, from the destination vehicle. Let $\overline{Y}_{RSS}(t-1)$ be the average RSS up to packet $t-1$; the estimated SNR of the $(t+1)$th packet, $\hat{Y}_{SNR}(t+1)$, by using an exponential moving average, is as follows [67].

$$\hat{Y}_{SNR}(t+1) = (1-\gamma)[\alpha\overline{Y}_{RSS}(t-1) + (1-\alpha)Y(t)] + \gamma\hat{Y}_{SNR}(t) \tag{3.5}$$

where $\hat{Y}_{SNR}(t)$ is the average SNR at the receiver, which is fed back to the RSU via the CCH. $0 \leq \alpha \leq 1$ and γ are the estimation parameters which capture the properties of the moving average. The receiver informs the RSU in two events: 1) when the difference between the average SNRs is more than a predefined threshold, Δ_{SNR}, or 2) when $\overline{Y}_{SNR}(t)$ stays in the channel longer than the channel coherence time, T_c, which is a function of the maximum Doppler frequency shift, f_m, i.e., $T_c = \frac{0.423}{f_m}$. The frequency decreases when the vehicle moves away from the RSU and increases when it moves toward the RSU according to the Doppler frequency shift, $f = (1 - \frac{V_s - V_r}{C})f_0$, where f_0 is the center frequency of the signal, while V_s, V_r and C are vectorial velocities of sender, receiver and waveform in free space, respectively. Obviously, in VANETs the velocity of RSU is equal to zero. Substitution of the updated SNR given by (3.5) into (3.1) results in the updated link loss probability between RSU and destination vehicle, which will be used in designing the quality-driven adaptive MAC retransmission protocol.

3.2 Formulation of Multi-Objective Optimization Protocol

A protocol is defined to be a collection of decision rules for each user that is utilizing the network resources [40]. The protocol must be symmetric (i.e., it must assign same

decision rules to all vehicles) since all the vehicles are identical in our model.

3.2.1 Performance Metrics

We assume that each vehicle has an infinite buffer, which is a reasonable assumption given the high storage capability that can be deployed in vehicles. The video playback process can be divided into two phases: 1) charging phase and 2) playback phase. The charging phase starts once the buffer becomes empty. Thus, the playback is kept frozen until the buffer is filled with b packets (i.e., b is a threshold of the playback). To derive an analytical formulation for streaming start-up delay (charging phase) in video streaming at the destination vehicle, the playout buffer can be modeled as a $G/G/1/\infty$ queue that follows the diffusion approximation method presented in [41]. By applying the diffusion approximation, the transient solution of the queue length can be exploited by obtaining its p.d.f. at any time instant t. The average start-up delay and its variance are given by

$$E(D_s) = \frac{b}{\lambda} \tag{3.6}$$

$$Var(D) = bv_a \tag{3.7}$$

where b is the playback threshold, v_a is the variance of inter-arrival rate of packets at the destination buffer and λ is the arrival rate of the packets at the destination vehicle. The playback terminates when the buffer becomes empty again. According to [41], the average number of streaming freezes after t seconds and its variance can be approximated using diffusion approximation as follows:

$$E(F) \approx -\frac{\lambda(\lambda - C)}{C \cdot b}t \tag{3.8}$$

$$Var(F) \approx \frac{\mu^3\lambda^3(v_a + v_s) + 3v_a\lambda^4(\lambda - \mu)}{b^2\mu^2} \tag{3.9}$$

where C is the service rate of the buffer and v_s is variance of service interval at the destination buffer.

3.2.2 MAC Retry Limit Adaptation

In existing WLANs environments, various protection strategies are available at various layers of the protocol stack for different tradeoffs among throughput, reliability and delay. These include 1) switching among different modulation and channel coding schemes, 2) retransmission and forward error correction (FEC) at the MAC layer, 3) FEC, automatic retransmission request (ARQ), or hybrid ARQ along with error resilient video coding schemes and error concealment strategies at the application layer, and 4) packetization optimization at the various layers [68]. In the following, we analyze how MAC retry limit may affect buffer overflow (loss due to congestion) and link loss due to packet drop. Analysis of such problem in general will be very generic.

Intuitively, when the wireless link experiences high packet error rate due to noise or collision of packets from parties which use the common wireless medium, more packet retransmissions are required in order to correctly deliver the packet. On the other hand, this will increase the probability of buffer overflow. Hence, there is a tradeoff between these two types of packet loss. To find the optimal retry limit, the transmitter is required to estimate the channel which it intends to use for packet transmission via transmission

of pilot signals. [1] Due to high mobility of vehicles, continuous channel estimation seems a formidable task. Hence, without loss of generality, a static log-normal shadowing channel model is considered, which is a practical model for urban environment. Of course, a series of channel estimation can be done in practice to find the channel characteristics closest to the environment which the network is designed for. In addition, packet drop due to overflow is performed before packets are put on the link. Let M_{ij} and $p_{ij}(x)$ be the link retry limit and packet error probability for link $(i \rightarrow j)$ between intermediate nodes (vehicles), i and j, respectively. The mean number of transmissions when the vehicle is at distance x from the RSU, $n(M_{ij}, p_{ij}(x))$, for a single packet until it is successfully received by node j or it reaches the retry limit is calculated as follows.

$$n(M_{ij}, p_{ij}(x)) = \frac{1 - p_{ij}(x)^{M_{ij}+1}}{1 - p_{ij}(x)} \tag{3.10}$$

where $p_{ij}(x)$ is the most recent updated link loss probability of the channel between the RSU and the vehicle at distance x from it. The blocking probability of the link, $P_{ij}^{B}(M_{ij}, p_{ij}(x))$, for the fluid model with arrival rate λ_{ij} (packets/s) and channel service rate (channel maximum capacity) C_{ij} (packets/s) is

$$P_{ij}^{B}(M_{ij}, p_{ij}(x)) = \frac{P_a(x)\lambda_{ij} \cdot n(M_{ij}, p_{ij}(x)) - C_{ij}}{P_a(x)\lambda_{ij} \cdot n(M_{ij}, p_{ij}(x))} \tag{3.11}$$

The link packet drop probability after M_{ij} unsuccessful retries can be calculated as

$$P_{ij}^{D}(M_{ij}, p_{ij}(x)) = p_{ij}(x)^{M_{ij}+1} \tag{3.12}$$

The overall packet loss probability, assuming $P_{ij}^{B}(M_{ij}, p_{ij}(x))$ and $P_{ij}^{D}(M_{ij}, p_{ij}(x))$ are

[1] Estimation is performed at the receiver and fed back to the transmitter, which also sustains transmission delay.

small, is approximated as

$$P_{ij}^{T}(M_{ij}, p_{ij}(x)) \cong P_{ij}^{B}(M_{ij}, p_{ij}(x)) + P_{ij}^{D}(M_{ij}, p_{ij}(x)) \tag{3.13}$$

3.2.3 Derivation of Quality Metrics

In this subsection we derive the effective packet rate at the destination. Inserting the effective rate in the performance metrics, gives the performance as a function os retransmission limit. The effective rate of the packets at the destination is a function of access probability, input rate, and total packet loss in the channel. The effective arrival rate of packet at the destination buffer is given by the following.

$$\lambda'_{ij} = P_a(x)\lambda_{ij}(1 - P_{ij}^{T}(M_{ij}, p_{ij}(x))) \tag{3.14}$$

where λ_{ij} is the packet input rate by the transmitter (RSU) before the signal enters the channel and λ'_{ij} is the effective packet arrival rate at the destination buffer. Substitution of the effective arrival rate in start-up delay (3.6) will result in the following:

$$\begin{aligned}
E(D_s) = \frac{b}{\lambda'_{ij}} &= \frac{b}{P_a(x)\lambda_{ij}(1 - P_{ij}^{T}(M_{ij}, p_{ij}(x)))} \\
&= \frac{b}{P_a(x)\lambda_{ij}(1 - p_{ij}(x)^{M_{ij}+1} - \frac{P_a(x)\lambda_{ij}\cdot n(M_{ij}, p_{ij}(x)) - C_{ij}}{P_a(x)\lambda_{ij}\cdot n(M_{ij}, p_{ij}(x))})} \\
&= \frac{b}{P_a(x)\lambda_{ij}(1 - p_{ij}(x)^{M_{ij}+1} - \frac{P_a(x)\lambda_{ij} - P_a(x)\lambda_{ij}p_{ij}(x)^{M_{ij}+1} - C_{ij} + C_{ij}p_{ij}(x)}{P_a(x)\lambda_{ij} - P_a(x)\lambda_{ij}p_{ij}(x)^{M_{ij}+1}})} \\
&= \frac{b(1 - p_{ij}(x)^{M_{ij}+1})}{P_a(x)\lambda_{ij}(1 - p_{ij}(x)^{M_{ij}+1})^2 - P_a(x)\lambda_{ij}(1 - p_{ij}(x)^{M_{ij}+1}) + C_{ij}(1 - p_{ij}(x))}
\end{aligned} \tag{3.15}$$

Hence, the effective arrival rate of packets at the destination buffer, λ', is

$$\lambda'_{ij} = \frac{b}{E(D_s)} = \frac{P_a(x)\lambda_{ij}(1-p_{ij}(x)^{M_{ij}+1})^2 - P_a(x)\lambda_{ij}(1-p_{ij}(x)^{M_{ij}+1}) + C_{ij}(1-p_{ij}(x))}{(1-p_{ij}(x)^{M_{ij}+1})} \tag{3.16}$$

Let $a = (1-p_{ij}(x)^{M_{ij}+1})$. The frequency of playback freezes in 1 second interval ($t = 1sec.$) can be derived by substitution of λ' from (3.16) in (3.8).

$$\begin{aligned}
E(F) =& \frac{\lambda'_{ij}(\mu_{ij}-\lambda'_{ij})}{\mu_{ij}b} \\
=& \frac{1}{\mu_{ij}b}\frac{P_a(x)\lambda_{ij}a^2 - P_a(x)\lambda_{ij}a + C_{ij}(1-p_{ij}(x))}{a}\times \\
& (\mu_{ij} - \frac{P_a(x)\lambda_{ij}a^2 - P_a(x)\lambda_{ij}a + C_{ij}(1-p_{ij}(x))}{a}) \\
=& \frac{1}{\mu_{ij}ba^2}[P_a(x)\lambda_{ij}a^2 - P_a(x)\lambda_{ij}a + C_{ij}(1-p_{ij}(x))]\times \\
& [\mu_{ij}a - P_a(x)\lambda_{ij}a^2 - P_a(x)\lambda_{ij}a + C_{ij}(1-p_{ij}(x))] \\
=& \frac{1}{\mu_{ij}ba^2}[-P_a^2(x)\lambda_{ij}^2a^4 + P_a(x)\lambda_{ij}\mu_{ij}a^3 + (P_a^2(x)\lambda_{ij}^2 - \mu_{ij}P_a(x)\lambda_{ij})a^2 + \\
& (-2P_a(x)\lambda_{ij} + \mu_{ij}) + C_{ij}(1-p_{ij}(x))a + C_{ij}^2(1-p_{ij}(x))^2]
\end{aligned} \tag{3.17}$$

3.2.4 Optimization Framework

Multi-objective optimization (or programming) is the process of simultaneously optimizing two or more conflicting objectives subject to certain constraints. The optimization problem that we consider involves the maximization of two objective functions are not necessarily of equal importance. The solution to this problem can be described in terms of decision vector $(x_1, x_2, ..., x_n)$ in decision space X. A function $f : X \to Y$ evaluates the quality of the specific solution by assigning it an object vector $(y_1, y_2, ...y_n)$ in object space Y.

The tradeoff between frequency of freezes and start-up delay can be calculated from (3.6) and (3.8), yielding

$$E(F) = \frac{b\mu_{ij}E(D_s) - b^2}{b\mu_{ij}E^2(D_s)} \tag{3.18}$$

where μ_{ij} is the packet playback rate at the destination buffer inside the vehicle. Let $\hat{D}$ and $\hat{F}$ be the maximum tolerable start-up delay and frequency of video playback freezes, respectively. Our objective is to manage the MAC retransmission limit M_{ij} to maximize the video perceived quality within the tolerable range of start-up delay, $\hat{D}$, and frequency of playback freezes, $\hat{F}$, i.e.,

- P1:$(\lambda'_{ij} > \mu_{ij})$

$$\begin{aligned} \min_{\lambda'_{ij}} \quad & P + w_1 E(D) \\ s.t. \quad & P\{D > \hat{D}\} \leq \xi \end{aligned} \tag{3.19}$$

- P2:$(0 < \lambda'_{ij} \leq \mu_{ij})$

$$\begin{aligned} \min_{\lambda'_{ij}} \quad & E(D) + w_2 E(F) \\ s.t. \quad & P\{D > \hat{D}\} \leq \xi \\ & P\{F > \hat{F}\} \leq \eta \end{aligned} \tag{3.20}$$

where $w_1, w_2 > 0$ are weighing factors that can be adjusted based on the user's requirements and $0 < \xi, \eta \ll 1$ are predefined scalars. P1 and P2 are nonlinear optimization problems which are computationally very intensive to perform in real-time streaming

systems. To reduce the search space, in order to decrease the computations, we can apply the one-sided Chebyshev inequality on (3.6),(3.7),(3.8) and (3.9), subjective to the constraints of (3.20).

$$(\hat{D} - \frac{b}{\lambda'_{ij}}) \geq \sqrt{\frac{bv_a}{\xi}(1-\xi)}. \tag{3.21}$$

Since $\xi \ll 1$, we can assume $\frac{1-\xi}{\xi} \approx \frac{1}{\xi}$, which will simplify the solution as follows.

$$\lambda'_{ij} \geq \frac{b}{\hat{D} - \sqrt{\frac{bv_a}{\xi}}} \tag{3.22}$$

By Applying the Chebyshev inequality on the maximum tolerable playback freezes constraint in (3.20), we have

$$\lambda'_{ij} \leq \sqrt[5]{\left(\frac{b\hat{F}\mu_{ij}^2}{3v_a}\right)^2} \tag{3.23}$$

Substitution of the constraints in P1 and P2, together with (3.22) and (3.23), will result in smaller yet more conservative search space which can significantly reduce the computational complexity at a cost of user's utility.

3.2.5 Analysis of Multi-hop V2I connection

The results of the preceding subsection are applicable in situations where there are direct (one-hop) connections between the vehicle and the RSUs. In this subsection, we extend our scheme to consider zones with no RSU coverage where two-hop connection can be a solution and improve the overall quality of streaming. To follow the footsteps of one-hop approach and extend it to the two-hop scenario, we first analyze the access probability for two hop communication. Also, when there are intermediate nodes (vehicles) to relay the packets, the problem of routing and path selection becomes important. Ideally, a

scheme which combines routing and MAC must be based on cross-layer (network layer and MAC layer) design, i.e., we must take full advantage of combined parameters of both layers in order to achieve near-optimal results. Hence, we start by introducing the set of parameters to be exchanged between the MAC sublayer and the network (routing) layer. Because of the highly dynamic nature of VANETs, routing is considered to be reactive and performed hop-by-hop in order to adapt to network dynamics.

Access probability for two-hop connection

The probability that a vehicle located at position x, $0 \leq x \leq L$, is connected to either RSU1 or RSU2 in at most two-hop connection, according to **Theorem 1** in [66], is given by:

$$p_a(x) = 1 - (1 - p_1(x))(1 - p_2(x)) \tag{3.24}$$

where

$$p_1(x) = 1 - (1 - g_b^{C\zeta}(x))(1 - g_b^{C}(L - x)) \tag{3.25}$$

$$p_2(x) = 1 - e^{-\int_0^L g_v^{\zeta}(\|x-y\|)\rho p_1(y)dy} \tag{3.26}$$

In the above, $p_1(x)$ is the probability of vehicle being directly connected to either RSU1 or RSU2; $p_2(x)$ is the probability of vehicle being connected to at least one vehicle which is connected to either RSU1 or RSU2; g_b^{ζ} and g_v^{ζ} are probabilities of connectivity of V2I and V2V channels, respectively, which are functions of the distance between transmitting and receiving nodes and channel model ζ. Under the log-normal shadowing model, the access probability for two-hop connection of vehicle to RSU can be obtained for different values of α and σ by computing (3.24) using numerical integration technique. The access probability computed by (3.24) is inserted in (3.17) to derive the expected frequency of freezes for two-hop scenario. Based on the estimated connectivity probability,

retransmission limit is determined by the RSU following the cross-layer optimization in conjunction with path selection described in subsection 3.2.5. Equal retransmission limit will be used for transmission over the first hop (transmission from RSU to intermediate vehicle) and for relaying over the second hop (transmission from intermediate vehicle to destination vehicle).

Cross-layer MAC-NET packet delivery

The routing protocol is location-based, which means that each mobile node incorporates its geographical location while broadcasting "Hello" messages for neighbor discovery. The algorithm is based on greedy geographic routing tailored for video streaming application. In other words the algorithm selects a path which optimizes the streaming metrics as follows. First, the RSU applies greedy geographic routing and selects a cluster of N_D nodes which are closest to the destination such that the distance of these nodes to the destination must be smaller than a predefined threshold, σ. If there are no nodes in the neighborhood of the relaying node which satisfy this condition and the destination node is not in the transmission range of the RSU, the closest node to the destination will be selected.

Let $M_i \in$ be the MAC retransmission value for all the packet of the ith video segment where M is the set of all admissible values. Let $\zeta_i \in Z$ denote the transmission path of packet π_i where Z is the set of all paths possible from the RSU to the destination vehicle by applying the modified greedy routing. The frequency of streaming freezes for the current video slice, $E[F_i]$, and the transmission delay D_i depends on M_i and ζ_i. The problem can be formulated as the selection of optimal transmission path and MAC parameter for all segments of the video clip so as to minimize the frequency of streaming freezes under a delay constraint in a multi-hop V2I channel. Let N be the total number of video segments of the clip and $D_T \approx \frac{1}{\mu}$ be the delay threshold for a video segment,

where μ is the video frame playback rate. The minimization problem is formulated as follows:

$$\min_{M_i \in M \ \zeta_i \in Z} \quad \sum_i^N E[F_i]$$
$$s.t. \quad \max\{D_i, \ldots, D_M\} \leq D_T \tag{3.27}$$

IP Mobility

To maintain continuous streaming, a change of RSU for controlling the transmission of video to the destination vehicle is necessary. We apply our proposed IP mobility management scheme (described in detail in [1]) to multi-hop VANETs with handover prediction. The scheme works in conjunction with a geo-routing algorithm and relies on the IPv6 support for VANET using geo-networking features. The geo-routing layer forwards IP packets in the multi-hop path, in a way that creates a virtual point-to-point link between the destination vehicle and the AR, without IP header forwarding at intermediate vehicles or the RSU.

3.3 Performance Evaluation and Model Validation

3.3.1 Adaptation of Retransmission Limit under Static Channel Condition

When the MAC protocol runs the adaptation algorithm in the background to calculate the optimum limit for retransmissions, it can achieve higher quality in terms of video playback freezes as shown in Figure 3.3(d)-3.3(f). However, the adaptation protocol introduces extra computations to the standard WAVE protocol that results in slight

increase of start-up delay as shown in Figure 3.3(a)-3.3(c). In addition, the impact of channel packet loss on the performance of the algorithm and its comparative performance with the standard WAVE protocol are shown in Figures 3.3(a)-3.3(f). Here, it is assumed that the channel state remains static and hence there would be no channel update during the streaming period. As we change the channel loss probability, the difference between the adaptive algorithm and the standard algorithm for both playback freezes and start-up delay, is more obvious. As packet transmission rate increases, we observe a decrement in the frequency of video playback freezes. However, the start-up delay shows increasing attribute for the non-adaptive protocol while having constant value for the adaptive protocol. The limit of retransmission is fixed in the standard protocol, while the adaptive protocol tends to select higher retransmission limit as the data rate increases which results in higher start-up delay within delay constraint (if the packet transmission delay exceeds the delay threshold, D_T, it will be dropped) for higher data rates. Having generally lower frequency of freezes in the adaptive protocol compared to the standard MAC protocol is a direct consequence of protocol optimization, which is basically designed to select the retransmission limit with consideration of updated channel conditions in order to minimize playback freezes. The drop in the number of freezes with increment of the transmission rate, is due to more successful transmissions and hence more video packet available to be streamed at the playback buffer for each time instant compared to the standard protocol.

3.3.2 Adaptation of MAC Retransmission Limit with Channel Dynamics

We simulate a drive-thru video streaming scenario as shown in Figure 3.1, in which RSUs are deployed along the road and the vehicles compete for communications using IEEE

802.11p. We validate the developed analytical model in two cases: 1) when the inter-RSU distance, L, is completely in the coverage range of the RSUs, i.e., $0 < L \leq 2R$; and 2) when L exceeds the communication range of base stations and hence there are areas in which the vehicle does not have access to an RSU. In each scenario, we change the transmission data rate and evaluate the achievable gain by the proposed scheme compared to the standard IEEE 802.11p protocol in terms of the defined video quality metric (frequency of playback freezes) and its tradeoff with other temporal metric, i.e., start-up delay. The impacts of (i) frequency of channel estimation updates, (ii) inter-RSU distance on access probability, and (iii) quality metrics on system performance are obtained.

Case 1: $0 < L \leq 2R$

The simulation parameters are selected as follows. The inter-RSU distance is $L = 1500m$, the speed of vehicles is fixed at $V_s = 50km/h = 13.88m/s$, maximum Doppler shift according to the speed of the vehicle and the base frequency of the WAVE protocol (5.9 GHz) is calculated to be $f_m = 273.15Hz$ and hence the coherence time is $T_c = 1.5$s. The log-normal shadowing parameters are as follows. The path loss exponent $\alpha = 2$, and variance of the normally distributed noise is $\sigma = 1$ with mean equal to zero. Accurate and timely update of channel state and selection of MAC parameter by consideration of current data rate, received signal strength, and quality of access to RSU in selecting an optimized retransmission limit is the contribution of the proposed adaptation scheme on and above those of the standard MAC protocol. This information can be embedded in the UDP packet header and hence it is compatible with the current IEEE 802.11p protocol.

In this scenario, the RSU maintains access to the destination vehicle while the vehicle gains the opportunity to access the channel. The vehicle is always within the transmission

range of the RSU (R=750m); however, as the vehicle moves away from the responsible RSU, with log-normal shadowing, there is a drop in RSS and hence it is expected that the vehicle experiences higher packet loss. Also, if the distance between the RSU and the vehicle is larger than the vehicle's transmission range, it can not inform the RSU of its RSS and hence the adaptation would be performed only with respect to the transmission rate of the video packet at the transmitter (RSU) side. As the vehicle moves toward the next RSU, the RSS will increase and so does the quality of the video streaming. The proposed adaptive protocol has considerable gain of more than 0.3 to 0.6 freezes/second compared with the standard protocol, which is equivalent to 30 to 60 percent in the number of streaming freezes (Figure 3.5). Such gain in frequency of freezes is achieved while the start-up delay is maintained below 4 seconds (Figure 3.4).

Case 2: $2R \leq L$

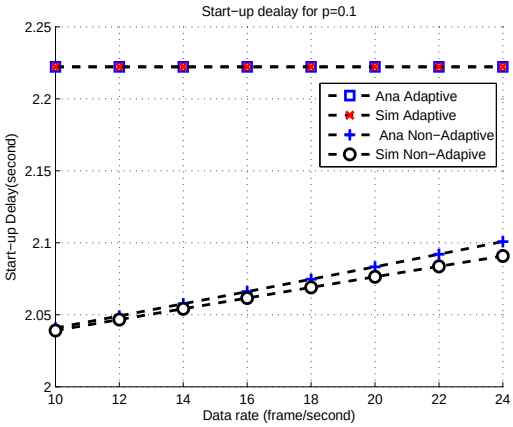
The inter-RSU distance is selected to be $L = 2000m$, which means that there are zones along the road in which the vehicle is not within the communication range of the RSU. Hence, the remaining packets in the receiver's playback buffer decreases with higher pace. Therefore, the likelihood of playback freezes increases during the intervals of low access to RSUs. The proposed adaptive protocol has generally better performance in terms of playback freezes compared to the standard IEEE 802.11p protocol; however, the extra steps for optimization causes the start-up delay to be slightly larger than the standard protocol (at most 2 sec.) which is negligible compared to much lower performance gain achieved in terms of playback freezes that has higher impact on the end user's level of satisfaction.

A zone with no coverage is expanded symmetrically about 500m around the middle of the inter-RSU distance. The access probability drops significantly in this zone and its impact on streaming quality is considerable as shown in Figures 3.6 and 3.7. The

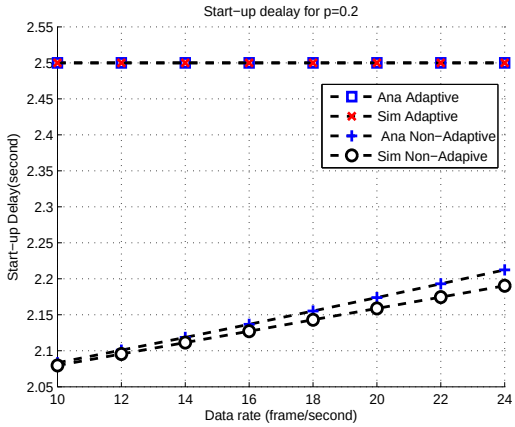
values for frequency of playback freezes have increased compared to the pervious scenario ($0 < L \leq 2R$). Our proposed method outperforms the standard algorithm in such scenarios with more sparse RSU deployment when the connectivity imposes challenge on maintaining high video streaming quality. We have performed experiments to asses the impact of Inter-RSU distance on the performance of the proposed algorithm and its performance compared to that of the standard protocol.

3.3.3 Impact of Inter-RSU Distance

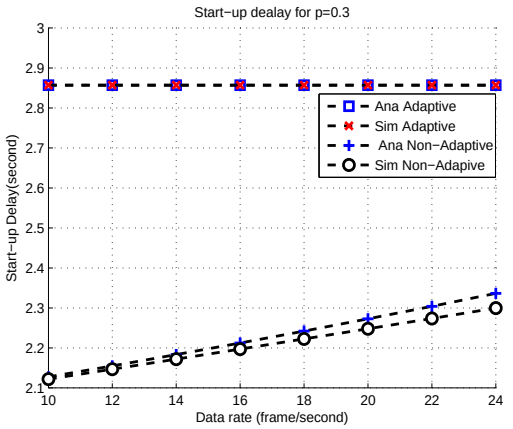
In each step, we change the inter-RSU distance, L, to show the impact of mobility and network access connectivity on the received video quality. The exact analytical results are verified by the simulation results. Fig. 3.8 shows the approximate analytical values which are reasonably close to the simulation results. Same situation can be observed for the result in Fig. 3.9. As L increases, it is more difficult for all vehicles to be connected to the RSUs due to the larger possible distances between the vehicles and the RSUs. This causes a drop in the access probability, and it tends to zero as L goes to infinity. Compared with the standard MAC, our algorithm has advantage in terms of frequency of playback freezes while introducing a slight increase in start-up delay.



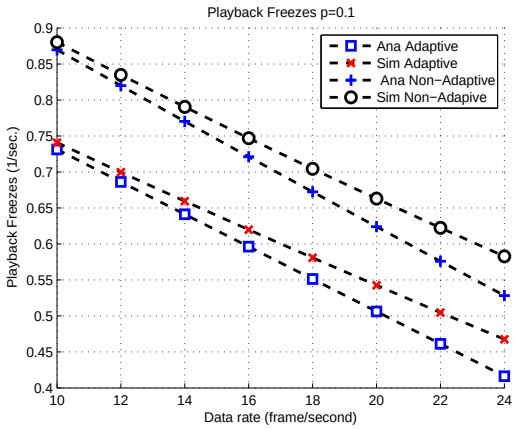
(a) Start-up delay



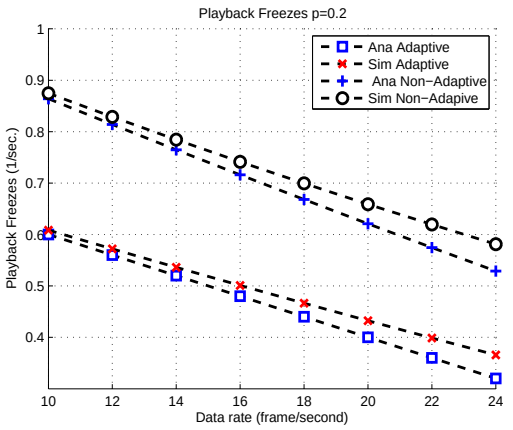
(b) Start-up delay



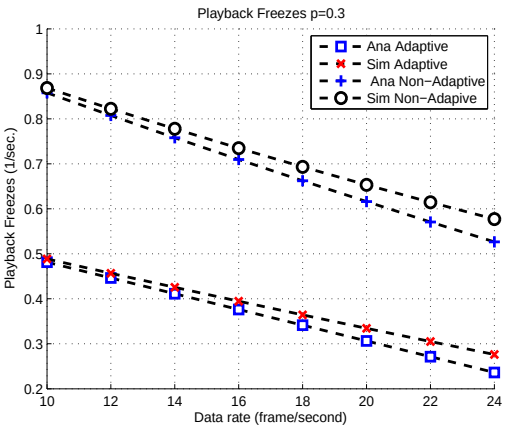
(c) Start-up delay



(d) Frequency of playback freezes



(e) Frequency of playback freezes



(f) Frequency of playback freezes

Figure 3.3: Performance evaluation of the proposed retransmission limit adaptation scheme vs. frame transmission rate

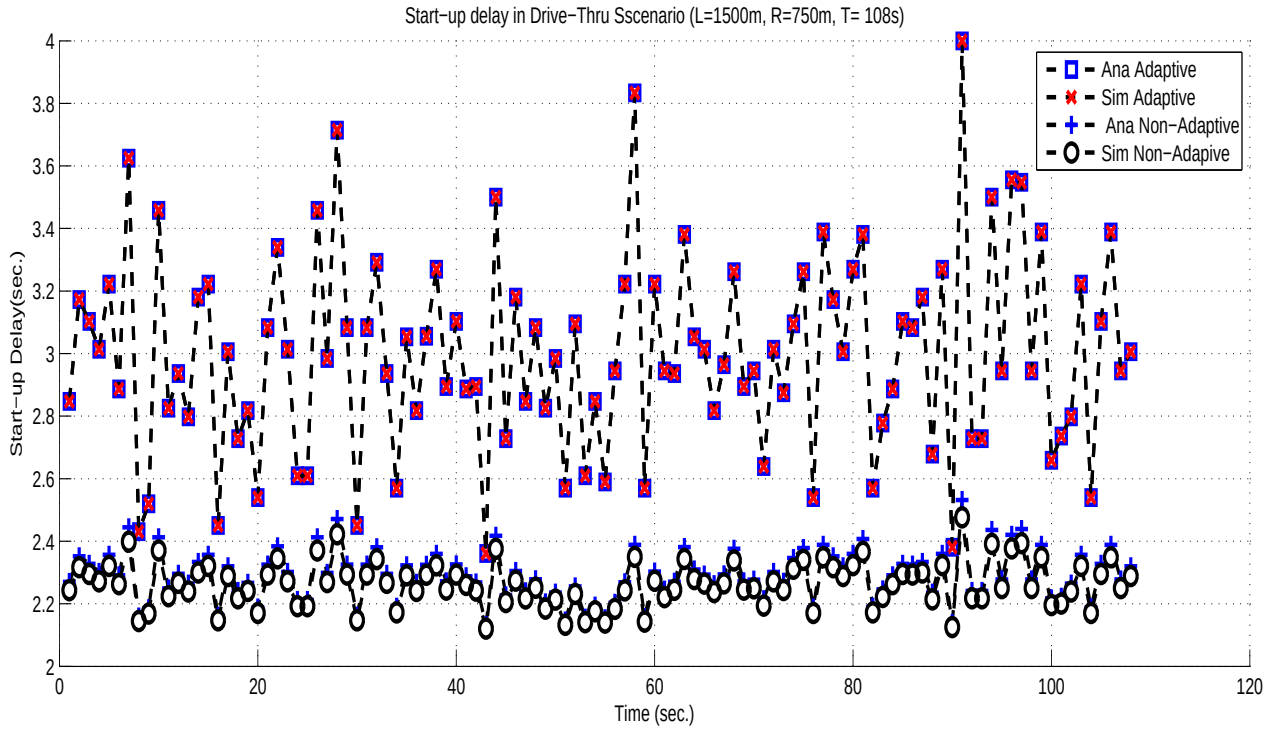


Figure 3.4: Start-up delay of the proposed retransmission limit adaptation vs. IEEE 802.11p in drive-thru scenarios ($L = 1500\text{m}, R = 750\text{m}$)

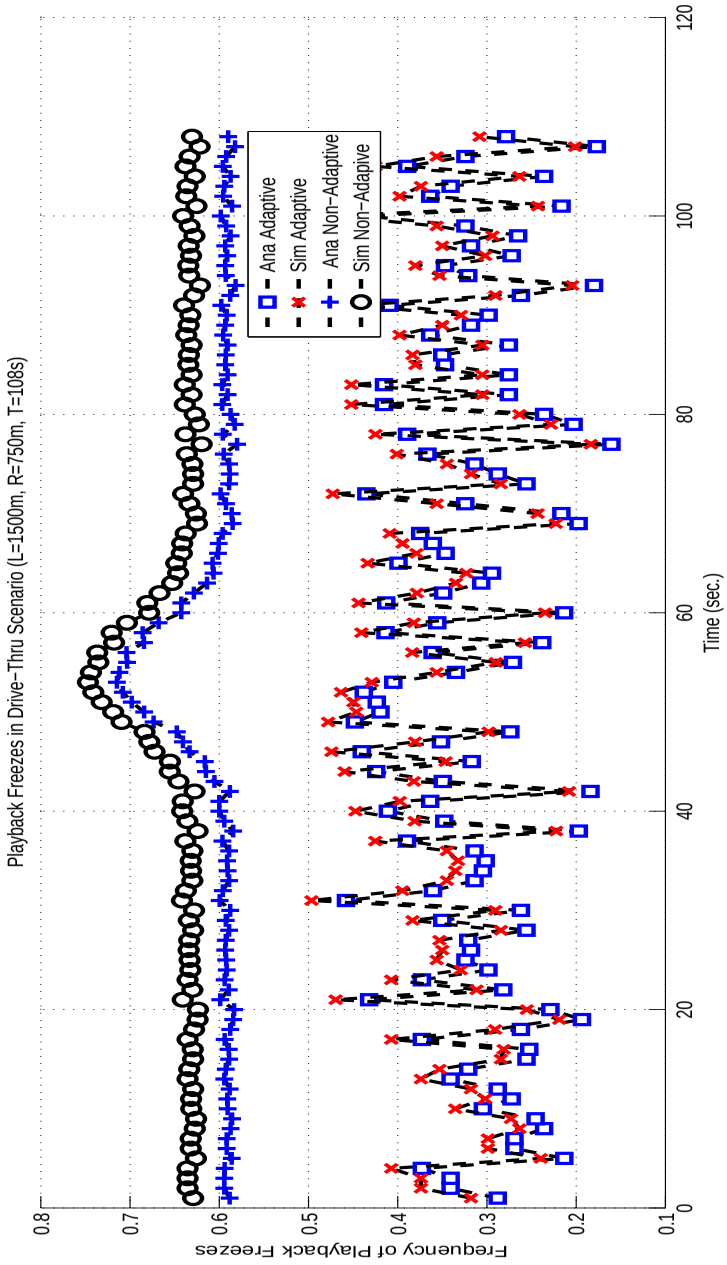


Figure 3.5: Frequency of playback freezes of the proposed retransmission limit adaptation vs. IEEE 802.11p in drive-thru scenarios ($L = 1500\text{m}, R = 750\text{m}$)

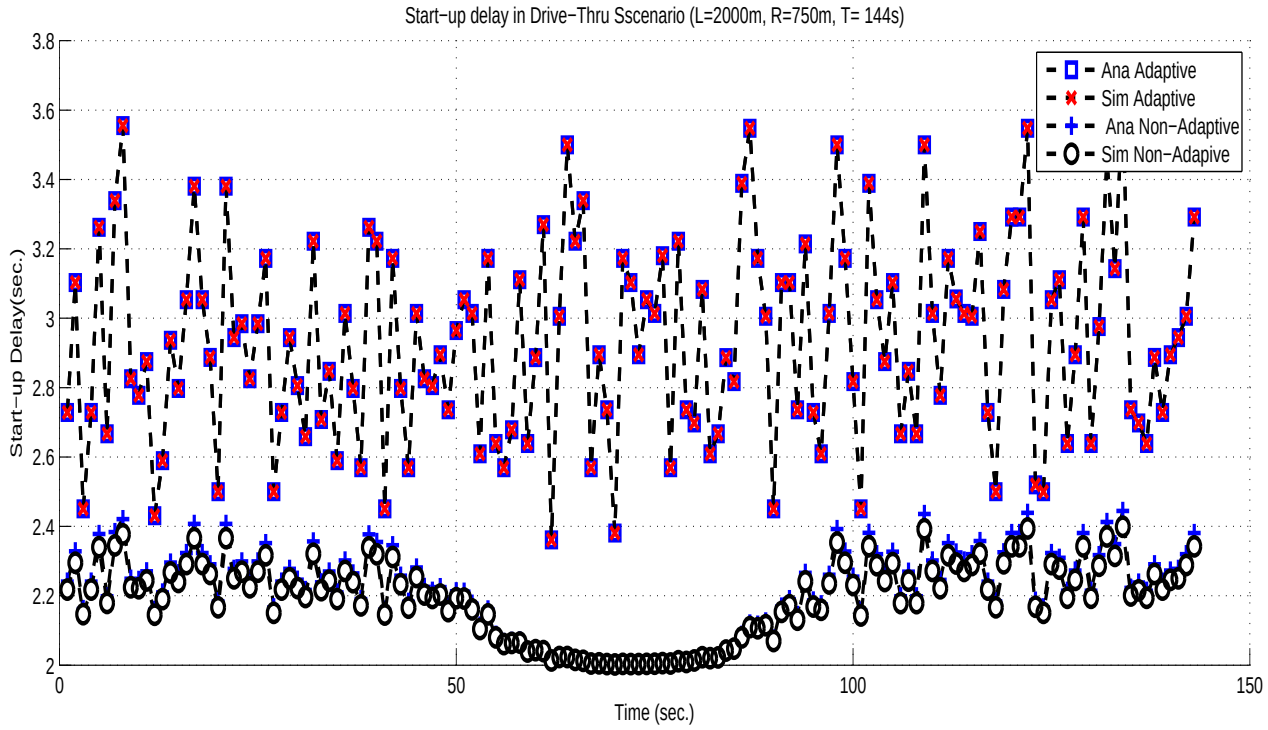


Figure 3.6: Start-up delay of the proposed retransmission limit adaptation vs. IEEE 802.11p in drive-thru scenarios ($L = 2000$m, $R = 750$m)

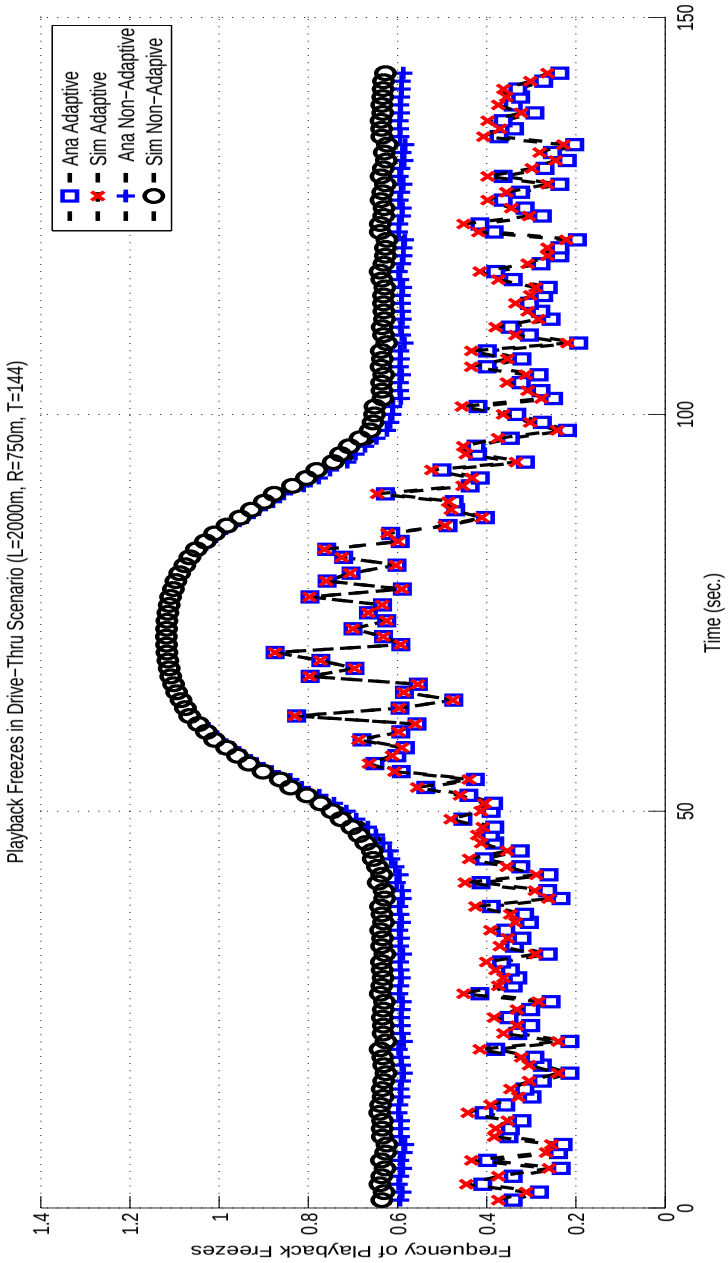


Figure 3.7: Frequency of playback freezes of the proposed retransmission limit adaptation vs. IEEE 802.11p in drive-thru scenarios ($L = 2000\text{m}, R = 750\text{m}$)

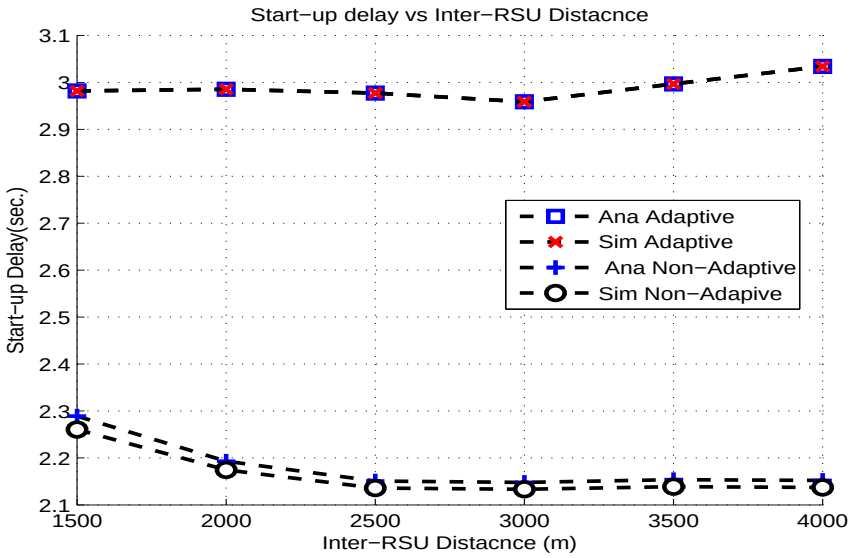


Figure 3.8: Start-up delay vs inter-RSU distance

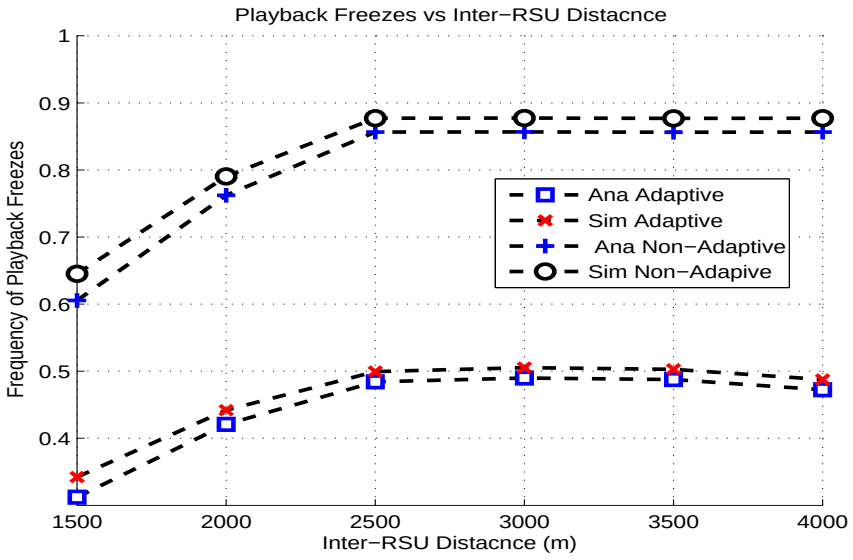


Figure 3.9: Frequency of playback freezes vs inter-RSU distance

3.4 Summary

In this chapter, an adaptive MAC retransmission limit adaptation scheme has been proposed in which the adaptation is based on an optimization of playback streaming quality. A multi-objective optimization framework is applied at the RSU, which jointly minimizes the probability of playback freezes and start-up delay of the streaming at the destination vehicle by tuning the MAC retransmission limit with respect to channel statistics (packet error rate and packet transmission rate). The proposed scheme can achieve significantly fewer playback freezes while introducing a small increase in start-up delay. Future work includes adaptation of other MAC parameters such as contention window (CW) size and consideration of more complex and comprehensive distribution models for deployment of RSUs.

Chapter 4

Application-Centric Routing Protocol over Multi-Hop Urban VANETs

Multi-hop urban vehicular networks in dense traffic conditions have similarities with mesh connected networks plus new challenges arising from highly dynamic nature of the network. First we investigate an application-centric routing scheme for video streaming over mesh connected overlays. Next, we introduce the challenges of urban VANETs compared to mesh networks and extend the proposed scheme in mesh network into a protocol for urban VANETs. Transmitting delay sensitive traffic like video across multi hop networks is a challenging task due to dynamically varying resources with limited support for video QoS [41]. Hence, video streaming solutions must consider time varying bandwidths and loss probabilities of wireless medium. Prior research on video transmission over multi-hop wireless networks has focused on constant video bit rate and exhaustive optimization solutions. In this chapter, we mathematically formulate the video streaming problem with the objective to minimize the video distortion via path selection. This problem has been addressed in [69] for constant bit rate (CBR) video. Given the importance of scalable

video encoders and their variable video bit rate, we extend the work to VBR data which makes the mathematic more complex, which is likely NP-hard. Since the optimization problem does not seem to have any special simplifying structure, a classification-based algorithm is applied to solve this optimization problem. Two major contributions in our work on video streaming over mesh networks are presented. First, we present the optimal path selection problem for video streaming over a multi-hop network as a classical classification problem. Second, we take the self-similar nature of the input traffic into consideration and formulate our problem based on a FBM traffic model, which is an appropriate model for state-of-the-art H.264 codecs.

Service-oriented vehicular networks, which rely on both V2I and V2V communications, can be regarded as a combination of infrastructure-based broadband wireless networks and ad-hoc networks [70]. High and dynamic vehicle movements impose challenging conditions for designing practical communication protocols over VANETs, especially in multi-hop communication. Multi-hop communication increases the transmission range while maintains low transmission power. However, transmitting video over multi-hop wireless networks encounters many challenges, such as unreliable link quality owing to multi-path fading, shadowing, signal interferences among nodes and dynamic connectivity outages. A common classification of the mobility models is based on the level of detail of the motion representation, namely, macroscopic, mesoscopic and microscopic models. Vehicular network simulations require a high level of detail in terms of car motion representation. Hence, microscopic or at most mesoscopic models are more appropriate [15].

In order to extend our work on mesh network to VANETs, we focus on designing a routing protocol for video streaming over multi-hop urban VANETs which can provide high frame quality. Research on routing over VANETs can be classified into several categories. Some existing works assume that routes are known a priori to the transmitter

to facilitate the selection of the best path based on the optimization criterion. However, exhaustive search among available paths imposes high computational complexity for the large set of paths in densely connected VANET. Also, because of rapid changes in topology and connectivity of the links, it is not reasonable to assume that valid paths are known to the transmitter. Hence, the logical conclusion is that a practical routing protocol over VANET must work based on hop-by-hop path selection with few candidate hops to transmit the packets. In our method, a cross-layer protocol is applied at each hop that relates the application layer's video quality metric to the next hop selected vehicle. The objective is to maximize the end user's PSNR of the received video frames, which is equivalent to minimize the end-to-end mean square error (MSE) of the delivered video frames. The vehicle in each hop uses a decode and forward scheme to relay the packets to the next hop without amplifying the noise effects. We limit the computational complexity of the algorithm by decreasing the size of search space at each hop. The carrying vehicle first selects a subset of candidate vehicles to forward its packets. Then, it applies the application-centric optimization to deliver the video frames with high quality.

The remainder of this chapter is organized as follows. Application-centric video packet routing over mesh networks is presented in section 4.1. In section 4.2, extension of the application-centric routing protocol to urban multi-hop VANET scenario is described. Numerical results and discussion for different urban scenarios and comparisons with greedy forwarding routing scheme are presented in section 4.3. Finally, concluding remarks are drawn in section 4.4.

4.1 Application-Centric Video Packet Routing over Mesh Networks

4.1.1 Problem Formulation and Network Topology

As mentioned before, we first propose an application-centric routing over mesh networks and then extend it to VANETs. A multi-hop network can be modeled as a graph with N vertices (nodes) and E edges (links). Each node uses OFDMA at the physical layer and 802.11e standard at the MAC layer. Consider video packets S in the network. Each video packet $\sigma \in S$ has video source z_σ and destination d_σ. The total rate of the video stream for packet σ is bounded by $\underline{R}_\sigma \leq R_\sigma \leq \overline{R}_\sigma$, $\sigma \in S$, while the upper and lower bounds are determined by the encoder used at the source node. The rate R_σ is split among the paths in P_σ.

$$\sum_{h \in P_\sigma} R_\sigma^h = R_\sigma, \quad R_\sigma^h \geq 0, \quad \forall h \in P_\sigma, \quad \forall \sigma \in S \tag{4.1}$$

Table 4.1 summarizes the notation used for formulation of video streaming problem over multi-hop mesh networks. Figure 4.1 shows a multi-hop mesh connected network topology being considered for streaming of video packets from source nodes to destination nodes.

An empirical rate distortion model for motion compensated video encoder has been developed in [71] which concludes for video sequence encoded at a target coding rate R_σ, the average end-to-end distortion D_σ^e consists of the distortion at the encoder D_σ^{enc}, the

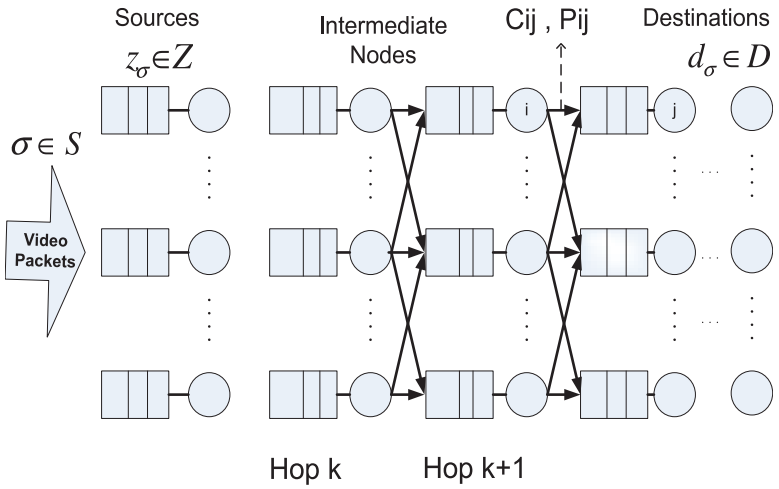


Figure 4.1: Topology of multi-hop mesh connected network for streaming packets from source nodes to destination node

Table 4.1: Notation used for video streaming over multi hop mesh network

Symbol	Definition
N	Set of nodes in network
E	Set of wireless links in network
$\{i,j\}$	Link from node i to j
$c_{i,j}$	Capacity of link $\{i,j\}$
p_{ij}	Packet loss probability on link $\{i,j\}$
λ_{ij}	Average aggregate traffic rate on link $\{i,j\}$
ρ_{ij}	Utilization of link $\{i,j\}$
t_{ij}	Delay on link $\{i,j\}$
$f_{ij}(x)$	Probability density function of t_{ij}
$M_{ij}(s)$	Moment generating function of t_{ij}
M_σ^h	Moment generating function of T_σ^h
S	Set of video packets in the network
z_σ	Source node of packet σ
d_σ	Destination node of packet σ
P_σ	Path set of packet σ from z_σ to d_σ
P_σ^h	A path in the set P_σ from z_σ to d_σ
Δ_σ	Decoding deadline for packet σ
T_σ^h	End-to-end delay on path $P_\sigma^h \in P_\sigma$
p_σ^h	End-to-end loss rate of $P_\sigma^h \in P_\sigma$
R_σ^h	Rate of video packet σ on path $P_\sigma^h \in P_\sigma$
R_σ	Rate of video packet σ
$\overline{R}_\sigma$	Maximum rate of video packet σ
$\underline{R}_\sigma$	Minimum rate of video packet σ
D_σ^e	End to end distortion of packet σ
D_σ^{enc}	Distortion of packet σ due to encoding
D_σ^{cg}	Distortion of packet σ due to congestion
D_σ^{loss}	Distortion of packet σ due to packet loss

distortion due to congestion D_σ^{cg} and the distortion due to packet loss D_σ^{loss}, i.e.,

$$D_\sigma^e = D_\sigma^{enc} + D_\sigma^{cg} + D_\sigma^{loss} \tag{4.2}$$

According to [71], encoding distortion can be modeled as

$$D_\sigma^{enc} = D_0 + \frac{\theta}{R_\sigma - R_0} \tag{4.3}$$

where θ, D_0 and R_0 are parameters for specific video encoder and video sequence. Congestion happens when packet is not lost, but its delivery exceeds packet delay deadline. Denote T_σ^h as delay on path P_σ^h. T_σ is the total end to end delay of all paths for packet σ.

$$T_\sigma = \frac{\text{total bit}}{\text{total rate}} = \frac{\sum_{h \in P_\sigma} R_\sigma^h \cdot T_\sigma^h}{R_\sigma}. \tag{4.4}$$

Since the delay on path T_σ^h is a summation of all the delays in the links of the path, we have

$$T_\sigma^h = \sum_{\{i,j\} \in P_\sigma^h} t_{ij}. \tag{4.5}$$

An upper bound for distortion due to congestion delay can be formulated as

$$D_{cg} = k \sum_{h \in P_\sigma} \frac{R_\sigma^h}{R_\sigma} (1 - p_\sigma^h) Pr(T_\sigma^h > \Delta_\sigma). \tag{4.6}$$

where k is a constant parameter in the time domain which depends on video encoder and video sequence. p_σ^h is loss probability of packet σ over path P_σ^h and can be computed as

$$p_\sigma^h = 1 - \prod_{\{i,j\} \in P_\sigma^h} (1 - p_{ij}), \quad \forall h \in P_\sigma, \quad \forall \sigma \in S \tag{4.7}$$

where p_{ij} is the loss probability of link $\{i, j\}$. So we assume that some channel estimation process is in place to estimate p_{ij}. Distortion caused due to packet loss can be written as

$$D_\sigma^{loss} = k \sum_{h \in P_\sigma} \frac{R_\sigma^h}{R_\sigma} p_\sigma^h \tag{4.8}$$

Equations (4.2), (4.3), (4.6) and (4.8) lead to end-to-end distortion

$$\begin{aligned} D_\sigma^e = D_0 + \frac{\theta}{R_\sigma - R_0} + k \sum_{h \in P_\sigma} \frac{R_\sigma^h}{R_\sigma} \\ (1 - p_\sigma^h) Pr(T_\sigma^h > \Delta_\sigma) + k \sum_{h \in P_\sigma} \frac{R_\sigma^h}{R_\sigma} p_\sigma^h \end{aligned} \tag{4.9}$$

Hence, we need to find the components in the above equation by deriving the link delay statistics. The traffic on link $\{i, j\}$ is the aggregate traffic from different paths that traverse the link. Let $P_\sigma^{h,ij}$ be the partial path for path P_σ^h up to link $\{i, j\}$. The average rate of the aggregate traffic on link $\{i, j\} \in L$ is

$$\lambda_{ij} = \sum_{\sigma \in S} \sum_{h \in P_\sigma} [\prod_{\{l,m\} \in P_\sigma^{h,il}} (1 - p_{lm})] \cdot R_\sigma^h \tag{4.10}$$

In other words, the traffic rate on a link is the sum of the rates of video packets that pass through the link minus the loss occurred on their path before reaching that link. The utilization of link $\{i, j\}$ is $\rho_{ij} = \frac{\lambda_{ij}}{c_{ij}}$.

Studies show that the traffic produced by VBR video encoder such as H.264, which is the subject of our studies, is self-similar (SS) in nature [72]. For such traffic, FBM input model has been introduced, mostly due to mathematical simplicity. A standard FBM process with Hurst parameter $H \in [\frac{1}{2}, 1)$ is a Gaussian centered process.For VBR video

traffic that exhibits LRD characteristics as input load on a link, if the service rate of each link is c_{ij}, we could model the queueing delay probability by a Weibull distribution [72] given by Norros [73] as

$$P(X > x) \sim \exp\left(-\frac{(C_{ij} - \lambda_{ij})^{2H}}{2k^2(H)c_v\lambda_{ij}^2}x^{2-2H}\right) \tag{4.11}$$

where λ_{ij} is the input traffic average rate, $c_v = \sigma_{ij}/\lambda_{ij}$ is the standard deviation divided by the mean input traffic load and $k(H) = H^H(1-H)^{1-H}$ where H is the Hurst parameter which appears in FBM distributions. The Hurst parameter can be determined via fitting a Weibull distribution to the queueing delay probability and using (4.11).

We first investigate a general upper bound for an arbitrary Weibull distribution using the Chernoff bound. Assume that we have the following Weibull distribution:

$$f_{ij}(t) = ba^{-b}t^{b-1}\exp(-(\frac{t}{a})^b) \tag{4.12}$$

where parameters b and a are the scale and location parameters, respectively. The Weibull distribution is close to a normal distribution. For $a \leq 1$ the density function of the distribution is L shaped and for values of $a > 1$, it is bell shaped. This distribution gives a failure rate increasing with time. For $a > 1$, the failure rate decreases with time. At $a = 1$, the failure rate is constant and the lifetimes are exponentially distributed.

For an independent (and not necessarily identically distributed) sequence of random variables, $\{X_i : i = 1...n\}$ and

$$S_n = \sum_{i=1}^{n} a_i X_i, \tag{4.13}$$

where the a_i's are constants, then the probability density function for S_n is the convolution of the probability density functions of each of the X_i and the moment-generating function

for S_n is given by

$$M_{S_n}(t) = M_{X_1}(a_1 t) M_{X_2}(a_2 t) \cdots M_{X_n}(a_n t) \tag{4.14}$$

Using the above and equation (4.5), we can find the moment generating function for T_σ^h which is needed for finding the Chernoff bound for $Pr(T_\sigma^h > \Delta_\sigma)$. For data links with long range dependent data such as VBR video, each t_{ij} is modeled as Weibull distributed. Knowing the moment generating function for t_{ij}, we can compute the moment generating function of T_σ^h. The moment generating function of t_{ij} with Weibull distribution is given by

$$M(t_{ij}) = E(e^{s \cdot t_{ij}}) = \sum_{n=0}^{\infty} \frac{s^n a^n}{n!} \Gamma(1 + \frac{n}{b}) \tag{4.15}$$

hence,

$$M(T_\sigma^h) = \prod_{\{ij\} \in P_\sigma^h} M(t_{ij}) = \prod_{\{ij\} \in P_\sigma^h} [(\sum_{n=0}^{\infty} \frac{s^n a^n}{n!} \Gamma(1 + \frac{n}{b})] \tag{4.16}$$

Clearly, we have assumed that the delays of different links are i.i.d random variables. The Chernoff bound for a random variable X, which is the sum of n independent random variables $X_1.X_2, ..., X_n$, is obtained by applying $E[e^{sX}]$ for some well-chosen value of s.

$$Pr(X \geq \Delta) \leq \min_{s>0} \frac{\prod_i E[e^{sX_i}]}{e^{s\Delta}}. \tag{4.17}$$

Hence,

$$Pr(T_\sigma^h \geq \Delta_\sigma) \leq \min_{s>0} \frac{\prod_{\{ij\} \in P_\sigma^h} E[e^{st_{ij}}]}{e^{s\Delta_\sigma}} \tag{4.18}$$

which leads to

$$Pr(T_\sigma^h \geq \Delta_\sigma) \leq \min_{s>0} \frac{\prod_{\{ij\}\in P_\sigma^h}[(\sum_{n=0}^{\infty} \frac{s^n a^n}{n!}\Gamma(1+\frac{n}{b})]}{e^{s\Delta_\sigma}}. \tag{4.19}$$

Let us define

$$F_\sigma^h(s) = \frac{\prod_{\{ij\}\in P_\sigma^h}[(\sum_{n=0}^{\infty} \frac{s^n a^n}{n!}\Gamma(1+\frac{n}{b})]}{e^{s\Delta_\sigma}} \tag{4.20}$$

and $s^*_{\sigma,h}$ is the value of $s > 0$ for which the above function is minimized.

We can now formulate the optimization problem for optimal path of streaming VBR scalable video over a multi-hop mesh network as a minimization problem. The optimal solution can be obtained by solving the following minimization of the end-to-end distortion:

$$\begin{aligned} Min\{D_{end} = \sum_{\sigma\in S}\{D_0 + \frac{\omega}{R_\sigma - R_0} + k\sum_{h\in P_\sigma}\frac{R_\sigma^h}{R_\sigma}\{p_\sigma^h + \\ (1-p_\sigma^h)\{\frac{\prod_{\{ij\}\in P_\sigma^h}[(\sum_{n=0}^{\infty}\frac{(s^*_{\sigma,h})^n a^n}{n!}\Gamma(1+\frac{n}{b}))]}{e^{s^*_{\sigma,h}\Delta_\sigma}}\}\}\}\} \end{aligned} \tag{4.21}$$

subject to

$$R_\sigma = \sum_{h\in P_\sigma^h} R_\sigma^h \tag{4.22}$$

$$\underline{R} \leq R_\sigma \leq \overline{R} \tag{4.23}$$

$$0 \leq R_\sigma^h \leq R_{max}^h \tag{4.24}$$

$$\lambda_{ij} = \sum_{\sigma\in S}\sum_{h\in P_\sigma}[\prod_{\{l,m\}\in P_\sigma^{h,il}}(1-p_{lm})]\cdot R_\sigma^h \leq c_{ij} \tag{4.25}$$

The Gamma functions in the last term of equation (4.21) make this minimization

problem very difficult to solve. Hence, we refer to [72] and use its result in order to simplify the problem. Specifically, our solution is based on minimizing video distortion over each hop which would consequently result in end-to-end distortion optimization. The optimization is done by a trained classifier deployed at each hop, with the classifier trained for that particular hop. The analysis given in [72] can be used to simplify the congestion distortion term to be discussed in subsection 4.1.2.

4.1.2 Solution Procedure

The above optimization problem is NP-hard, nonlinear, and non-convex, which is not solvable by methods normally applied in convex optimization problems. Since each hop has its own characteristics, the end-to-end minimization problem may be decomposed into per hop minimization. It has been suggested in [69] to linearize the objective function and then use a branch and bound technique to solve the linear optimization problem. We do not attempt to linearize the minimization problem in (4.21). Our approach simplifies the real-time optimization problem using classification. The complexity associated with performing optimization in real-time is very high. Thus, low complexity systems are required for determining optimal rate allocation in real-time whenever a packet needs to be transmitted. We discuss such a possible approach based on classification and machine learning techniques. Applying classification methods in optimization over wireless networks has been suggested in [11] to determine optimal retry limit at the MAC layer in a local area network. We formulate the problem of minimizing the utility (end-to-end video quality distortion) as a standard classification problem. In order to apply a classifier at each hop, we follow these steps:

1. ***Offline generation of ground truth:*** Collect a set of packets with a variety of encoder parameters under different channel conditions. The extracted features

include VCF extracted from statistics of video sequences, VEP determined based on encoder configuration/parameter set. Class prototypes includes WCC. WCC include maximum bandwidth R_{max} and error probability P_e of the next hop that video packets are going to be transmitted through. The features are selected according to formula (4.21). It shows that packet type, encoding and channel features have an impact on video distortion. The selected VEP feature is intra-coding percentage, β, which is used to estimate ω, R_0 and D_0. WCC features include P_{ij} and R_{ij}^{max}. VCF features are explained in detail in the next subsection. The distance measure is distortion per hop which is based on MSE or PSNR. The formula for the distance measure is provided in subsection 4.1.2.

2. ***Training the classifier:*** at this step, the goal is training classifier parameters to map each packet j to a class label L_j which corresponds to optimal link selection (partial path selection). To describe it in simple language, we are classifying each feature vector including VCF and VEP to a class presented by prototype WCC.

3. ***Real time strategy selection:*** based on classification results the optimized strategy is used to determine the optimized parameters and configurations of wireless multimedia system.

Video Content Features (VCF)

We follow the video sequence study results in [74]. The video sequences used for the statistics presented are the ten minute Sony Digital Video Camera Recorder demo sequence (17,682 frames at 30 frames/sec), the first half hour of the Silence of the Lambs movie (54,000 frames at 30 frames/sec), the Star Wars IV movie (54,000 frames at 30 frames/sec), and the first hour of the Tokyo Olympics video (133,128 frames at 30 frames/sec), 30 minutes of the NBC 12 News (49,523 frames at 30 frames/sec), includ-

ing the commercials. The video sequences Silence of the Lambs, Star Wars IV, Tokyo Olympics, and NBC 12 News can respectively be described as drama/thriller, science fiction/action, sports, and news video. Choosing a variety of video sequences with a variety of natures and with enough length is recommended in [5]. Key frame size, GOP size and quality statistics for H.264/AVC, H.264/SVC and MPEG-4 encoders are derived in [74] using the above mentioned video sequences. VEP features affect the last term (Gamma function) of the distance measure in (4.21).

A right strategy for selecting video features would be to compute mutual information (MI) between any two features used in feature vectors and selecting those features with less MI. It is important to avoid very large feature sets in order to avoid high complexity. The distortion formula presented in the pervious section deals with queueing delay which itself is relevant to input traffic bit rate. We have selected mean traffic bit rate as one of the features for training the classifiers; other potential feature is Hurst parameter which appears in the distortion formulation and can be estimated for each video sequence. There are many ways to estimate Hurst parameter [75] [76].

Video Encoding Parameters (VEP)

In our video streaming architecture, we use a rate distortion (RD) model to analyze the performance of video encoder. For motion compensated video encoders like H.263, H.264, and MPEG-4, empirical self similar RD models obtained from actual video sequences are more practical compared to theoretical models [71]. Hence, we use a self-similar RD model and derive the required parameter for each video sequence using data fitting. The fitting was done by minimizing the sum of squared MSE differences between the model and the measured points. This resulted in sets of parameters $\{\theta, R_0, D_0\}$ for each encoder parameter settings. It is illustrated in [71] that this parameter set is related to the intra coding percentage factor β. That is, each of this parameter set corresponds

to a specific value of β. Hence, we use this parameter as one of our features and avoid very high dimensional classification problem that would likely result in high complexity. VEP parameters affect the first term in (4.21).

Wireless Channel Condition (WCC)

We exclude the WCC features as they are common to all the packets. Channel loss probability P_L and maximum channel rate R_{max} are considered as WCC features for classification of video packets. In our simulations, we take R_{max} uniformly distributed in $[0,\ 200]$ kbps and p_L uniformly distributed in $[0,\ 0.2]$.

Classifier Design

Our optimization problem is about specifying optimal link at each hop for each packet such that the expected overall decoded distortion is minimized. At each hop, a classifier decides the next link (class) for each video packet (data sample to be classified). Hence, at each hop the number of classes is equivalent to the number of available links at that hop (between that hop and next hop). Each classifier is trained for only one hop and classification is done per packet. We encounter a multi-class classification problem which in general is more challenging than the two-class classification problems. At each hop only one link can be selected as an optimal partial path for one video packet. We first reformulate (4.21) for a one-hop case to get some insight on optimal transmission link at each hop for packet σ. This formulation serves as a distance measure. Our goal is to minimize this distance (distortion) at each hop for each video packet. It is shown in [72] that the end-to-end delay of the tagged flow in both tandem and tree queueing networks is completely dominated by one of the queues. The dominant queue is the one with the maximal Hurst parameter. If several queues have the same maximal Hurst parameter, then we have to compare the ratio $\frac{(1-\rho)^H}{\sigma}$ to determine the dominant queue, where ρ is

the mean load of the queue and σ is the standard deviation of the load.

$$P(W > x) \sim P(D > x) \tag{4.26}$$

which means that $P(D > \Delta_k)$ is also Weibullian where $\Delta_k = \Delta/K$ is the delay deadline of hop k and K is the total number of hops. The results further imply the round-trip time (RTT) of transport control protocol (TCP) sessions is asymptotically Weibullian. They also provide upper and lower bounds on the constant that determines the scale parameter of the corresponding Weibull distribution.

$$\begin{aligned} D_{\sigma}^{hop} = D_0 + \frac{\omega}{R_e - R_0} + \sum_{ij} \frac{R_{\sigma}^{ij}}{R_{\sigma}} \cdots \\ \cdots \left\{ p_{\sigma}^{ij} + (1 - p_{\sigma}^{ij}) \exp\left(-\frac{(C_{ij} - \lambda_{ij})^{2H}}{2k^2(H) c_v \lambda_{ij}^2} (\frac{\Delta}{K})^{2-2H} \right) \right\} \end{aligned} \tag{4.27}$$

where R_e is the encoder output rate. Hence, we can interpret the above distortion formula (distance measure for classification) as follows.

$$D_{\sigma}^{hop} = f(\text{VCF,VEP,WCC}) \tag{4.28}$$

where σ is the packet index and h is the index for each hop. The encoding distortion term is only considered once, i.e., at the first hop. A support vector machine classifier is applied at each hop to deal with the multi-feature multi-class classification problem. For more detailed information about the basics of support vector machines and their application in pattern classification, the reader is referred to [77]. Here, we discuss in full detail how we train and apply support vector machine (SVM) (appendix C) in the path selection problem formulated in the pervious section. Given the labeled training data

$$\mathcal{D} = \{(\vec{x}_i, y_i)\}_{i=1}^{\ell}, \quad \vec{x}_i \in \vec{X}, \quad y_i \in \vec{Y}$$

where $\vec{x}_i$ is the data sample (feature vector) to be classified and y_i is the class label, the support vector machine constructs a maximal margin non-linear classifier in high dimensional feature space.

4.1.3 Simulation Results

We analyze the performance of our algorithms in terms of the received video quality (average end-to-end distortion) of the various video packets. The sample elementary multi-hop structures (Figure 4.2) are selected in our simulation for demonstration. We assume that each hop has its own characteristics and is independent from other hops since nodes at each hop have maximum transmission range up to the next hop and would not interfere with further hops. We assume that we have a fixed architecture as shown perviously in Figure 4.1. We also assume that the total number of hops is fixed and is given. We compare our simulation results based on the proposed classification method for optimal path selection with the work in [78], which is a cross-layer approach based on priority queueing for video streaming in a similar architecture as ours.

In the simulations, each video session has rate between 20 kbps and 200 kbps. We use H.264 codec and the first 200 frames of sequence "Foreman" trace. The video is encoded at an intra rate of 3. The rate distortion parameters are obtained from [71]. The bandwidth of a link is selected from a uniform distribution between [50, 400] kbps. The proposed solution procedure is implemented in Matlab. In each sets of simulation we change one of the parameters to analyze the distortion performance of the multi-hop video streaming network against that parameter. For the architectures in Fig. 2, the delay deadline is assumed to be 0.5 second and the Hurst parameter is 0.88. Table 4.2

shows that the average distortion performance is close to those of the state-of-the-art method [78]. Table 4.3 and Table 4.4 show how we can tune the classification parameters to affect the performance. Table 4.5 shows the performance versus increase in the number of the hops (network size).

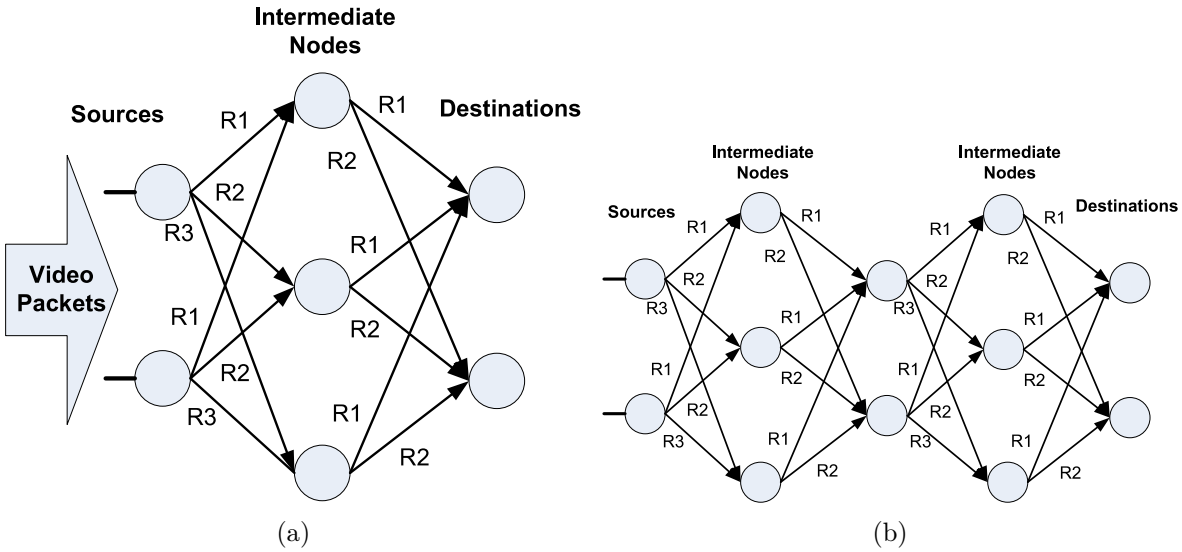


Figure 4.2: (a) Network design for the elementary structure and (b) Network design for a cascade of the elementary structure

Table 4.2 gives end-to-end video distortion using classification-based path selection scheme. The selected encoding feature is fixed ($\beta = 1$). We change the WCC conditions and record the performance of our scheme in terms of end-to-end distortion. In general, Table 4.2 shows that applying our scheme results in very close performance with other state-of-the-art methods such as [78]. In the following tables, we investigate how tuning classification parameters would affect the performance results in our proposed scheme. The path with significantly higher capacity is more often selected by our scheme as the optimal path. Acquired results are in the range of 29.8440 dB (lowest distortion) to 30.4894 dB (highest distortion). In Table 4.3, the impact of changing encoding parameter, β, is considered. As the intra encoding percentage increases, the distortion decreases at

Table 4.2: Average distortion performance results for video "Foreman" versus different WCC parameters. ($\beta = 1$) and ($R_e = 60kbps$)

Max Rate (kbps)			Link loss probability			Distortion (dB)
R_1	R_2	R_3	p_1	p_2	p_3	D_{total}
46	121	97	0.08	0.07	0.04	31.5
130	164	88	0.06	0.07	0.09	31.6
147	35	81	0.09	0.09	0.04	31.3
178	11	70	0.08	0.00	0.03	31.1
40	39	120	0.10	0.19	0.01	31.7
149	89	186	0.04	0.01	0.08	31.6

Table 4.3: Average distortion performance results for video "Foreman" versus intra-rate parameter β. ($R_e = 60$, $R_1 = 46$, $R_2 = 121$, $R_3 = 97$) kbps, $p_1 = 0.08$, $p_2 = 0.07$ and $p_3 = 0.15$.

Intra Rate (β)	1	3	6	11	22	33
Dist. (dB)	32.9	32.7	32.3	31.7	30.4	29.1

Table 4.4: Average distortion performance results for video "Foreman" versus Hurst Parameter. $\beta = 11$,($R_e = 60$, $R_1 = 46$, $R_2 = 121$, $R_3 = 97$) kbps, $p_1 = 0.08$, $p_2 = 0.07$ and $p_3 = 0.15$.

Hurst	0.880	0.895	0.883	0.910	0.546	0.721
Dist.(dB)	31.5367	31.5348	31.5363	31.5330	31.5519	31.5510

Table 4.5: Average distortion performance for video "Foreman" versus different WCC parameters for 4-hop structure which is cascade of elementary network structure. ($\beta = 1$) and ($R_e = 60$kbps)

Max rate (kbps)			Loss probability			Distortion (dB)
R_1	R_2	R_3	p_1	p_2	p_3	D_{total}
46	121	97	0.08	0.07	0.04	32.0
130	164	88	0.06	0.07	0.09	32.3
147	35	81	0.09	0.09	0.04	31.7
178	11	70	0.08	0.00	0.03	31.3
40	39	120	0.10	0.19	0.01	32.5
149	89	186	0.04	0.01	0.08	32.1

the cost of higher complexity. Table 4.4 shows the impact of Hurst parameter (VCF feature) which is related to the choice of video sequence. Results show that end-to-end distortion has low sensitivity to the choice of video sequence. The results of simulation on a 4-hop architecture are shown in Table 4.5. Comparison of the results in this table with those in Table 4.2 shows that from hop-to-hop we have a slight increase in total distortion. Basically, at the price of higher distortion, the network can cover a larger area.

Our observations indicate that the major contribution to distortion comes from the encoding distortion term. The end-to-end distortion is not very sensitive to the Hurst parameter while it is more sensitive to intra coding parameter. Comparison of our work with other proposed solutions in the literature, such as [78], shows that the results are very close. Since the most complex part, i.e. feature extraction and classifier training, is done off-line in our method, the complexity of our method is lower, resulting in shorter delay in real time video streaming.

4.2 Application-Centric Video Packet Routing over Urban VANETs

4.2.1 Network Topology and System Model

Network Topology

Figure 4.3 shows the topology of the considered VANET. RSUs are connected to the video server via wired connections. Each RSU transmits video packets to the destination vehicle through multi-hop communication where the intermediate vehicles act as communication hops to relay the video packets to the destination vehicle. Switching of the transmission between RSUs over a video session is controlled by an access router (AR) which controls several RSUs. When the vehicle exits the coverage range of an AR, handover happens and a new AR controls the transmission via a new set of RSUs which cover the current location of the destination vehicle. The wireless links between RSU and vehicles (V2I links) and vehicle to vehicle (V2V links) are assumed to be memoryless channels with Rician fading. The vehicular network is based on 5.9 GHz DSRC [47] standard and each vehicle is broadcasting information of position, current time, direction, speed, acceleration/deceleration, and traffic events. Hence, the RSU and vehicles can update information regarding locations and direction of the movement of the neighbor nodes (vehicles).

Mobility Model

A network scenario of vehicular mobility for deriving stochastic properties of traffic over straight ways and intersection is shown in Fig. 4.4. Mobility of vehicles can be considered independent from each other in sparse traffic scenarios, however, for dense traffic situations, this assumption is not correct. In order to model vehicular mobility, we

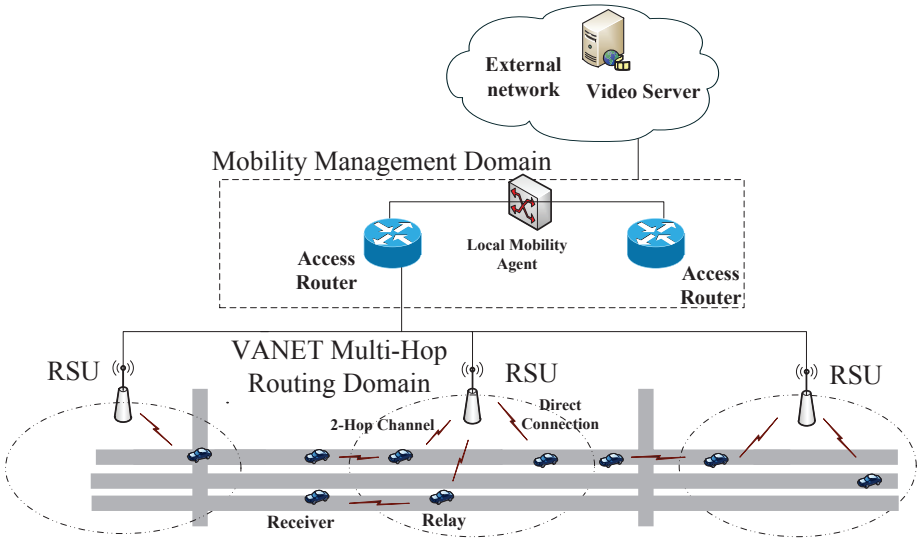


Figure 4.3: Network topology: RSUs receive video from VOD server via wired Internet connection. The RSU is the video streaming source over the wireless part of the network. The vehicles act as communication hops to relay the video packets to the destination vehicles.

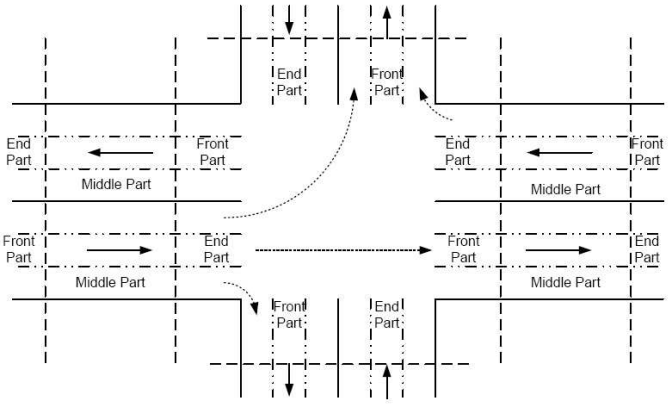


Figure 4.4: A network scenario of vehicular mobility

follow the stochastic model in [63], which is capable of including the behavior of vehicles separately in sparse and dense scenarios.

In this model, different speed categories are considered at each section of the street separated by dashed lines. In this respect, the mobility can be modeled by a network of service queues (sections of streets) in cascade and each queue has customers with several classes (speed categories). At each queue, service rate is equivalent to the rate of passing of vehicles from current section to the next section of the street, i.e., a vehicle with higher speed is associated with the customer class with higher service rate.

In a sparse scenario, vehicles have independent mobility patterns. Therefore, each queue is an $M/G/\infty$ queue, i.e., the service rate does not depend on the number of customers in the queue (number of vehicles in that part of the street). When all parts of the street are in a sparse situation, we actually have a queueing network comprising of $M/G/\infty$ queues with two, three, and two customer classes, corresponding to the front, middle, and end parts, respectively. A section of the street with dense traffic can be modeled by $M/G(n)/1$ queue. The mobility model is used to extract the probability of connectivity between two vehicles which is later used to formulate the QoS objective function to be optimized in the proposed application-centric routing protocol. The prob-

ability of connectivity is defined as the probability that the vehicles along the streets can communicate with each other through a radio link. In general, it is possible to have several radio links between two vehicles. In addition, each radio link may consist of multiple hops. Two vehicles are able to directly communicate if they are within the radio transmission range of each other. The lower and upper bounds for the probability of connectivity at a typical street is given in [63]. The bounds for probability of connectivity are related to the state probabilities. State probabilities indicate the distribution of vehicles in the street. For example, if we divide the street into smaller sections (sub-streets), the probability that there are n_1 vehicles in section 1, there are n_2 vehicles in section 2 and so on is called the state probabilities. The difference between sparse and dense scenario comes from difference in state probabilities for these two scenarios. Hence, we discuss derivation of state probabilities for sparse and dense scenarios, separately. According to [63], the upper and lower bounds for probability of connectivity is as follows.

$$\begin{aligned} P(\text{Connectivity}) &\leq \sum_{\substack{n_i \geq 1 \\ 1 \leq i \leq N_S(R)}} \pi(n_1, n_2, ..., n_{N_S(R)}) \\ &= \prod_{i=1}^{N_S(R)} (1 - \pi(n_i = 0)) \\ &= \prod_{i=1}^{N_S(R)} \sum_{n_i \geq 1} \pi(n_i) \end{aligned} \tag{4.29}$$

$$\begin{aligned} P(\text{Connectivity}) &\geq \sum_{\substack{n_i \geq 1 \\ 1 \leq i \leq N_S(R/2)}} \pi(n_1, n_2, ..., n_{N_S(R/2)}) \\ &= \prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i) \end{aligned} \tag{4.30}$$

where R is the transmission range and $N_S(R)$ is the total number of subsections of length R along the street and n_i is the number of vehicles in section i. The graphical illustration for deriving the above bounds for probability of connectivity is provided in Appendix A.

4.2.2 State Probabilities for Sparse and Dense scenarios

1) Sparse traffic scenario: Let α_{ju} be the arrival rate of class u customers at node j in the network, which represents an exogenous arrival rate as well as the arrivals routed from other nodes. The routing probabilities are determined according to the mobility patterns [63]. The spatial traffic distribution for the queueing network (street) comprising $M/G/\infty$ nodes can be obtained as follows [63] [79]:

$$\pi(n) = \prod_{j=1}^{N} \pi_j(n_j), \quad n_j = (n_{j1}, ..., n_{jC(j)}) \tag{4.31}$$

$$\pi_j(n_j) = e^{-\rho j} \prod_{u=1}^{C(j)} \frac{\rho_{ju}^{n_{ju}}}{n_{ju}!} \quad \rho_{ju} = \frac{\alpha_{ju}}{\mu_{ju}}, \quad \rho_j = \sum_{u=1}^{C(j)} \rho_{ju} \tag{4.32}$$

where μ_{ju} is the service rate of class u customers at node j. $\pi(n)$ represents the probability of n_{11} class 1 customers at node 1, n_{12} class 2 customers at node 1, and so on. By (4.32), we obtain spatial traffic distribution comprised of stationary probabilities for different numbers of customers of different speed categories in different parts of the streets.

2) Dense traffic scenario: With respect to the preceding discussions, a street in a dense situation cannot be mapped onto an $M/G/\infty$ node. In other words, when a part of a street is considered in a dense situation, we need to include the effect of the number of customers on the service rate at the corresponding node. To include the foregoing dependency, we employ a queueing network comprising nodes with state-dependent service times, i.e., $M/G(n)/\infty$ nodes. Detailed discussion on how to derive spatial distribution

in the dense scenario is provided in [63] and [79]. In the dense traffic scenario, the sojourn time of a vehicle is dependent on the number of vehicles at the same street. In the dense scenario, the discussed dependency of vehicle sojourn time can be described by a network of $M/G(n)/\infty$ nodes. Only the service rates are going to change compared to the sparse scenario. The stationary probability distribution, i.e., equations (4.31) and (4.32), are modified as follows [63]:

$$P(n) = b_n \Phi(n) \pi(n) \qquad \pi(n) = \prod_{j=1}^{N} \pi_j(n_j) \tag{4.33}$$

where $P(n)$ is the stationary probability distribution of the modified state dependent network and b_n is the normalizing factor. $\Phi(n)$ is the function that represents the dependency of the service rate on the traffic state (number of the vehicles in the street). To determine $\Phi(n)$, assumptions on laws for overtaking and maneuvers of the vehicles are needed. An example for such assumptions and computation of the state probabilities are provided in [63]. In the dense traffic scenario, the accumulation of vehicles in each segment affects the service rate of the previous nodes. Hence, in such scenario we are dealing with a non-separable queueing network. For our queueing model with multi-class queues and state dependent service rate, Baskett, Chandy, Muntz, Palacios (BCMP) is a suitable queueing model. BCMP queueing model (Appendix B) describes networks with several customer classes, different service strategies (priorities), and different inter-arrival or service time distributions. When the service rate depends on the number of customers in the queue, for First Input First Output (FIFO) service policy there exists a product form solution for the probability of the number of customers [80]:

$$\pi(n_1, n_2, ..., n_N) = \frac{1}{G(K)} d(n) \prod_{i=1}^{N} f_i(n_i) \tag{4.34}$$

where K is the total number of classes of all queues, $n = (n_1, ...n_N)$ is the state vector and N is the total number of queues. $G(K)$ is the normalization constant.

$$K = \sum_{i=1}^{N} C(i) \tag{4.35}$$

where $C(i)$ is the total number of classes in queue i. Let the arrival process be Poisson where all customers (vehicles) are from one source with an overall arrival rate λ, then $d(n)$ which is a function of the number of jobs in the network of queues [80]. $f_i(n_i)$ is a function which depends on the service rate and state of each node [80] [81]. Marginal distributions, i.e., $\pi_i(n_i)$s, can be computed from (4.34). Accordingly, upper bound for probability of connectivity in dense scenario can be derived which will be used in the next section to formulate the path selection scheme.

4.2.3 Application-Centric Routing Protocol

Formulation of Application-Centric Metric

The proposed routing protocol is based on exchanging information between the application layer and the network layer in order to select the path which minimizes the application layer's frame distortion function. The objective function is the average distortion of the video frames at the destination vehicle. Formulation of the objective function is described here and the routing protocol is presented in the next subsection.

In the case of VANETs, the network topology is constantly changing. We extend the solution for multi-hop fixed mesh architecture to the multi-hop vehicular network scenario. We assume that channel estimation process is in place to estimate link loss probabilities. θ, D_0, R_0 and k are parameters for specific video encoder and video sequence. p_σ^h is the loss probability of packet σ over path P_σ^h. The objective function to be minimized is the

summation of all the packet distortions of a video stream as given in (4.9). We extend the formulation for multi-hop mesh scenario to vehicular network by analyzing the three terms which comprise the total end-to-end video distortion, i.e., encoding distortion, distortion due to channel noise and distortion due to delay in VANET, and consider the effect of topology change on each of these three terms in order to formulate the end-to-end optimization problem for vehicular network. D_σ^{enc} is the error related to the encoding and decoding procedure and does not correspond to the link conductivities. Hence, it maintains its original form (i.e., same as fixed multi-hop network scenario). The mobility of nodes in vehicular networks causes uncertainty in existence of connectivity between adjacent nodes in contrast to mesh networks. This behavior can be considered by introducing the probability of connectivity. The lower and upper bounds for the probability of connectivity is given in [63]. We can find the upper bound for the end-to-end distortions due to packet loss and delay. Therefore, the optimization problem should reduce to minimizing the upper bound of end-to-end distortion. The results are nearly optimal and not exactly an optimal solution. The upper bound for D_σ^{loss}, denoted by $D_\sigma^{loss,upper}$, can be derived as follows

$$D_\sigma^{loss,upper} = \frac{k \sum_{h \in P_\sigma} \frac{R_\sigma^h}{R_\sigma} p_\sigma^h}{\prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i)} \tag{4.36}$$

where,

$$p_\sigma^h = 1 - \prod_{ij \in P_\sigma^h} (1 - p_{ij}) \tag{4.37}$$

$\pi(n_i)$ is the stable probability of the event n_i vehicles of class i are in the section of street with length equal to R meters. $\prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i)$ is the lower bound for probability of connectivity derived in [63]. To find the upper bound for $Pr(T_\sigma^h > \Delta_\sigma)$, we apply the *Chernoff bound* and denote it by $Pr^{'}(T_\sigma^h > \Delta_\sigma)$. Using the upper bound for connectivity

probability along the path, $D_{\sigma}^{delay,upper}$ for VANET is derived as follows

$$D_{\sigma}^{delay,upper} = \frac{k \sum_{h \in P_{\sigma}} \frac{R_{\sigma}^{h}}{R_{\sigma}} (1 - p_{\sigma}^{h}) Pr^{'}(T_{\sigma}^{h} > \Delta_{\sigma})}{\prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i)} \tag{4.38}$$

Substituting (4.36) and (4.38) into (4.2), the end-to-end upper bound for distortion of the video packet σ is shown in (4.39). Minimizing (4.39) over all the video packets of a video stream S ($\sigma \in S$) by selecting among the paths gives the nearly optimal solution for minimized end-to-end video distortion over VANETs. Each hop is associated with a rate R_{σ}^{h} and a packet loss probability p_{σ}^{h}. Hence, the optimization problem is the selection of hops which minimize the following distortion function.

$$\begin{aligned} D_{\sigma}^{e,upper} = D_0 + \frac{\theta}{R_{\sigma} - R_0} + \cdots \\ \frac{k \sum_{h \in P_{\sigma}} \frac{R_{\sigma}^{h}}{R_{\sigma}} p_{\sigma}^{h} + k \sum_{h \in P_{\sigma}} \frac{R_{\sigma}^{h}}{R_{\sigma}} (1 - p_{\sigma}^{h}) Pr^{'}(T_{\sigma}^{h} > \Delta_{\sigma})}{\prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i)} \end{aligned} \tag{4.39}$$

The above equation gives an upper bound for video frame distortion over multi-hop transmission in VANETs. The proposed routing protocol minimizes (4.39) at each hop's retransmission in order to increase the quality of the transmitted video frames in terms of MSE at the destination vehicle.

Algorithm 1 Application-centric routing protocol

Notations:
L_{sent} : current location of the packet carrier
I_n: the current intersection p: packet to be forwarded
σ : video packet N_n: the number of outgoing roads at I_n
D_σ^{hop} : hop distortion $P(r)$: the priority of road r to forward packet p
$E[]$: a sorted list of all outgoing roads at I_n,
V_{next} : next hop vehicle for p
C_{next} : cluster of next candidate vehicles
$P(r)$: the priority of road r to forward packet p
$I_{next}(r_{nj})$: neighbor intersection I_j (connected to I_n by r_nj)
Algorithm:
Repeat
$P(E[i])$ is ordered based on the distance to the destination.
if $L = intersection$, then
1. Enter Intersection:
$d_{sent} \leftarrow$ moving direction of the current packet carrier
1.1. $counter = 0$
while $counter < N_n$ and $P(d_{sent}) < P(E[counter])$ do
if there is a loop, then
$P(E[i])$ is ordered based on the direction of nodes
go back to **1.1.**
end if
$S \leftarrow$ all neighbors moving towards road $E[counter]$
$C_{next} \leftarrow$ the closest nodes in S moving toward $I_{next}(E[counter])$
$V_{next} \in C_{next}$ is the vehicle which minimizes the hop distortion D_σ^{hop}
$counter = counter + 1$
apply application based next hop selection in the cluster of neighbor vehicles.
if V_{next} is found then
Break
end if
end while
apply application based next hop selection in the cluster of neighbor vehicles.
if V_{next} is found then
send the packet p to V_{next}
delete the packet from the buffer
end if
else
2. Straight Way
$counter = 0$
while $counter < N_n$ and $P(d_{sent}) < P(E[counter])$ do
$S \leftarrow$ all neighbors moving towards road $E[counter]$
$C_{next} \leftarrow$ the closest nodes in S moving toward $I_{next}(E[counter])$
$V_{next} \in C_{next}$ is the vehicle which minimizes the hop distortion D_σ^{hop}
$counter = counter + 1$
apply application based next hop selection in the cluster of neighbor vehicles.
if V_{next} is found then
send the packet p to V_{next}
delete the packet from the buffer
end if
end while
end if
Until packet reaches the destination vehicle

4.3 Numerical Results

Description of the Routing Protocol

Routing protocols for MANETs are highly dependent on the flooding of the control messages. Geographic routing is the exception in which the location of the destination is known and the packet can be routed by taking local decisions. We focus on geographic routing protocols as they are believed to be more suitable for VANETs. The standard geographic routing protocol normally consists of the following four steps [82]:

1. Determining the destination coordination: this is determined by the source and the coordinates are included in the packet's header to inform intermediate relays.

2. Determining 1-hop neighbors's coordinates: this is done by interchanging the beacon messages which contain the identifier of the sender and its coordinates.

3. Determining the next relay: the next relay is selected based on the cardinality of the sender, 1-hop neighbors and the destination node.

4. Message delivery: the 1-hop neighbors receive the message, but only one of them will contribute to the routing which contains the identifier of the selected node.

We have considered different speed categories along a straight way as well as scenarios where vehicle reaches the intersection. Traffic is modeled by a network of queues in cascade. According to Fig.4.4, each section of the street (front, middle or end part of the street) is a queue with multiple classes of customers (vehicles) corresponding to different speed categories. Hence, a vehicle with higher speed will experience a higher service rate. The source of packet is always an RSU along the road and intermediate vehicle nodes relay the packets to the destination vehicle.

In urban vehicular networks, finding the best path at an intersection is challenging due to possibility of routing loops as well as different paths available at the intersection to choose from. Routing loops which tend to cause high delay on end-to-end packet transmission may happen, which can result in packet dropping and lowering of video quality. Algorithm 1 summarizes the routing protocols which can be described as follows. The following steps happen at each hop of the multi-hop communication to select the best candidate node which minimizes the video frame distortion in terms of MSE.

Location and direction of the node, video packet rate on the link between the carrier node and the neighbor node, packet error probability on the connecting links and total transmission time of the packet are the information exchanged between cooperating neighboring nodes to find the next vehicle (node) for transmission of video packets. On straight way, the carrying vehicle applies the greedy geographic approach to select a cluster of hops which are closest to the destination vehicle. Among the selected vehicles, one is selected for packet forwarding, which transmission over that link results in minimum frame distortion given by (4.39). Upon entering an intersection, the geographical location of vehicles (closeness to the destination vehicle) has a higher priority for selecting the cluster of candidate nodes for forwarding the packets. If a routing loop is detected and the delivery time has not exceeded the transmission delay deadline, the packet forwarding mechanism changes the priority from location to the direction of movement of candidate vehicles in order to prevent recurrence of the routing loop.

First, we present the frame distortion results using the proposed application-centric routing protocol in three different vehicular scenarios. We change the vehicular entrance rate and compute the resultant probability of connectivity. Then, we apply it on the objective function used by application-centric routing as discussed in section 4.2.3.

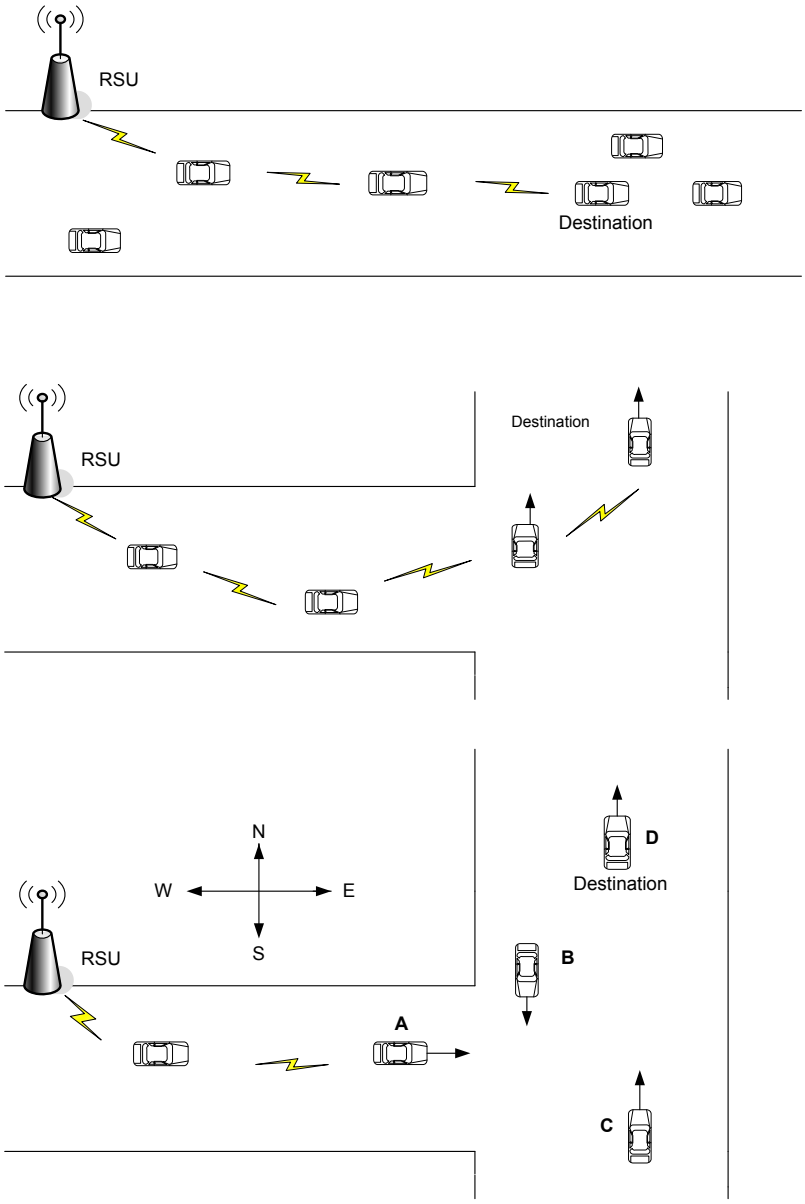


Figure 4.5: (a) Routing over a straight path (b) Routing over an intersection without routing loop (c) Routing over an intersection with routing loop

Scenarios under consideration are shown in Figure 4.5. We consider routing over a straight path, routing over an intersection without occupance of routing loops, and routing over an intersection where a routing loop happens. Figure 4.6 shows the video

distortion of the proposed multi-hop routing protocol in straight way scenario. Increment in number of hops, increases the distortion. In other words, it is possible to deploy fewer RSUs and apply multi-hop communication to cover the areas not covered by RSU radio coverage at the expense of lower frame quality. However, if the traffic density is high due to higher entrance rate of vehicles, the level of distortion can be maintained at low levels (less than 32dB). Video distortion in multi-hop communication at intersection without routing loop is shown in Figure 4.7. The trend is similar to that of straight way scenario, however, the level of distortions is higher and increment in the number of hops causes more distortion compared to straight way scenario. Figure 4.8 shows an intersection scenario with occurrence of routing loop. In this scenario, due to excessive delivery delay more packets are dropped and level of distortions are higher compared to pervious scenarios.

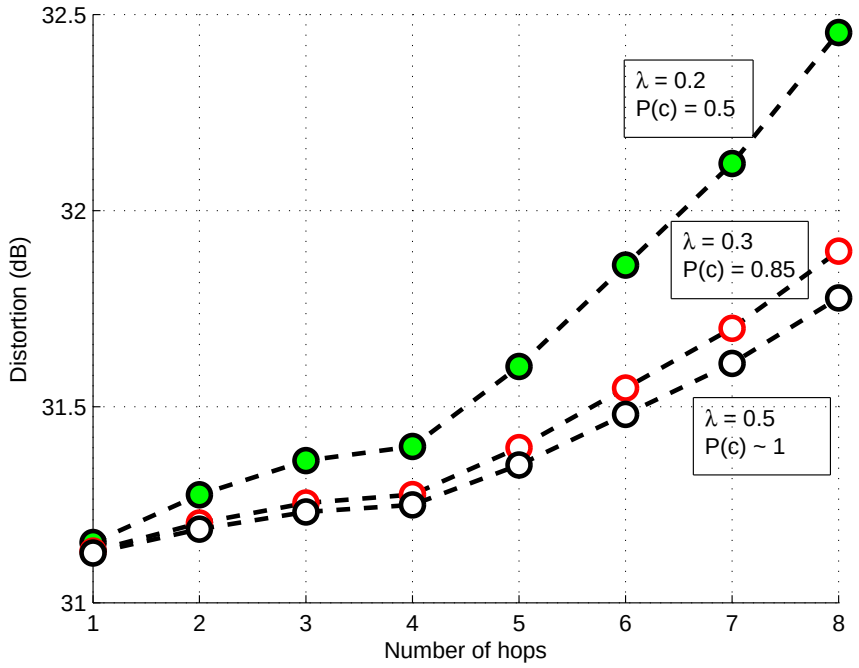


Figure 4.6: Frame distortion over straight way vs. number of hops for different vehicle densities

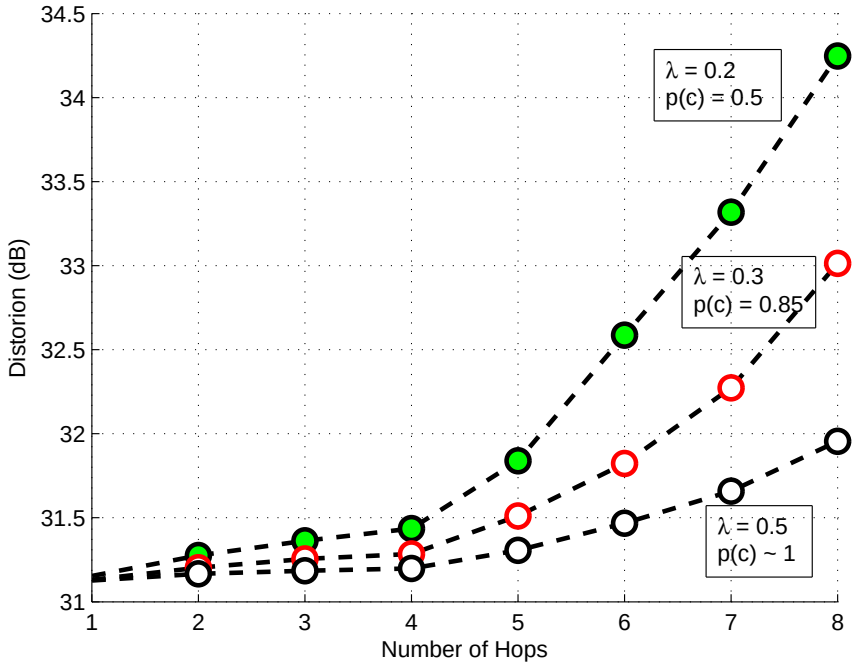


Figure 4.7: Frame distortion over intersection without loop vs. number of hops for different vehicle densities

Connectivity probability versus arrival rate of vehicles to the street is shown in Figure 4.9. As the entrance rate of the vehicles into the street increases, so does the connectivity probability of the vehicles until it reaches the upper bound which is equal to one. To compare the proposed application-centric protocol with greedy forwarding approach we have the following setup and compare performance of the two protocols over a 3-hop connection. Let RSUs be located 1.2Km from each other. Each RSU has an effective radio of coverage of 350m. There is a 500m gap between two consecutive RSUs that multi-hop V2V communication covers this area and is the focus of this comparison. The radio of coverage of each vehicle is approximately 200m.

Frame distortion comparisons are shown in Figure 4.10. It can be seen that there is an average of 2dB gain in our method versus greedy forwarding approach because application-centric protocol always selects the next candidate hop which results in the

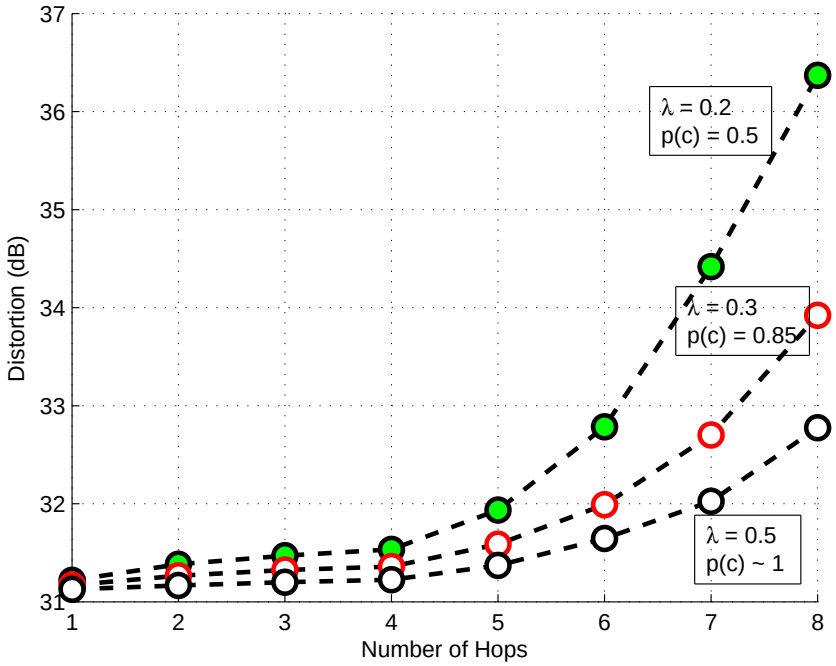


Figure 4.8: Frame distortion over intersection with loop vs. number of hops for different vehicle densities

minimum frame distortion given by (4.9). As the frame rate increases, the distortion decreases for both protocols due to decrease in encoding distortion.

Comparison of start-up delays for application-centric protocol and traditional greedy forwarding scheme is shown in Figure 4.11. It can be seen that for standard data rate of 30 frame/second, there exists a gain in the order of 2 seconds. Increment in data rate results in lower start-up delay for both routing protocols due to higher arrival rate of packets at the destination buffer according to start-up delay analysis in [41].

Compared to the greedy forwarding scheme, the application centric protocol has more computation steps per hop. The objective function in (4.39) is a non-convex function of R_σ^h and p_σ^h. This problem is NP-hard in general and does not appear to possess any special simplifying structure. However, the computation at each hop to select the optimum next hop in the application-centric method is limited to a few candidate nodes

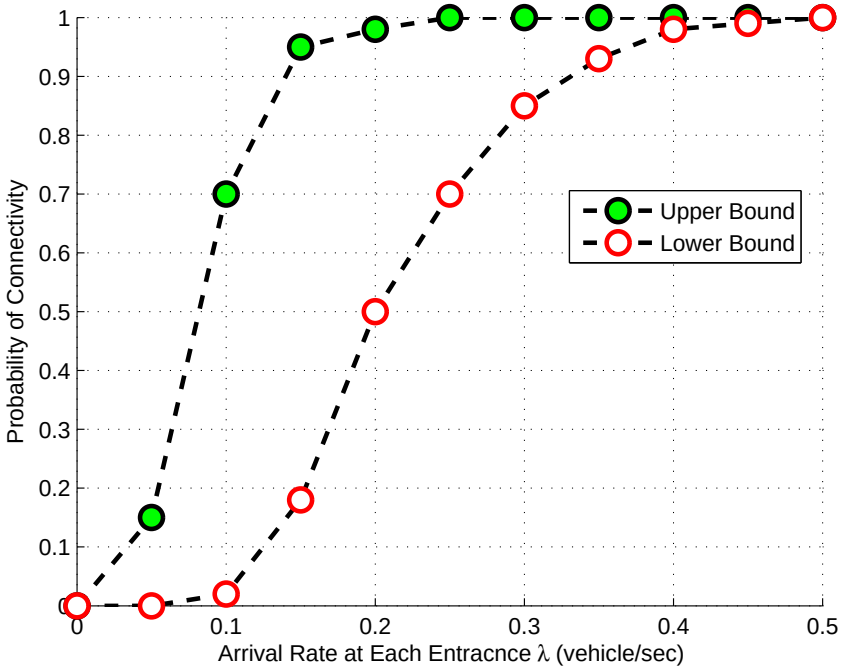


Figure 4.9: Probability of connectivity vs. vehicle entrance rate

by first selecting the cluster of candidates using the greedy geographic approach. In other words, we decrease the size of the search space in order to be able to run the protocol in polynomial time.

We have also measured the performance of the algorithm in terms of PSNR versus number of vehicles in the street and decoding deadline for video packets. The video stream is downloaded from an RSU to the destination vehicle in an urban architecture shown in Figure 4.3 Following the queueing based mobility model described in section 4.2.1, the selected speed for each part of the street as well as the length of the parts are as follows. At the front part of the street, low and medium speed categories are considered. The probability distribution for the low and medium speed categories are $(Unif[0.3,\ 3]\ m/s)$ and $(Unif[3,\ 14]m/s)$, respectively. At the middle part of the street, three speed categories, i.e, low $(Unif[3,\ 14]m/s)$, medium $(Unif[14,\ 22]m/s)$

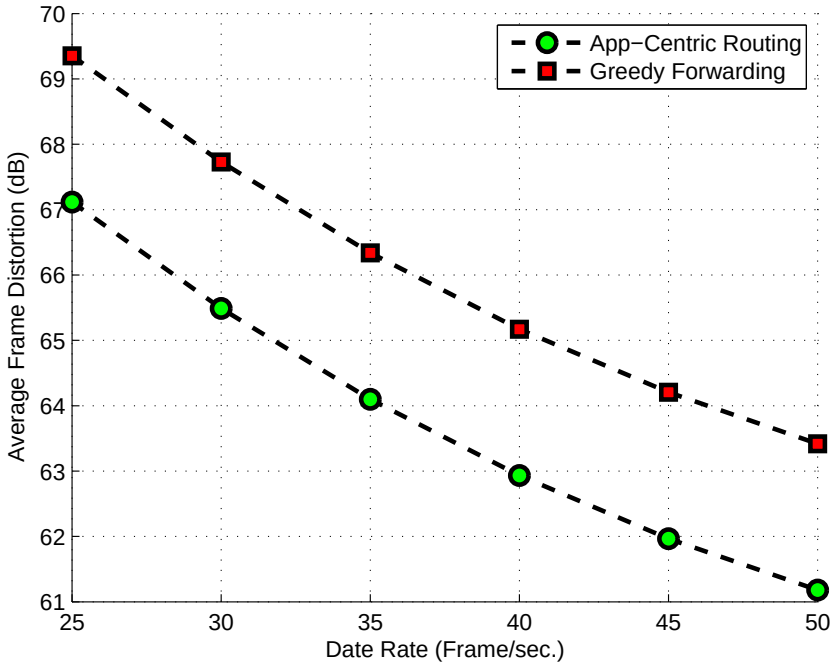


Figure 4.10: Frame distortion vs. data rate (frame/sec.) in comparison with greedy geographic routing

and fast ($Unif[22,\ 33]m/s$) are considered. Finally, at the end part of the street, low ($Unif[0.3,\ 1.5]m/s$) and medium ($Unif[1.5,\ 14]m/s$) speed categories are considered. The length of the front and the end parts are each equal to 200 meters and the length of middle part is 1600 meters. The upper and lower bounds for the probability of connectivity are derived in [63].

To derive the probability distribution for the number of nodes (vehicles) in the street, a state-independent model is suitable for the number of vehicles approximately less than 250. For higher number of nodes (lower arrival rate due to traffic density), simulation results match those of the state-dependent model. Lower and upper bounds for the probability of connectivity are dependent on the entrance rate. The average arrival rate of the vehicles is fixed and same for all the entrances ($\lambda = 0.8(1/sec)$). A 20 seconds

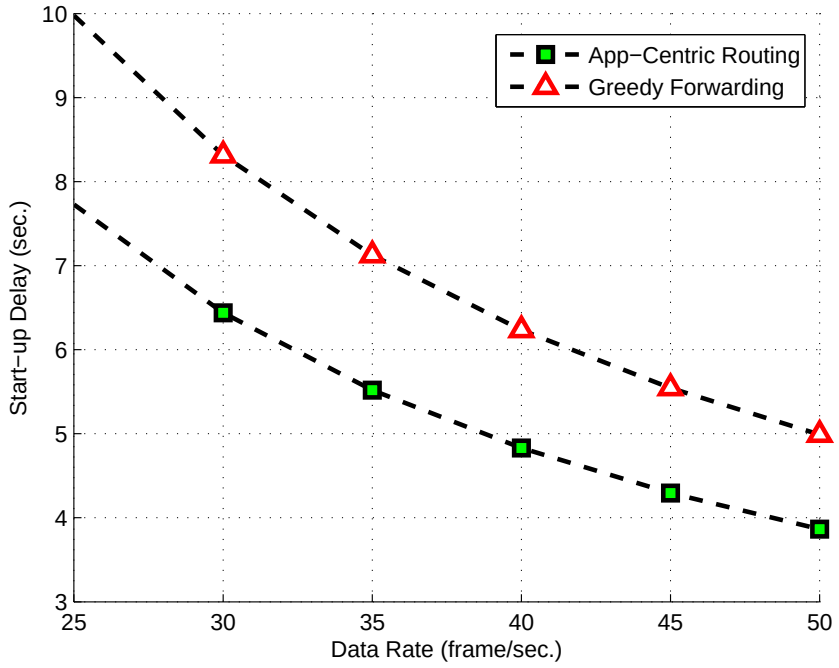


Figure 4.11: Frame distortion vs. data rate (frame/sec.) in comparison with greedy geographic routing

of MPEG encoded video with 30 frames per second rate is considered. Frames are packetized into packets of 1024 bits. The source and the decoder record the departure and arrival time of each packet, respectively. This information is used to reconstruct the video frames. PSNR is calculated per frame and the mean of these PSNR values is reported as the average quality of the video. IEEE 802.11e MAC is adopted and the data rate is set to 6Mbps. The maximum transmission range of the vehicles is taken to be 100 meters which needs an antenna with not more than 20 dBm transmission power.

We are interested in evaluating the performance of the proposed scheme in terms of the average PSNR of the delivered video packets as the total number of vehicles increases from 50 to 500, i.e., starting with sparse traffic scenario and gradually entering the dense traffic scenario. Analytical solution derived from closed form solution in section 4.2.3 is

presented along with simulation results. Figure 4.12 shows the simulation results of the average PSNR versus the total number of vehicles in the street.

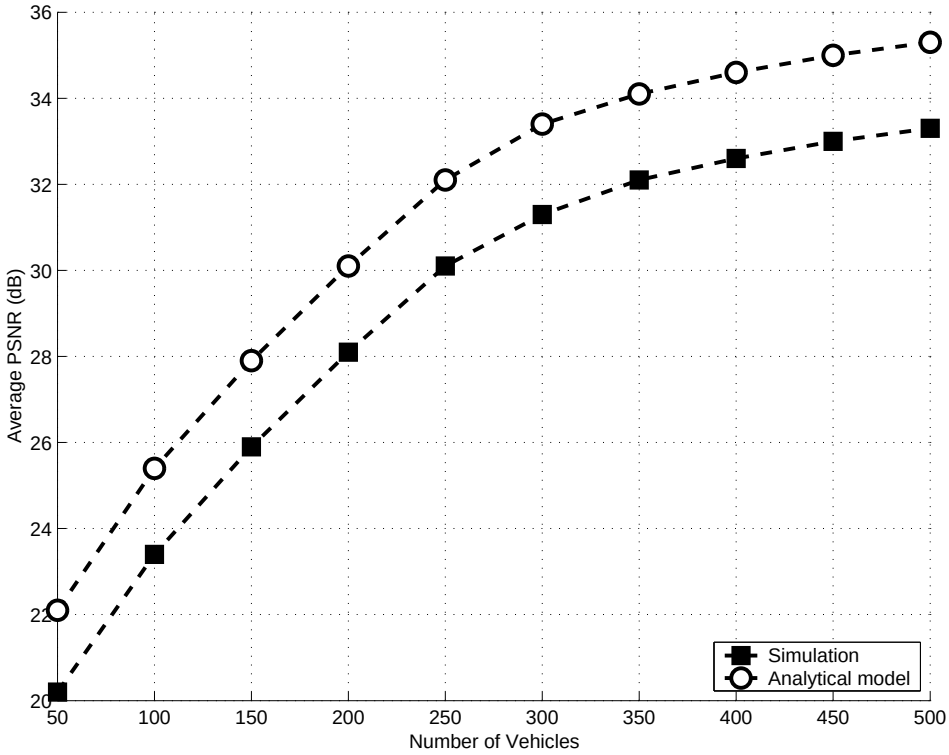


Figure 4.12: Simulation results of average PSNR (dB) vs. the number of the nodes

It can be seen that the average PSNR value of the delivered video packets ascends as the number of nodes (vehicles) on the road increases. The reason is that when the total number of vehicles increases, the probability of connectivity between vehicles increases which consequently results in decrement of the distortion due to the packet loss (the second term in equation (4.39)) in the channel. It is observed that the simulation results exhibit the same trend as the analytical results, however, there is approximately 5% to 10% difference between the simulation and the analytical solution. Considering the complexity of analytical formulation for streaming in VANET, we conclude that the applied queueing network model can predict the performance of the designed scheme

with approximately 5% precision in the highly dense areas.

Analyzing the distortion due to congestion and the distortion due to delay in the end-to-end distortion formula (4.39), we observe that the impact of the delay distortion is weaker due to multiplication by the $Pr(T_\sigma^h \geq \Delta_\sigma)$. The channel loss probability is assumed to be estimated and available at the RSU and obviously is not in the control of the transmitter or receiver. The parameter that can be determined is the decoding deadline. If the decoding deadline is taken to be large, it causes the $Pr(T_\sigma^h \geq \Delta_\sigma)$ to approach 0. This will eliminate the distortion due to the delay term and will increase the end-to-end PSNR. We analyze the impact of the delay deadline in the very dense (600 vehicles) scenario where the major distortion is caused by delay deadline rather than packet loss in the channel. Figure 4.13 shows that the distortion due to delay decreases as the decoding deadline Δ_σ increases. It also shows that distortion due to delay has a maximum impact of 2 dB on the video quality which is lower compared to the maximum distortion due to the channel loss. The simulation result closely matches the analytical results. The analytical solution and simulations are closer for high deadline values, because in such scenarios, the distortion due to delay is very small. Error bars for frame distortion over different scenarios are shown in Figures 4.14, 4.15 and 4.16. Error margins are shown for different connectivity probabilities versus number of hops between source and destination. The error margins increases with the number of hops in all three scenarios (straight line, intersection without routing loop, intersection with loop) because multi-path fading and shadowing, signal interferences among nodes show more dynamic when the number of hops increases.

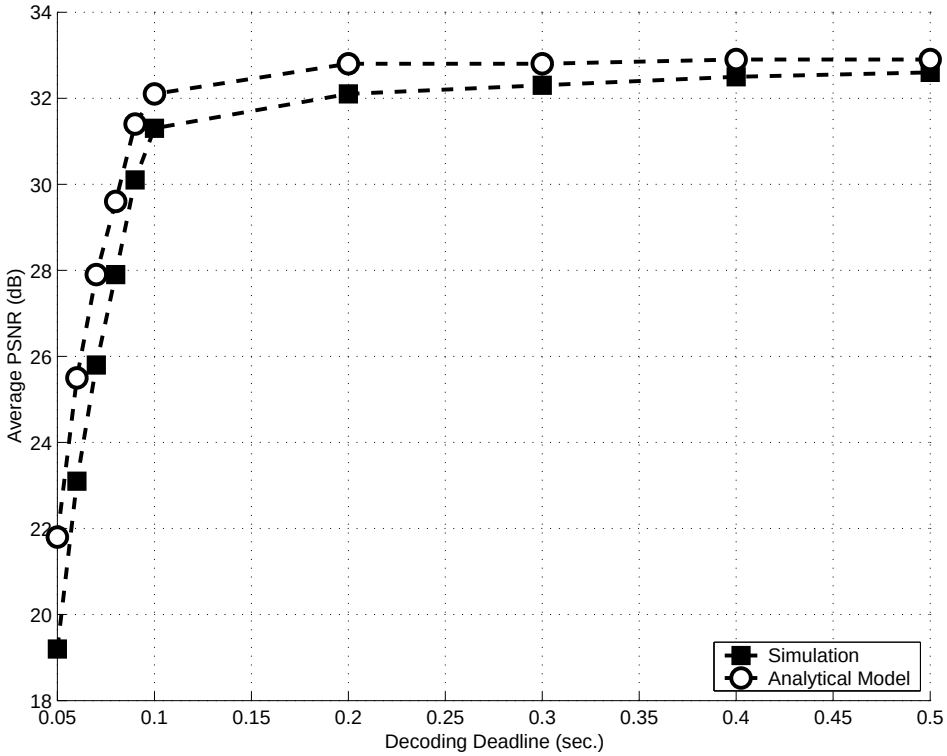


Figure 4.13: Simulation results of average distortion (dB) vs. decoding deadline (sec.)

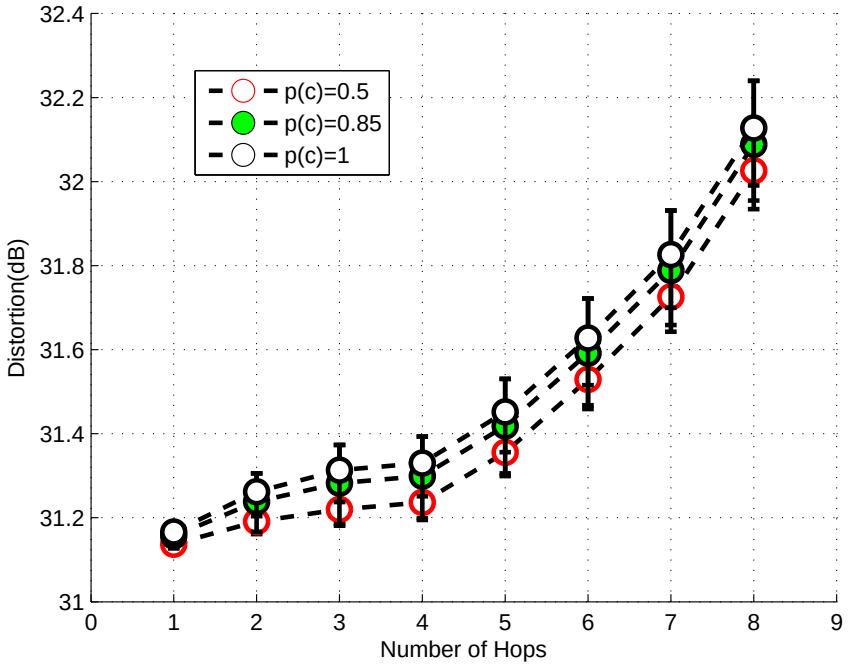


Figure 4.14: Error margins of frame distortion over straight path

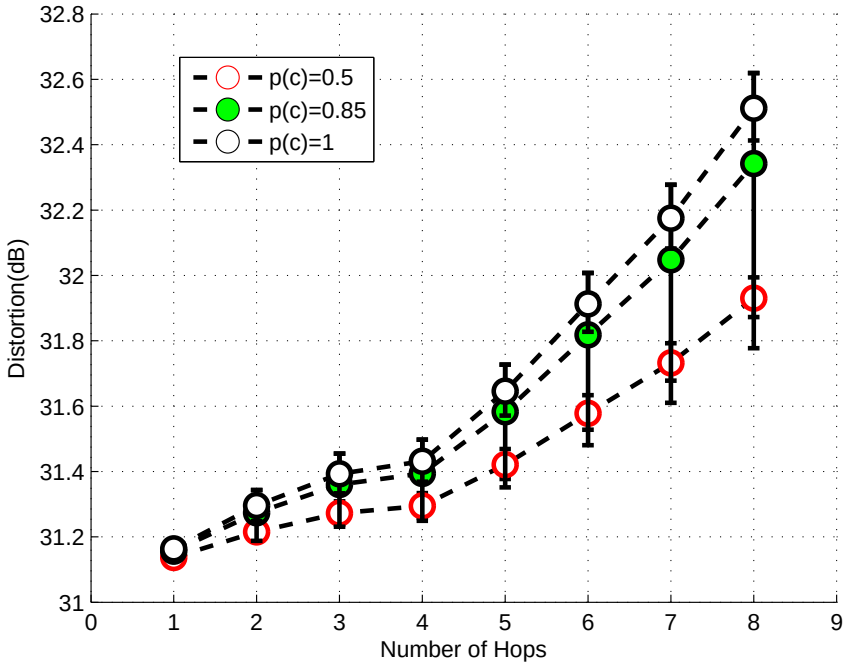


Figure 4.15: Error margins of frame distortion over an intersection

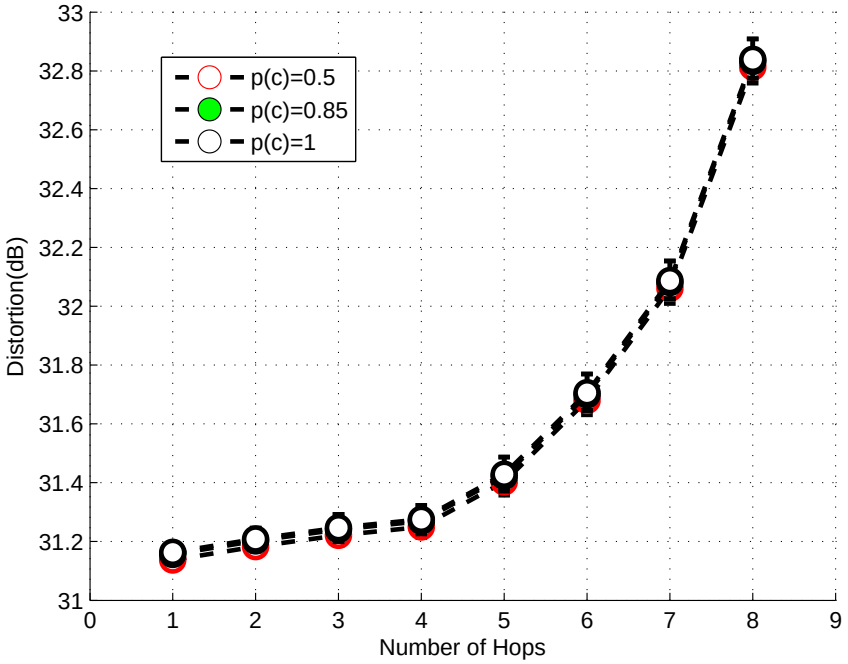


Figure 4.16: Error margins of frame distortion over intersection with loop

4.4 Summary

In this chapter, a novel application-centric multi-hop routing protocol for video transmission in urban VANET scenarios has been proposed. First, we presented the optimal path selection problem for video streaming over multi-hop mesh network as a classical classification problem. Second, we take the self-similar nature of the input traffic into consideration and formulate our problem based on a FBM traffic model, which is an appropriate model for state of the art H.264 codecs. Comparison of our results against other state of the art solutions gives very similar range of results while the complexity of our method is lower compared to complex optimization solutions that other methods propose which must be done real time. The complex part of our method, i.e., feature extraction and classifier training, is done off-line, which has no impact on real time complexity and hence reduces end-to-end delay. To extend the work to a vehicular network scenario, we have translated the design of routing protocol into an optimization problem using a probabilistic model to formulate average end-to-end distortion of the delivered video frames over the entire video session with consideration of the connectivity probability of vehicle to the RSU along the roads in an urban area. Performance of the protocol in different vehicular scenarios versus increment in coverage range through more communication hops has been evaluated. In addition, numerical results have been given to demonstrate the proposed application-centric routing protocol achieves lower start-up delay and frame distortion compared to the traditional greedy forwarding protocol. Decrement in the size of search space at each hop has made it possible to run the protocol in polynomial time.

Chapter 5

Seamless Quality-Driven Multi-Hop Routing in Urban VANET Scenarios

Video streaming is a delay sensitive application. Therefore, in order to have a smooth playout, it is necessary to have enough packets in the playback buffer at the destination [41]. In addition, the perceptual quality of the reconstructed frames should also be taken into consideration. To address the aforementioned concerns, we propose a new quality-driven routing scheme for delivering video streams from a RSU to a destination vehicle. The routing scheme also considers the use of multi-hop paths, given that V2V communications are highly feasible in urban scenarios. In this way, the packet delivery is supported by a routing mechanism that can guarantee a smooth video streaming with high visual quality. On the other hand, given that the destination vehicle is in motion, it is possible that it moves to a service area controlled by a different AR. Therefore, to address the problem of handovers experienced by the vehicle, a network mobility management scheme is proposed in [1], which works in conjunction with the quality-driven geo-routing protocol, for a seamless delivery of video packets. The applied mobility management

scheme is based on Proxy Mobile IPv6 (PMIPv6) [83], and is combined with the geo-networking features present in VANET.

Although PMIPv6 has a good acceptance in vehicular scenarios, it has an important restriction for its deployment in VANET: it requires a direct connection between the mobile node and the AR (also known as mobile access gateway (MAG)). Therefore, [1] proposes a solution in which the multi-hop paths are considered by adapting the geo-routing protocol proposed in this chapter. Specifically, the following challenges are addressed in [1] to manage mobility: 1) the definition of a mechanism to detect new connections in the domain; 2) the delivery of IP-related signalling through the multi-hop path; and 3) the prediction of handover events. We emphasize that only the geo-routing protocol directly comes from the content of this chapter and the discussion on the mobility management scheme, which is based on collaborative work in [1], is only provided to show the application of the proposed routing scheme in drive-thru scenarios where IP management is necessary in order to have continues access to the RSUs.

The remainder of this chapter is organized as follows: In section 5.1, we describe our system model. Section 5.2 describes the proposed quality-driven geo-routing scheme. In section 5.3, we introduce the multi-hop PMIPv6 management scheme. Numerical results are provided in section 5.4. Section 5.5 describes the concluding remarks.

5.1 System model

We assume a vehicular system based on Dedicated Short Rage Communications, in which each vehicle is equipped with OBUs and broadcasts the location, direction, speed, acceleration and traffic events to its neighbors [45]. The network topology is shown in Figure 5.1. It is considered that RSUs are spaced by 1.2Km from each other, with an effective radio of coverage of 350m each. Thus, there is a 500m-area between two consecutive RSUs

that relies on multi-hop links for Vehicular to Infrastructure (V2I) communications. The radio of coverage of each vehicle is approximately 200m.

The video is streamed from an AR to the proper RSU, and then from the RSU to the destination vehicle. While the destination vehicle is in the transmission range of the RSU, they connect directly in a one-hop fashion. Once the vehicle exits the coverage range of the RSU, video packets are transmitted to the vehicle using multi-hop paths, in which the intermediate vehicles serve as relays. According to the considered values for RSU coverage, vehicle transmission range, and distance between RSUs, there would be at most a 3-hop connection between the RSU and the destination vehicle for video streaming. When the destination vehicle gets closer to the next RSU compared with its distance to the previous RSU, the AR switches the video streaming to the next RSU. Hence, the transmission would follow a repetitive sequence of {1-hop, 2-hop, 3-hop, 2-hop}.

The service area of each AR contains several RSUs and is defined by the network operator [84]. For IP configuration, the AR sends Router Advertisement (RA) messages towards certain geographical area based on the information received from the vehicles. Hence, the AR can advertise its services via geocast messages.

In this way, vehicles in the infrastructure-connected VANET may learn the exact position of an AR and directly request the assignment of an IP prefix by adaptation of PMIPv6 in the VANET multi-hop domain using the proposed geo-routing protocol.

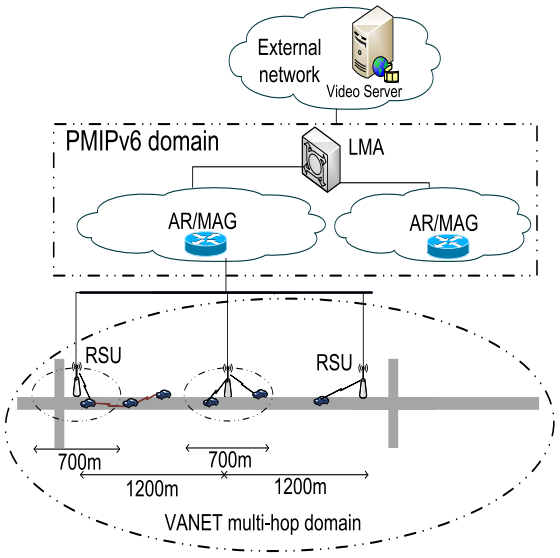


Figure 5.1: Network topology

5.2 Quality-Driven Routing Protocol

The proposed routing protocol is designed based on the optimization of video quality in terms of improving the visual quality of the transmitted video frames, as well as the smoothness of the streaming. In the following subsections, we first describe the video quality metrics, and the last subsection introduces the proposed quality-driven routing protocol.

5.2.1 Startup Delay and Frequency of Streaming Freezes

We consider an infinite buffer capacity at vehicles, which is a reasonable assumption given the high storage capability that can be deployed in vehicles. The video playback process can be divided into two phases: 1) charging phase and 2) playback phase. The charging phase starts once the buffer becomes empty. Thus, the playback is kept frozen until the buffer is filled with b packets (i.e., b is a threshold of the playback). The charging phase

of the destination's vehicle's buffer or start-up delay , D_s, is given by:

$$D_s = min\{t|X(0) = 0,\ X(t) = b,\ t > 0\} \tag{5.1}$$

where $X(t)$ is the number of packets in the buffer at time t. Due to dynamic packet arrivals and departures during the playback phase, the playback phase may stall when the buffer becomes empty. Denote the playback phase by a random variable T. Charging and playback phases iterate until the whole video is played. In this work, we analyze the start-up delay in the proposed video streaming framework over an urban VANET. A larger playback threshold will result in larger start-up (charging) delay. We keep the playback threshold fixed and, instead, analyze the start-up delay according to the dynamics of the vehicular density (i.e., sparse and dense scenarios).

To derive an analytical formulation for start-up delay in video streaming at the destination vehicle, the playout buffer can be modeled as a $G/G/1/\infty$ queue that follows the diffusion approximation method presented in [41]. By applying the diffusion approximation, the transient solution of the queue length can be exploited by obtaining its p.d.f. at any time instant t.

$$E(D_s) = \frac{b}{\lambda} \tag{5.2}$$

where b is the playback threshold and λ is the arrival rate of the packets at the destination vehicle. The playback terminates when the buffer becomes empty again. According to [41], the average number of streaming freezes after t seconds, which can be approximated using diffusion approximation as follows:

$$E(F) \approx -\frac{\lambda(\lambda - \mu)}{\mu b}t \tag{5.3}$$

Both (5.2) and (5.3) indicate that as the arrival rate of video packets at the destination

vehicle increases, the performance of streaming improves in terms of startup delay and frequency of streaming freezes. Hence, the goal of routing at the link layer must focus on minimizing the transmission delay, in order to improve the streaming performance in terms of the above mentioned metrics. However, these metrics do not consider the visual quality of the video frames because they are not concerned in delivering the specific high priority video packets (e.g., packets that construct I-frames). As long as enough packets are in the buffer, these metrics may show a high performance. Hence, in addition to smoothness of streaming, we also consider the video frame qualities in terms of the PSNR, and design the inter-vehicle routing protocol to be capable of improving both frame quality (PSNR) and streaming quality (startup delay and streaming freezes).

5.2.2 PSNR of Delivered Video Frames

A multi-hop network can be modeled as a graph with N vertices (nodes) and L edges (links). The intermediate hops are in fact mobile relays which use decode and forward scheme in order to prevent amplification of the noise along hop to hop transmission. Let S denote a set of video sessions to be transmitted. Each video packet $\sigma \in S$ has video source z_σ and destination d_σ. The rate of the video stream for packet σ is bounded by $\underline{R}_\sigma \leq R_\sigma \leq \overline{R}_\sigma$, $\sigma \in S$, while the upper and lower bounds are determined by the encoder used at the source node. The end-to-end video frame distortion is given in (5.4). Details of the derivation of end-to-end video frame distortion over a VANET scenario similar to this work can be found in [64].

$$D_\sigma^e = \underbrace{D_0 + \frac{\theta}{R_\sigma - R_0}}_{\text{encoding distortion}} + D_{loss} \tag{5.4}$$

where θ, D_0 and R_0 are parameters for the specific video encoder and video sequence. D_{loss} is the mean square error due to channel noise or pack drop due to exceeding the

transmission delay deadline. The total distortion is the summation of all the frame distortions of a video stream.

5.2.3 Impact of Mobility

One of the critical issues consists of the design of scalable routing algorithms that are robust to frequent path disruptions caused by vehicles mobility [36]. When a path breaks, not only portions of data packets are lost, but also in many cases, there is a significant delay in establishing a new path. This delay depends on whether another valid path already exists (in the case of multipath routing protocols) or whether a new route-discovery process needs to take place. By clustering, since control messages are only forwarded within the same group, the scheme prevents flooding of control packets throughout the entire network. Hence, the achieved throughput of the network will be more evident than in the case of traditional algorithms that do not take into account mobility, as will be demonstrated later in the numerical results. The goal is to select stable and more durable paths, so that there will be fewer path breaks and handoffs.

While it is all but impossible to come up with a routing approach that can be suitable for all VANET applications and can efficiently handle all their inherent characteristics, attempts have been made to develop some routing protocols specifically designed for particular applications. For the proposed clustering routing protocol in [36], the key idea behind the scheme is to group vehicles according to their velocity headings. If the two vehicles belong to two different groups, the link between the two vehicles is judged to be unstable.

We consider a scenario of high practical relevance, in which the network infrastructure is represented by a wireless mesh network and the vehicular node is a mesh node itself. Mesh networks are typically freestanding, robust systems that can be conveniently inte-

grated with the existing infrastructure and offer high bit-rate services. We specifically address the case in which the vehicular node is a mesh router. The advantage of our solution is twofold: (i) it allows the routing protocol to be run on the mobile node itself, thus better adapting to the high-mobility profile of the node; (ii) the OBU on the vehicle can act as a gateway towards the mesh infrastructure for all client devices in the vehicle.

5.2.4 Inter-Vehicle Routing Protocol

The vehicular network can be represented as a directed graph. The intersections are graph nodes and roads are edges of graph [45]. Only edges with vehicles on it can be selected for packet forwarding. The proposed data delivery model has two modes: 1) straight way and 2) intersection. On the straight way, the packet carrying vehicle selects N_D neighbor vehicles that are geographically closer to the destination vehicle and are in transmission range of the carrying vehicle. Then, it selects the next candidate hop which minimizes the frame distortion with minimum channel loss probability. At the intersection, the vehicle must select the next straight path in order to forward the packet. The delivery delay can be estimated using the stochastic model proposed in [45] by solving the following set of equations:

$$D_{mn} = d_{mn} + \sum_{j \in N(n)} (P_{nj} \times D_{nj}) \tag{5.5}$$

D_{ij} is the expected packet delivery delay from intersection I_i to destination, if the packet is forwarded trough road r_{ij}. P_{ij} is the probability that a packet is forwarded through road r_{ij} at I_i. $N(j)$ is the set of neighbor intersections of I_j. If any routing loop happened at the intersection, the packet will be dropped since video streaming is a delay sensitive application and can not tolerate large delivery delays. In fact, our scheme maximizes the arrival rate of packets at the destination because the end-to-end packet loss probability

will be minimized according to $PL = 1 - \prod_{ij}(1 - p_{ij})$ where p_{ij} is the packet loss probability of link between node i and j. The effective arrival rate at destination will increase according to $\lambda^* = \lambda(1 - PL)$. Hence, the start-up delay and average number of streaming freezes will decrease. The steps of the data routing scheme are given in Table 5.1.

Table 5.1: Inter-vehicle routing protocol

On Straight Way

If any vehicle in the transmission range and closer to the destination,

then

- Apply greedy geographic routing and select N_D candidate nodes, $\{c_i | i = 1..N_D\}$, such that $|D(c_i) - D(c_{i+j})| \leq \delta$ for $j = 1$ to $N_D - 1$.
- Select a node among candidates which gives minimum packet distortion.

else drop the packet.

end

At the Intersection

If any vehicle **NOT** in the transmission range, **then**

- Drop the packets.

else (vehicles in transmission range at intersection)

- Solve the set of equations in (5.5).
- Select road r_{ij} which results in minimum delay.
- Apply greedy geographic routing, select N_D candidate nodes, $\{c_i | i = 1..N_D\}$, such that $|D(c_i) - D(c_{i+j})| \leq \delta$ for $j = 1$ to $N_D - 1$.
- Select a node among candidates which gives minimum packet distortion.
- if routing loop happened at intersection, then drop the packet.

end

5.3 IP Mobility for Seamless Quality-Driven Data Delivery

In section 5.2, we select the multi-hop path that offers the minimum end-to-end distortion for the delivery of video packets. However, it is unrealistic to assume that the destination vehicle will remain connected to the same RSU while it is moving along the city. Taking into account the network topology shown in Fig. 5.1, if a change of RSU occurs, it may also carry out a change of the AR (i.e., the vehicle enters to a service area under the control of a different AR). Therefore, to show the complete scenario where change of AR happens, an IP mobility management scheme is proposed in [1] using the proposed geo-routing protocol that allows for a seamless video streaming service when an active connection may be affected by a change of IP addresses at the destination vehicle.

The proposed IP management scheme in [1] is a novel adaptation of PMIPv6 for a multi-hop VANET with a handover prediction mechanism. It works in conjunction with the geographical routing algorithm described in Table 5.1, and relies on the IPv6 support for VANET using geo-networking features [85]. The geo-routing layer is in charge of the forwarding of IP packets in the multi-hop path, so that it creates a virtual point-to-point link between the destination vehicle and the AR, with no need of processing IP headers at any intermediate vehicle, nor at the RSU.

5.3.1 IP configuration in the PMIPv6 domain

An IP address is assigned to the vehicle by the network domain when it enters to the vehicular network in order to deliver video packets from the video server to the vehicle. PMIPv6 is defined for the case when a mobile node connects directly to an AR. Therefore, the challenge is to adapt the original PMIPv6 to multi-hop scenarios. Tables 5.2 and

5.3 in [1] describe the steps to request and obtain a valid IP prefix using our proposed geo-routing scheme to extend the original PMIPv6 to multi-hop scenarios, respectively. Both processes make use of the geo-routing layer at each vehicle for the virtual point-to-point link. They also rely on standard control messages defined in Neighbor Discovery (ND) for IPv6 [86] for finding routers in the domain, and for keeping track of reachability between IP neighbors. The way in which those messages are processed is modified in order to take advantage of the geo-networking features of the VANET.

Once the IP configuration at the destination vehicle has been completed, the regular IPv6 control information related to ND is exchanged between the IP peers, i.e., between the vehicle and the AR. In order for this exchange of information to happen, the tunnel established at the geo-routing layer is always used for the multi-hop delivery of packets between the two peers. For the proper operation of PMIPv6 in the multi-hop scheme, it is necessary to maintain state information at the entities involved in the exchange of IP packets as follows [1]:

- Neighbors table: This table stores the information of nodes in VANET from which a link layer beacon has been received. Based on this information, every node is able to execute the geo-routing protocol and locate the best candidate to forward packets to the destination.

- IP Neighbors Table: Table in which the IP neighbors are stored. It corresponds to nodes from which messages that belong to the ND protocol have been received. At the vehicle, this table contains the information about the AR that is serving its current area. It also stores the information about the candidate ARs available in the domain (the list is extracted from the RA messages). At the AR, this table contains all the nodes in the connected VANET for which the AR is serving as

the IP next hop. This table could be integrated to that of the ND protocol by appending the additional fields for location information.

5.3.2 Geo-Routing Protocol for Handover Predictions

Location information of VANETs can be directed to AR using the proposed geo-routing protocol. This location information can utilized in order to predict occupance of a handover. Hence, a pre-configuration of the IP setting can be performed at the AR/MAG before the vehicle moves to a new service area. Every node that has a valid IP prefix from the PMIPv6 domain, exchanges ND messages with the AR. Therefore, based on the information received from NeighborDiscovery/NeighborAdvertisement messages, the AR is able to update the location of a vehicle. and predicts if it is going to leave the current AR service area and report this to LMA. The LMA then chooses the next candidate AR based on the position and direction of the vehicle, and establishes an inactive tunnel with it. Once the handover happens, this pre-established tunnel with the LMA will be activated. Once the LMA receives the notification for the activation of the tunnel, it redirects the forwarding of IP packets to the new vehicle's location, and the handover process is terminated. The handover prediction mechanism will reduce the handover delay and results in a better video streaming performance [1].

Table 5.2: IP request for multi-hop PMIPv6 [1] which applies our proposed geo-routing protocol to establish a connection between AR and the mobile nodes.

At destination vehicle

1. Complete layer 2 connection to an intermediate vehicle.

2. IP layer processing:

- Generate a Router Solicitation (RS) message (all-routers multicast address as the IP destination).
- Pass RS packet to the geo-routing layer.

3. Geo-routing layer processing:

- Create a geo-header where the all-routers multicast address is translated to a geo-cast address.
- Set the flag `RouterRequired` in the geo-header.
- Forward packet using the VAAD protocol.

At intermediate vehicle or RSU

1. Geo-routing layer processing:

- **if** `RouterRequired` is set and AR location exists in location table, **then** change geo-header destination to the geo-unicast address of AR.
- Forward packet using the VAAD protocol.

At the AR/MAG

1. Geo-routing processing:

- Store location of destination vehicle.
- Pass RS packet to the IP layer.

2. IP layer processing:

- Start PMIPv6 signalling to the LMA.

At the LMA

1. If MN already exists in the domain, **then**

- Update MN location.
- Keep the same network prefix assignment.

2. else

- Create binding cache for the MN.

3. Create list with candidate ARs, locations and service areas.

4. Send PBA to MAG with home network prefix and candidate ARs.

Table 5.3: IP advertisement for multi-hop PMIPv6 proposed in [1] which applies our proposed geo-routing protocol to establish a connection between AR and the mobile nodes.

At the AR/MAG

1. IP layer processing:

- Create RA message with prefix and list of candidate AR received from the LMA.
- Pass RA packet to the geo-routing layer.

2. Geo-routing layer processing:

- Set the destination vehicle geo-id as tunnel end-point for packets directed to the vehicle's network prefix.
- Create geo-header with the geo-unicast address of the vehicle.
- Set the flag `AccessRouter` in the geo-header.
- Forward packet to the RSU.

At intermediate vehicle or RSU

1. Forward packet using the VAAD protocol.

At destination vehicle

2. Geo-routing layer processing:

- Store AR location in location table
- Set AR's geo-id as tunnel end-point for packets directed to external IP addresses (IP next-hop)
- Pass RA packet to IP layer

2. IP layer processing:

- Process RA message (perform stateless address configuration and set default router)

5.4 Performance Evaluation

To evaluate the performance of the quality-driven routing, we make a comparison between the proposed quality-driven routing and traditional greedy geo-routing scheme in terms of quality of streaming (start-up delay and number of freezes) as well as visual quality (PSNR of delivered video frames). Fig. 5.2 shows the performance of the data routing scheme in terms of start-up delay of video streaming for two scenarios: 1) streaming starts when vehicle maintains 2-hop connection with RSU, and 2) Vehicle has 3-hop connection with RSU. The results are also compared with greedy geographic routing scheme. Increment in data sending rate in each scenario results in lower start-up delay for both routing schemes because higher data rate increases the arrival rate of video packets at the destination buffer, i.e, buffer will be filled faster and reaches the streaming threshold in a shorter time. Comparison of 2-hop and 3-hop scenarios for our quality scheme shows that in 2-hop scenario we have slightly shorter start-up delay compared to 3-hop, however the difference is very small which shows high stability for our scheme. The difference is due to two reasons: 1) faster delivery of packets due to fewer number of hops and 2) lower packet loss (higher arrival rate at the destination) due to fewer number of hops. Finally, the reason for lower performance of our method compared to greedy scheme is selection of path with lowest loss probability in our method which results in higher arrival rate at the destination although our method uses the greedy approach in its core.

Average number of streaming freezes is selected as another metric to evaluate the performance of the quality-driven routing approach in terms of smoothness of video streaming. Fig. 5.3 shows the number of streaming freezes versus data sending rate for 2-hop and 3-hop connection scenarios. The results are also compared with greedy

geographic routing performance for the session length of 300 seconds. The results show that for low data rates the greedy algorithm achieves a slightly lower frequency of freezes compared to our quality-driven scheme. However, as the data rate increases our method results in significantly smaller number of freezes compared to greedy algorithm. Our method guarantees higher arrival rate by selecting links with lower loss probability. Lower packet loss is the main attribute of our routing scheme compared to greedy scheme and its impact on the arrival rate of packets at the destination increases as the input data rate is increased (above 30 frames/sec.). According to (5.3) for higher values of arrival rate the number of freezes decreases because the second derivative of (5.3) is negative.

In Fig. 5.4, the distortion of delivered video frames using our routing scheme is compared against greedy scheme. The quality-driven algorithm always selects the next candidate hop which results in minimum frame distortion given by (5.4). As it was expected, our method results in lower frame distortion compared to greedy approach.

Video sessions lengths are from 30s to 1800s. ρ_s is defined as λ/μ_s, where the subnet crossing rate μ_s is calculated as $vL_s/\pi A_s$. The average velocity v is 50Km/h, L_s and A_s are the perimeter and area of the router's service area. Total cost is $C_{BU} + C_{PD}$, where C_{BU} is the signalling cost (bytes) transmitted per hop $C_{BU} = \sum_i i \times BU \times \alpha(i)$, where BU is customized for each scheme according to how many hops the signalling messages have to cross to reach the anchor point. C_{PD} is total overhead to deliver packets in one video session. Video frame size is $7.7^{\exp 02} bytes$, the transmission delay between vehicles is assumed to be 5ms, and from the vehicle to RSU is 10ms. Figure 5.5 shows the performance in the compared schemes in terms of handover delay. The multi-hop PMIPv6 with the prediction mechanism allows for resuming the reception of video packets approximately $1.6 \sim 2.1$ times faster than MANET-centric NEMO, thanks to the pre-established tunnel at the new MAG using the proposed geo-routing protocol.

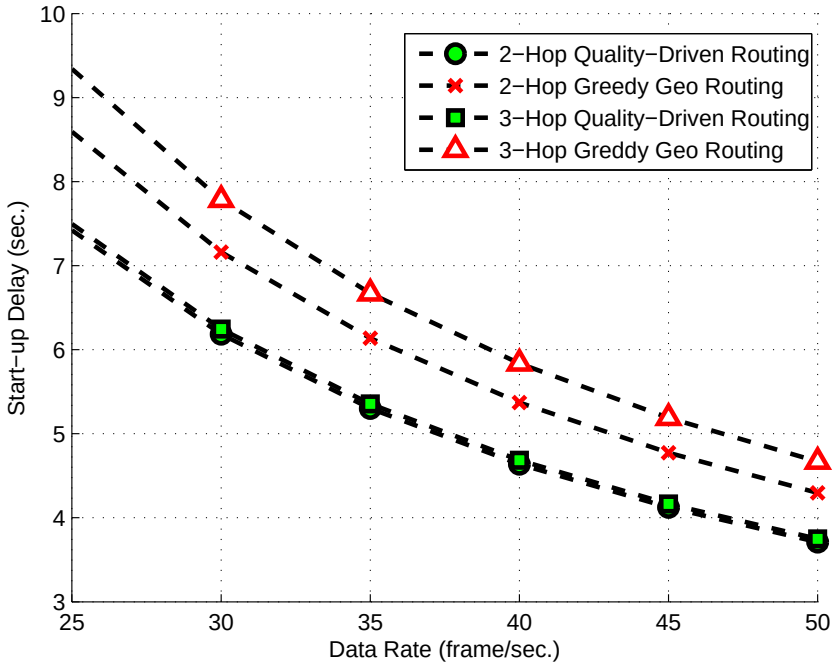


Figure 5.2: Start-up delay vs. data rate (frame/sec.)

Finally, Figure 5.6 shows the impact of the vehicle's velocity in the average of lost frames during a video session, when our integrated scheme is applied for the routing and handover management. We consider a session length of 900s, and three different average velocities for the vehicle. The results show that the behavior of the integrated scheme is fairly stable under an increase in velocity. Fig. 5.6-2 shows more detail of the behavior for high data rates- high velocities scenarios. The increase in losses are due mainly to the increase in the number of handovers during the session length.

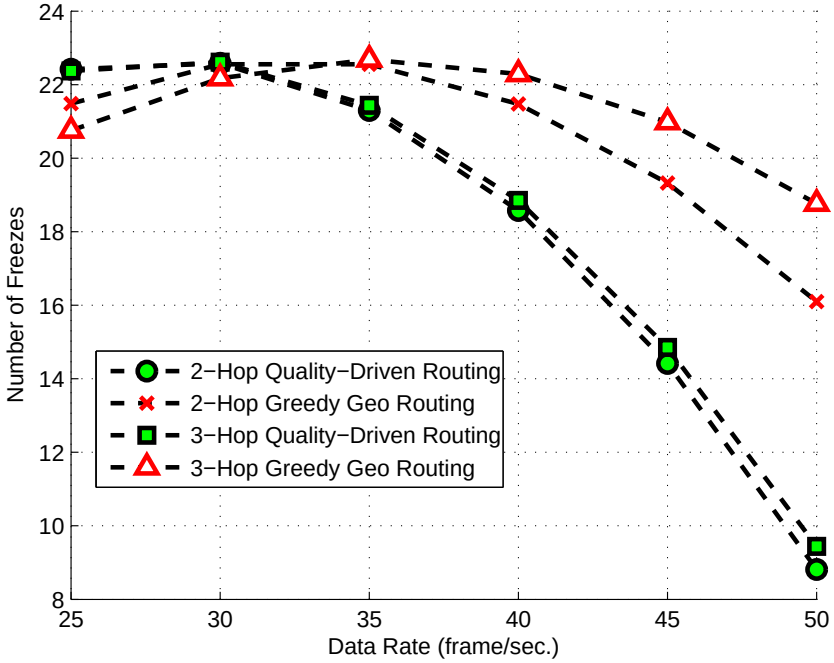


Figure 5.3: Number of freezes vs. data rate (frame/sec.)

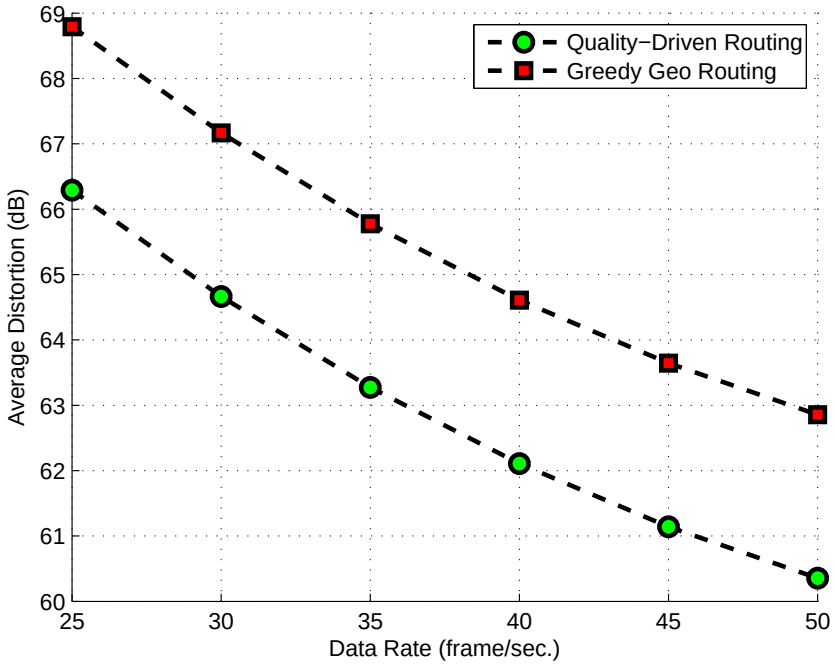


Figure 5.4: Frame distortion vs. data rate (frame/sec.)

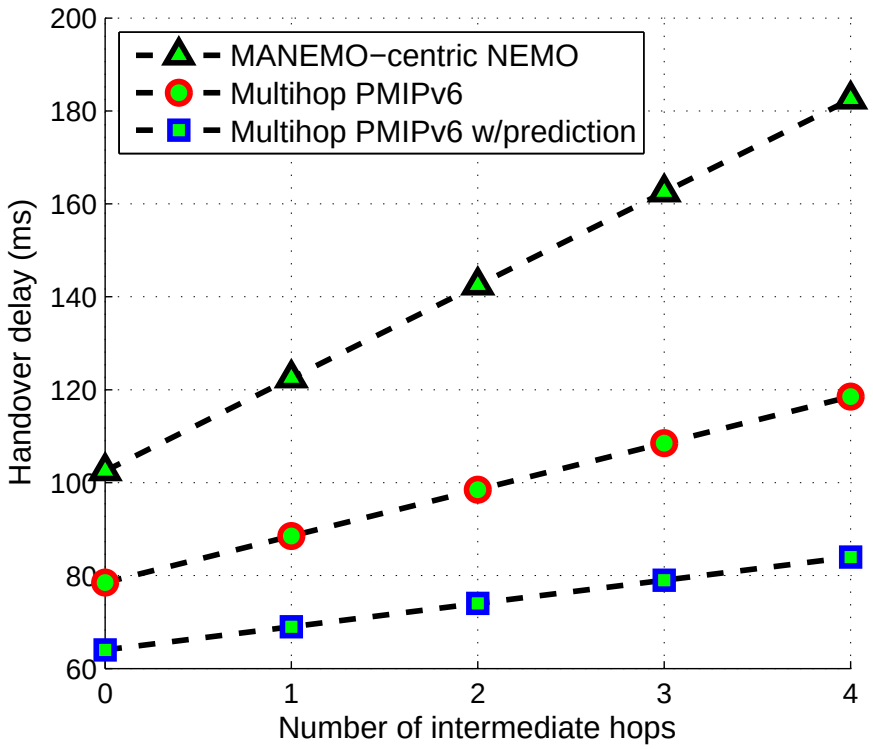


Figure 5.5: Handover latency comparison vs. number of intermediate hops [1]

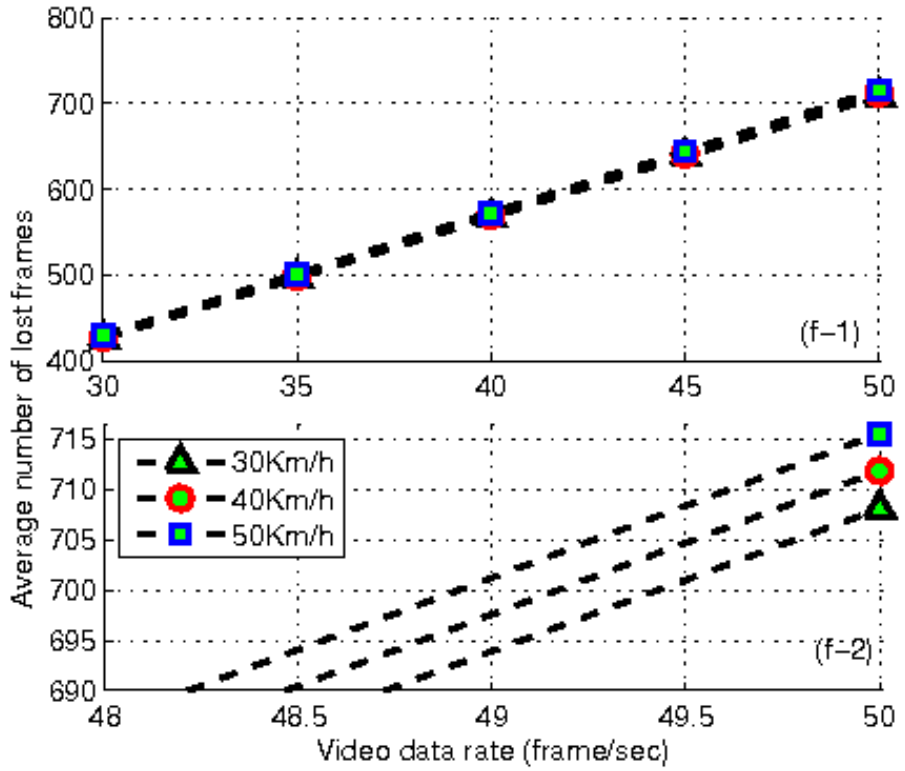


Figure 5.6: Frame loss vs. data rate (frame/sec.)

5.5 Summary

Vehicular networks have been envisioned to play an important role in the future wireless communication service market. Yet, the stream of high quality video to fast-moving vehicles is still fraught with fundamental challenges, attributed to the high mobility and dynamic nature of the network. In this chapter, integration of a quality-driven geo-routing protocol with a PMIPv6-based multi-hop scheme for seamless video streaming in urban vehicular scenarios, is considered. The integrated scheme takes advantage of the geographic features in VANET for data delivery. On one hand, the geo-routing protocol considers important metrics for video quality such as distortion, start-up delay and frequency of freezings, so that the multi-hop path is selected accordingly. On the other hand, the multi-hop PMIPv6 scheme provides a seamless delivery of data packets under the event of handovers, and incorporates a prediction mechanism based on the vehicle's location. The numerical results demonstrate that the proposed quality-driven routing protocol achieves in lower start-up delay, number of streaming freezes and frame distortion compared to the traditional geo-routing scheme. Similarly, the multi-hop PMIPv6 scheme achieves a low handover latency.

Chapter 6

Summary of Contributions and Future Research

6.1 Major Research Contributions

The major contributions of this thesis are mainly in the area of cross-layer protocol design for video streaming application over VANETs. In what follows, we provide a summary of contributions that have been achieved in this thesis and potential future research directions that can be followed in order to provide further contributions in this area.

- First, we started our cross-layer protocol design from MAC sub-layer. Specifically, we proposed a multi-objective optimization framework, which jointly minimizes the probability of playback freezes and start-up delay of the streamed video at the destination vehicle by tuning the MAC retransmission limit with respect to channel statistics as well as packet transmission rate, applied at the RSU. This scheme can be categorized as a cross-layer protocol between Link and Application layers since application layer streaming quality information are fed back to MAC sub-layer to

adjust the retransmission limit. Major attributes of this scheme are due to periodic channel state estimation at the RSU using RSS and Doppler shift effect. Estimates of access probability between the RSU and the destination vehicle is incorporated in the design of the adaptive MAC cross-layer scheme. The adaptation parameters are embedded in the UDP packet header. Two-hop transmission is applied in zones in which the destination vehicle is not within the transmission range of any RSU. In the multi-hop scenario, we discuss the computation of access probability used in the MAC adaptation scheme and proposed cross-layer path selection scheme followed by a discussion on mobile IP management to maintain continuous video streaming. Significantly fewer playback freezes at the cost of slight increase in start-up delay has been achieved using the proposed adaptive MAC retransmission scheme.

- Secondly, we have utilized the similarities of multi-hop urban vehicular networks in dense traffic conditions with mesh connected networks to extend our cross-layer routing protocol over mesh networks to VANET scenarios by considering new challenges arising from the highly dynamic nature of the network. A classification-based cross-layer protocol with exchange of information between Network and Application layers is proposed to select an optimal path for video streaming over multi-hop mesh networks. Our main contribution in this cross-layer routing protocol design is the translation of path selection over multi-hop networks to a standard classification problem. The classification is based on minimizing the average video packet distortion at the receiving nodes. We have also considered LRD characteristics of VBR video encoders by modeling the link as a FBM queue. Support vector machine has been applied at each hop to optimally assign the next hop (partial path) for each video packet received at that hop. WCC, including packet loss probability of the channel and maximum achievable rate, have been used as class prototypes.

Sample feature vectors include VCF and VEP which are extracted from video sequences. The classifiers are trained offline using a vast collection of video sequences and wireless channel conditions in order to yield optimal performance during real time path selection. Our method substantially reduces the complexity of conventional exhaustive optimization methods and results in high quality (low distortion). Simulations are conducted over an elementary multi-hop structure and cascades of such structure which not only proves the superiority of our method in terms of low complexity and high PSNR performance, but also provides important insights that can guide the design of network infrastructures and streaming protocols for video streaming. Next, we extend the proposed routing protocol over mesh network to real-time video transmission over urban multi-hop VANET scenarios. Queueing based mobility model, spatial traffic distribution and probability of connectivity for sparse and dense VANET scenarios are taken into consideration in designing the routing protocol. The numerical results demonstrate the gain achieved by the proposed routing protocol versus geographic greedy forwarding in terms of video frame distortion and streaming start-up delay in several urban communication scenarios for various vehicle entrance rate and traffic densities.

- Finally, in order to provide seamless and continuous streaming in drive-thru scenarios, we further proposed an efficient mobility management protocol. The proposed network mobility management scheme incorporates a handover prediction mechanism, and is a novel adaptation of PMIPv6 for multi-hop VANET. Numerical results show that our routing scheme achieves less frame distortion, lower start-up delay and fewer streaming freezes compared to the greedy geographical routing protocol. Also, the proposed multi-hop PMIPv6 achieves lower handover delay with less signalling cost when compared to the MANET-centric NEMO protocol.

6.2 Future Research

6.2.1 Quality-driven cooperative MAC protocol design

The probability distribution of the number of nodes within the transmission range of a randomly chosen node can be used to recognize the nodes that may interfere with the transmission of the random node. Clearly, this result will be useful in the design of effective MAC for the VANETs [56]. Cooperative communication can be used in order to extend the MAC retransmission limit adaptation scheme proposed in chapter 3. Also other contention parameters such CW size can be optimized jointly with retransmission limit in order to achieve results which are closer to the optimal performance.

6.2.2 Scalable Protocol Design via Node Clustering

VANETs are normally composed of a large number of nodes. Therefore, clustering VANETs into smaller manageable groups can be beneficial in terms of reducing energy consumption, improving resource utilization, balancing the load, reducing redundancy, retransmissions and collisions, or avoiding broadcasting storm. Partitioning the network into smaller manageable node clusters clearly supports scalability. Clustering can be used for saving energy, either by balancing the load among nodes so as not to drain certain nodes or by assigning sleep schedules for cluster members (CMs) by cluster heads (CHs). The application-centric routing protocol proposed in chapter 4 can be extended using node clustering in order to achieve better performance results with less complexity.

6.2.3 IP Management via Mobility Prediction

In situations where whole network moves such as VANETs, the standard NEMO protocol (which is designed for single-hop access) does not address the problem of seamless

multi-hop connection to the RSUs . Based on studies in [87], MANET-centric mobility management approaches seem to be more appropriate compared to NEMO-Centric Approaches because it allows a cleaner, modular integration at lower complexity with respect to economic, functional and performance criteria. They are more cost-efficient and have easier direct V2V communication with intermittent infrastructure access. Also they impose less complex support of geographic addressing while having better routing performances, because of easier integration with reactive, VANET-specific ad hoc routing protocols. While the above issues indicate an advantage of the MANET-centric approach over the NEMO-centric approach, currently no readily available solution for NEMO in VANETs exists and important aspects like routing consistency, privacy, and security need considerable research efforts.

Future wireless networks are expected to improve and safeguard QoS agreements (despite the users movement), hence, it is required to ensure availability of network services along the path for the lifetime of a connection. Therefore, mobility prediction schemes can be used in order to further improve the proposed mobile IP solution in chapter 5. Most models heavily rely on historical data which can be related to individual or aggregate mobility. Most work done for location predictability can be categorized in the following mobility prediction models:

- Order-k Markov predictor: take a sequence of symbols as a history string, and try to predict the next symbol from the current context [88].

- Hidden Markov model (HMM): output variables (observable variables) are distinguished from state variables (hidden variables). There is a one-to-one mapping between output/state variables. The system goes through some learning process in order to estimate hidden parameters and can be used to predict the output

parameters (future location of the vehicle) [89].

- Markov renewal theory: holding times (time between state transitions) are random variables with any distribution (semi-Markov process). Sojourn time in any state depends on current-state and next state transition [90].

In VANET, if we consider sections of the road between two intersections as location areas, Order-k predictor can be used for location prediction using current state (velocity, direction) for temporal prediction. MAG can perform the predictions and inform the local mobility anchor (LMA) about the result. Hence, LMA can pre-establish the tunnel with the most likely next AR based on prediction.

6.2.4 Energy Efficient Protocol Design in Green VANETs

Green communication and networking solutions can not only improve the efficiency of energy consumption and resource utilization, but also benefit global environment. Cross-layer protocol design across multiple OSI layers can be utilized to improve energy and spectrum efficiency of green communication networks, especially in emerging markets such as VANETs. In terms of analysis, queueing models can be applied in order to study the behavior of energy packets with similarities to what have been previously done for analysis of data packet transmission.

Appendix A

Derivation of Connectivity Probability Bounds

The probability of connectivity is a key factor in designing efficient routing schemes because packet forwarding through radio propagation is faster than packet carrying by a vehicle. In fact, sometimes, routing the packets from a longer path is more suitable than shorter paths where the former could provide higher connectivity which consequently will result in less packet delivery delay.

As defined in chapter 4, probability of connectivity is the probability that the vehicles along the streets can communicate with each other through a path which could be direct connection or multi-hop connection. Two vehicles are able to directly communicate if they are in the radio transmission range of each other. Therefore, to find the probability of non-connectivity, in general, we need to find the probability of interruption of all radio links between any two vehicles.

Lower and upper bounds for the probability of connectivity along a typical street in the road topology is derived in [63]. Along a typical street, the radio link between two

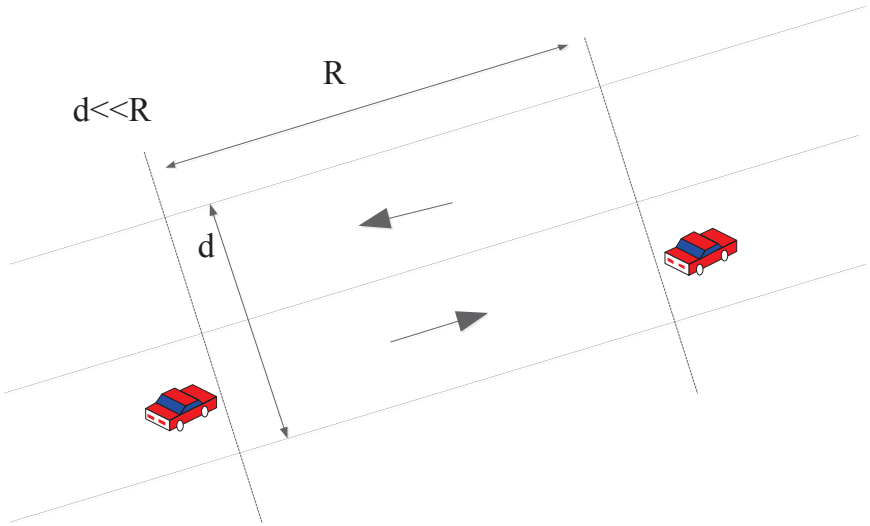


Figure A.1: Certain non-connectivity when any of the substreets with length R is empty.

vehicles will be disconnected if there are two consecutive vehicles with a distance greater than the radio transmission range R. In their approach, they have divided each part of the street (i.e., front, middle, and end parts) into substreets, each with length R. With respect to the previous approach in mapping parts of streets onto queues in a queueing network, we will have a queueing network similar to the one proposed in Section 4.2.1 but with greater number of queues.

In this queueing network, a departed vehicle (customer) with a specific class from a queue corresponding to a substreet is routed to the queue corresponding to the following substreet as a customer with the same class. Moreover, if the substreets belong to different parts of the street, the customer class may change. To find the probability of non-connectivity along a typical street, we find the probability that there is at least one pair of side-by-side empty substreets (queues) with opposite directions (Figure A.1).

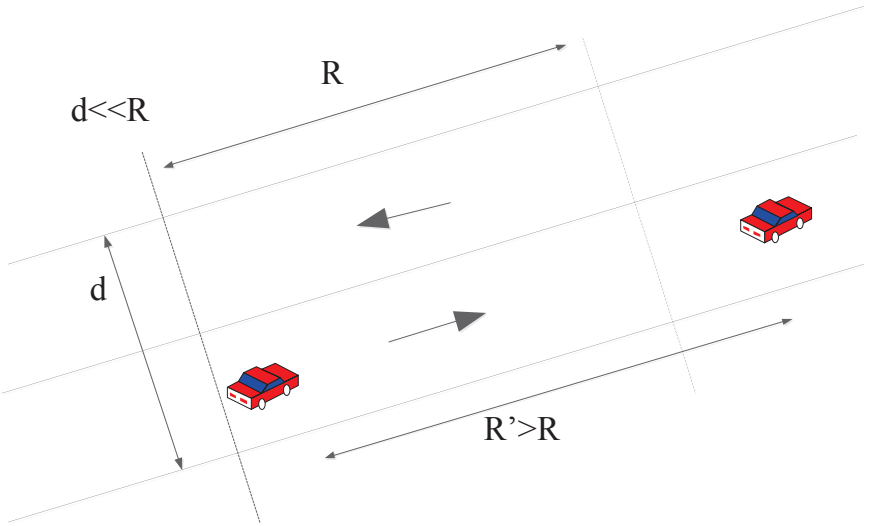


Figure A.2: Possible non-connectivity with substreets with length R are non-empty.

However, if we have such a pair of queues, the connectivity does not necessarily hold along the street (e.g., Figure A.2). Therefore, the computed probability is only a lower bound for non-connectivity, which is equivalent to an upper bound for connectivity.

On the other hand, consider the case in which we divide the streets into substreets with length $R/2$. Then, if we have at least one vehicle at each substreet (irrespective of its direction and speed), we will have full connectivity for that street. However, if we have an empty pair of side-by-side substreets in this new structure, connectivity may also hold (see Figure A.3).

In this case, connectivity depends on the locations of the vehicles in the following and preceding substreets. Therefore, by computing the probability with which all pairs of side-by-side queues are nonempty, we will find a lower bound for connectivity. With respect to the foregoing discussions, we have the following relations for the lower and upper bounds for the probability of connectivity at a typical street:

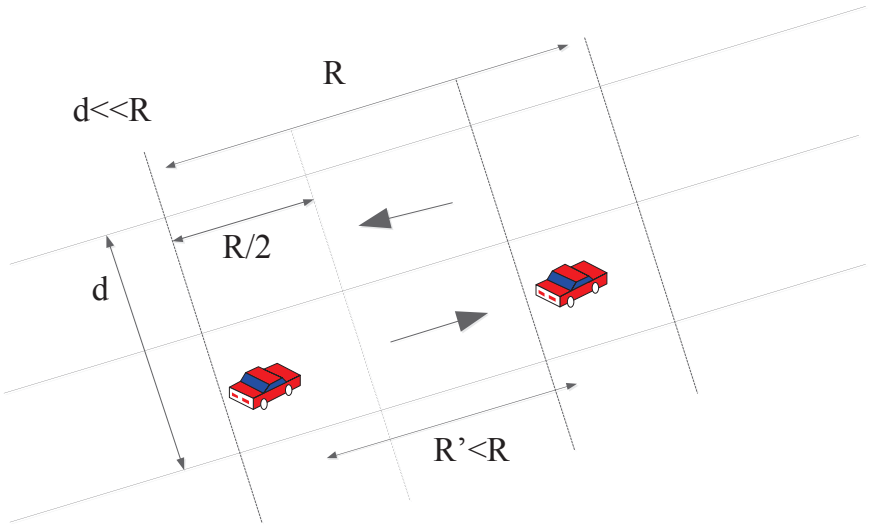


Figure A.3: Possible connectivity even if a substreet with length $R/2$ is empty.

Upper bound:

$$\begin{aligned} P(\text{Connectivity}) &\leq \sum_{\substack{n_i \geq 1 \\ 1 \leq i \leq N_S(R)}} \pi(n_1, n_2, ..., n_{N_S(R)}) \\ &= \prod_{i=1}^{N_S(R)} (1 - \pi(n_i = 0)) \\ &= \prod_{i=1}^{N_S(R)} \sum_{n_i \geq 1} \pi(n_i) \end{aligned} \tag{A.1}$$

Lower bound:

$$\begin{aligned} P(\text{Connectivity}) &\geq \sum_{\substack{n_i \geq 1 \\ 1 \leq i \leq N_S(R/2)}} \pi(n_1, n_2, ..., n_{N_S(R/2)}) \\ &= \prod_{i=1}^{N_S(R/2)} \sum_{n_i \geq 1} \pi(n_i) \end{aligned} \tag{A.2}$$

where $N_S(R/2)$ and $N_S(R)$ denote the number of substreets for a typical street S, where the length of each substreet equals R and $R/2$, respectively.

Appendix B

BCMP product form queueing networks

BCMP network is a class of queueing network for which a product form equilibrium distribution exists. The class of systems under consideration contains an arbitrary but finite number of service centers denoted by N. There is an arbitrary but finite number R of different classes of customers. Customers travel through the network and change class according to transition probabilities. Thus a customer of class r who completes service at service center i will next require service at center j in class s with a certain probability denoted by $P_{i,r,j,s}$. The transition matrix $P = [P_{i,r,j,s}]$ can be considered as defining a Markov chain whose states are labeled by the pairs (i, r). The Markov chain is assumed to be decomposable into m ergodic subchains which is utilized in [91] in order to derive statitics of such queueing networks.

According to [91], a network of m interconnected queues is known as a BCMP network if each of the queues is of one of the following four types:

1. The service discipline is first-come-first-served (FCFS); all customers have the same

service time distribution at this service center, and the service time distribution is a negative exponential. The service rate can be state dependent where $\mu(j)$ will denote the service rate with j customers at the center.

2. There is a single server at a service center, the service discipline is processor sharing, i.e., when there are n customers in the service center each is receiving service at a rate of $1/n$, and each class of customer may have a distinct service time distribution. The service time distributions have rational Laplace transform.

3. The number of servers in the service center is greater than or equal to the maximum number of customers that can be queued at this center in a feasible state, and each class of customer may have a distinct service time distribution. The service time distributions have rational Laplace transforms.

4. There is a single server at a service center, the queuing discipline is preemptive-resume last-come-first-served (LCFS), and each class of customer may have a distinct service time distribution. The service time distributions have rational Laplace transforms.

In the final three cases, service time distributions must have rational Laplace transforms. This means the Laplace transform must be of the form

$$L(s) = \frac{N(s)}{D(s)} \tag{B.1}$$

where $N(s)$ and $D(s)$ are polynmials in s, degree of $N(s)$ is less than that of $D(s)$, all roots of $D(s)$ are real, and $L(0) = 1$.

For a BCMP network of m queues which is open, closed or mixed in which each queue is of type 1, 2, 3 or 4, the equilibrium state probabilities (probability of number of

customers in each queue) are given by a product form as follows.

$$\pi(n_1, ..., n_m) = C\pi_1(n_1)\pi_2(n_2)...\pi_m(n_m) \tag{B.2}$$

where C is a normalizing constant and $\pi(.)$ represents the equilibrium distribution for queue i. In the case of an open network , C is always 1. For a closed network, C is determined by the constraint which makes the summation of the state probabilities equal to 1; i.e.,

$$C = \frac{1}{\sum_{\text{all states}\in\{n'_1,...n'_M\}} \pi_1(n'_1)\pi_2(n'_2)...\pi_m(n'_m)} \tag{B.3}$$

where M is total number of classes and N is total number of customers in the system, i.e., $\sum_{m=1}^{M} n'_m = N$.

Appendix C

Support Vector Machines

C.1 Supervised Learning and Generalization

The input/output pairings typically reflect a functional relationship which maps the inputs to outputs. This is not always the case as for example when the outputs are corrupted by noise. When an underlying function from inputs to outputs exists it is referred to as the target function. The estimate of the target function is known as the solution of the learning problem. In the case of classification this function is sometimes referred to as the decision function. The solution is chosen from a set of candidate functions which map from the input space to the output domain. Usually we will choose a particular set or class of candidate functions known as hypotheses before we begin trying to learn the correct function. For example, so-called decision trees are hypotheses created by constructing a binary tree with simple decision functions at the internal nodes and output values at the leaves. Hence, we can view the choice of the set of hypotheses (or hypothesis space) as one of the key ingredients of the learning strategy. The algorithm which takes the training data as input and selects a hypothesis from the hypothesis space

is the second important ingredient. It is referred to as the learning algorithm [92]. Hence the challenges can be summarized as follows.

- The first challenge is that the target function may not have a simple representation.
- The second problem is that training data are usually noisy and so there is no guarantee that there is an underlying function which correctly maps the training data.
- There is, however, a more fundamental problem with this approach in that even when we can find a hypothesis that is consistent with the training data, it may not make correct classifications of unseen data.

The ability of a hypothesis to correctly classify data not in the training set is known as its generalization, and it is this property that must be optimized.

C.2 Kernel induced feature spaces

In general, complex real-world applications require more expressive hypothesis spaces than linear functions. Another way of viewing this problem is that frequently the target concept cannot be expressed as a simple linear combination of the given attributes, but in general requires that more abstract features of the data be exploited. Multiple layers of linear functions were proposed as a solution to this problem, and this approach led to the development of multi-layer neural networks and learning algorithms such as back-propagation for training such systems [92].

Kernel representations offer an alternative solution by projecting the data into a high dimensional feature space to increase the computational power of the linear learning machines. The advantage of using the machines in the dual representation derives from the fact that in this representation the number of tunable parameters does not depend on the

number of attributes being used. By replacing the inner product with an appropriately chosen Śkernel' function, one can implicitly perform a non-linear mapping to a high dimensional feature space without increasing the number of tunable parameters, provided the kernel computes the inner product of the feature vectors corresponding to the two inputs.

The kernel technique provides one of the main building blocks of Support Vector Machines (SVMs). One of the remarkable features of SVMs is that to a certain extent the approximation-theoretic issues are independent of the learning-theoretic ones. One of the remarkable features of SVMs is that to a certain extent the approximation-theoretic issues are independent of the learning-theoretic ones. One can therefore study the properties of the kernel representations in a general and self-contained way, and use them with different learning theories. Another attraction of the kernel method is that the learning algorithms and theory can largely be decoupled from the specifics of the application area, which must simply be encoded into the design of an appropriate kernel function. Hence, the problem of choosing an architecture for a neural network application is replaced by the problem of choosing a suitable kernel for a SVM.

The complexity of the target function to be learned depends on the way it is represented, and the difficulty of the learning task can vary accordingly. Ideally a representation that matches the specific learning problem should be chosen. So one common preprocessing strategy in machine learning involves changing the representation of the data:

$$\mathbf{x} = (x_1, ...x_n) \rightarrow \phi(\mathbf{x}) = (\phi_1(\mathbf{x}), ...\phi_N(\mathbf{x})) \tag{C.1}$$

The quantities introduced to describe the data are usually called features, while the

original quantities are sometimes called attributes. The task of choosing the most suitable representation is known as feature selection. The space X is referred to as the input space, while $F = \{\phi(x)|x \in X\}$ is called the feature space (Figure C.1). Different approaches to feature selection exist. Frequently one seeks to identify the smallest set of features that still conveys the essential information contained in the original attributes. This is known as dimensionality reduction.

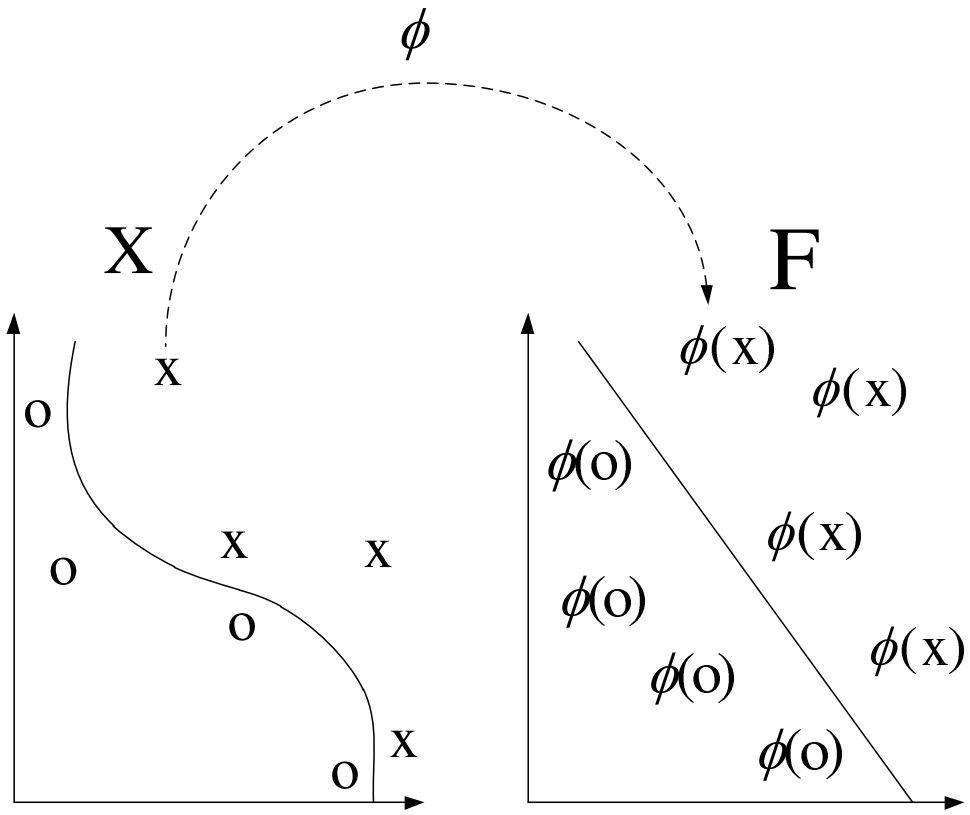


Figure C.1: Simplification of the classification task by feature mapping

In order to learn non-linear relations with a linear machine, we need to select a set of non-linear features and to rewrite the data in the new representation. This is equivalent to applying a fixed non-linear mapping of the data to a feature space, in which the linear machine can be used. Hence, the set of hypotheses we consider will be functions of the

type:

$$f(\mathbf{x}) = \sum_{i=1}^{N} w_i \phi_i(\mathbf{x}) + b \tag{C.2}$$

where $\phi : X \to F$ is a non-linear map from the input space to some feature space. one important property of linear learning machines is that they can be expressed in a dual representation. This means that the hypothesis can be expressed as a linear combination of the training points, so that the decision rule can be evaluated using just inner products between the test point and the training points:

$$f(\mathbf{x}) = \sum_{i=1}^{l} \alpha_i \lambda_i \langle \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}) \rangle + b \tag{C.3}$$

If we have a way of computing the inner product $\langle \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}) \rangle$ in feature space directly as a function of the original input points, it becomes possible to merge the two steps needed to build a non-linear learning machine. We call such a direct computation method a kernel function. Hence, a kernel function K is defined such that for all $x, z \in X$

$$K(x, z) = \langle \phi(x) \cdot \phi(z) \rangle \tag{C.4}$$

where ϕ is a mapping from X to an (inner product) feature space F.

C.3 Maximal Margin Classifier

The goal of classification via SVMs is to devise a computationally efficient way of learning good separating hyperplanes in a high dimensional feature space [92]. The simplest model of Support Vector Machine, which was also the first to be introduced, is the so-called maximal margin classifier. It works only for data which are linearly separable in the feature space, and hence cannot be used in many real-world situations. Nonetheless it is

the easiest algorithm to understand, and it forms the main building block for the more complex Support Vector Machines. The maximal margin classifier forms the strategy of the first Support Vector Machine, namely to find the maximal margin hyperplane in an appropriately chosen kernel-induced feature space. This strategy is implemented by reducing it to a convex optimization problem, i.e., minimizing a quadratic function under linear inequality constraints. If w is the weight vector realizing a functional margin of 1 on the positive point x^+ and the negative point x^-, we can compute its geometric margin as follows. A functional margin of 1 implies:

$$\mathbf{w} \cdot \mathbf{x}^+ + b = 1 \tag{C.5}$$

$$\mathbf{w} \cdot \mathbf{x}^- + b = 1 \tag{C.6}$$

while to compute the geometric margin we must normalize $\mathbf{w}$. The geometric margin γ is then the functional margin of the resulting classifier.

$$\gamma = \frac{1}{2}\left(\frac{\mathbf{w}}{\| \mathbf{w} \|_2} \cdot \mathbf{x}^+\right) - \left(\frac{\mathbf{w}}{\| \mathbf{w} \|_2} \cdot \mathbf{x}^-\right) = \frac{1}{\| \mathbf{w} \|_2} \tag{C.7}$$

The hyperplane $(\mathbf{w}, b)$ that solves the optimization problem for linearly sparable points $(\mathbf{x_1}, y_1)...(\mathbf{x_l}, y_l)$ is as follows.

$$\begin{aligned} \text{minimize} \quad & \mathbf{w} \cdot \mathbf{w} \\ \text{subject to} \quad & y_i(\mathbf{w} \cdot \mathbf{x}^+ + b) \geq 1, \\ & i = 1...l \end{aligned} \tag{C.8}$$

The use of linear machines in the dual representation (dual representation of C.8) makes it possible to perform this step implicitly.

Bibliography

[1] M. Asefi, S. Cespedes, J. W. Mark, and X. Shen, "A seamless quality-driven multi-hop data delivery scheme for video streaming in urban vanet scenarios," *In Proc. of IEEE ICC, Kyoto, Japan*, June 2011.

[2] H. T. Cheng, H. Shan, and W. Zhuang, "Infortainment and Road Safety Service Support in Vehicular Networking: From a Communication Perspective," *Mechanical Systems and Signal Processing*, vol. 26, pp. 2020–2038, 2010.

[3] A. Boukerche, H. A. B. F. Oliveira, E. F. Nakamura, and A. A. F. Loureiro, "Vehicular Ad Hoc Networks: A New Challenge for Localization-Based Systems," *Computer Communications*, vol. 31, no. 12, pp. 2838 – 2849, 2008.

[4] T. H. Luan, X. Ling, and X. Shen, "MAC in Motion: Impact of Mobility on the MAC of Drive-Thru Internet," *IEEE Trans. Mobile Comput.*, to appear.

[5] P. Seeling, M. Reisslein, and B. Kulapala, "Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial," *IEEE Communications Surveys and Tutorials*, vol. 6, pp. 58–78, Third Quarter 2004.

[6] P. Seeling, F. Fitzek, and M. Reisslein, *Video Traces for Network Performance Evaluation*. Springer Netherlands, 2007.

[7] G. Van Der Auwera, P. T. David, and M. Reisslein, "Traffic and Quality Characterization of Single-Layer Video Streams Encoded with the H.264/MPEG-4 Advanced Video Coding Standard and Scalable Video Coding Extension," *IEEE Trans. Broadcas.*, vol. 54, pp. 698 –718, Sept. 2008.

[8] M. Van Der Schaar and S. Shankart, "Cross-Layer Wireless Multimedia Transmission: Challenges, Principles, and New Paradigms," *IEEE Wireless Communications*, vol. 12, pp. 50–58, 2005.

[9] S. Shankar and M. V. D. Schaar, "Performance Analysis of Video Transmission over IEEE 802.11a/e WLANs," *IEEE Trans. on Veh. Technol.*, vol. 56, no. 4, pp. 2346–2364, 2007.

[10] S. Shakkottai and T. S. Rappaport, "Cross-Layer Design for Wireless Networks," *IEEE Communications Magazine*, vol. 41, pp. 74–80, Oct. 2003.

[11] M. Van Der Schaar, D. S. Turaga, and R. Wong, "Classification-Based System For Cross-Layer Optimized Wireless Video Transmission," *IEEE Trans. Multimedia*, vol. 8, pp. 1041–1078, Oct. 2006.

[12] "IEEE Std. 802.11-1999, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-speed Physical Layer Extension in the 2.4 GHz Band, IEEE Std. 802.11b-1999, 1999,"

[13] M. Van Der Schaar and P. A. Chou, *Multimedia Over IP Wireless Networks*. Elseivier, 2007.

[14] M. Emmelmanmn, B. Bochow, and C. Kellum, *Vehicular Networking: Automotive Applications and Beyond*. WILEY, 2010.

[15] E. Schoch, F. Kargl, M. Weber, and T. Leinmuller, "Communication Patterns in VANETs," *IEEE Communications Magazine*, vol. 46, no. 11, pp. 119–125, 2008.

[16] C.-F. Chiasserini and R. Gaeta and M. Garetto and M. Gribaudo and M. Sereno, "Efficient broadcasting of safety messages in multihop vehicular networks," *In Proc. of 20th IEEE International Parallel and Distributed Processing Symposium, Washington, DC, USA*, April 2006.

[17] M. Khabazian, S. Aissa, and M. Mehmet Ali, "Performance Modeling of Message Dissemination In Vehicular Ad Hoc Networks with Priority," *IEEE JSAC*, vol. 29, pp. 61 –71, Jan. 2011.

[18] F. Bai, H. Krishnan, V. Sadekar, G. Holl, and T. Elbatt, "Towards characterizing and classifying communication-based automotive applications from a wireless networking perspective," *In Proc. of IEEE Workshop on Automotive Networking and Applications (AutoNet), San Francisco, CA, USA*, 2006.

[19] J. Camp and E. Knightly, "Modulation Rate Adaptation in Urban and Vehicular Environments: Cross-Layer Implementation and Experimental Evaluation," *IEEE/ACM Trans. Netw.*, vol. 18, pp. 1949 –1962, Dec. 2010.

[20] J. Kim, S. Kim, S. Choi, and D. Qiao, "CARA: Collision-aware rate adaptation for IEEE 802.11 WLANs," *In Proc. of IEEE INFOCOM, Barcelona, Spain*, pp. 1–11, April 2006.

[21] G. Holland, N. Vaidya, and P. Bahl, "A rate-adaptive mac protocol for multi-hop wireless networks," *In Proc. of ACM MobiCom, Rome, Italy*, pp. 236–251, 2001.

[22] H. Zhai, J. Wang, and Y. Fang, "DUCHA: A New Dual-Channel MAC Protocol for Multihop Ad Hoc Networks," *IEEE Trans. on Wireless Commun.*, vol. 5, pp. 3224 –3233, Nov. 2006.

[23] Y. Kwon, Y. Fang, and H. Latchman, "A novel MAC protocol with fast collision resolution for wireless LANs," *In Proc. of IEEE INFOCOM, San Francisco, CA, USA*, vol. 2, pp. 853 –862, April 2003.

[24] Y.-C. Lai, P. Lin, W. Liao, and C.-M. Chen, "A Region-Based Clustering Mechanism for Channel Access in Vehicular Ad Hoc Networks," *IEEE JSAC*, vol. 29, pp. 83 –93, Jan. 2011.

[25] G. Korkmaz, E. Ekici, and F. Ozguner, "A Cross-Layer Multihop Data Delivery Protocol with Fairness Guarantees for Vehicular Networks," *IEEE Trans. Veh. Technol.*, vol. 55, pp. 865 –875, May 2006.

[26] H. Menouar, F. Filali, and M. Lenardi, "A Survey and Qualitative Analysis of MAC Protocols for Vehicular Ad Hoc Networks," *IEEE Wireless Communications*, vol. 13, pp. 30 –35, Oct. 2006.

[27] F. Borgonovo, A. Capone, M. Cesana, and L. Fratta, "ADHOC MAC: New MAC Architecture for Ad Hoc Networks Providing Efficient and Reliable Point-to-Point and Broadcast Services," *Wirel. Netw.*, vol. 10, pp. 359–366, July 2004.

[28] G. Jakllari, W. Luo, and S. V. Krishnamurthy, "An integrated neighbor discovery and mac protocol for ad hoc networks using directional antennas," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1114 –1024, March 2007.

[29] T. Korakis, G. Jakllari, and L. Tassiulas, "CDR-MAC: A Protocol for Full Exploitation of Directional Antennas in Ad Hoc Wireless Networks," *IEEE Trans. on Mobile Comput.*, vol. 7, pp. 145 –155, Feb. 2008.

[30] H. Zhang, Y. Ma, D. Yuan, and H.-H. Chen, "Quality-of-Service Driven Power and Sub-Carrier Allocation Policy for Vehicular Communication Networks," *IEEE JSAC*, vol. 29, pp. 197 –206, Jan. 2011.

[31] S.-H. Wu, C.-M. Chen, and M.-S. Chen, "An Asymmetric and Asynchronous Energy Conservation Protocol for Vehicular Networks," *IEEE Trans. on Mobile Comput.*, vol. 9, pp. 98 –111, Jan. 2010.

[32] H. Xiong, R. Li, A. Eryilmaz, and E. Ekici, "Delay-Aware Cross-Layer Design for Network Utility Maximization in Multi-Hop Networks," *IEEE JSAC*, vol. 29, pp. 951 –959, May 2011.

[33] Y. Chang, C. Lee, and J. Copeland, "Goodput Enhancement of VANETs in Noisy CSMA/CA Channels," *IEEE JSAC*, vol. 29, pp. 236 –249, Jan. 2011.

[34] A. Benslimane, T. Taleb, and R. Sivaraj, "Dynamic Clustering-Based Adaptive Mobile Gateway Management in Integrated VANET-3G Heterogeneous Wireless Networks," *IEEE JSAC*, vol. 29, pp. 559 –570, March 2011.

[35] Z. Ding and K. Leung, "Cross-Layer Routing Using Cooperative Transmission in Vehicular Ad-hoc Networks," *IEEE JSAC*, vol. 29, pp. 571 –581, March 2011.

[36] T. Taleb, E. Sakhaee, A. Jamalipour, K. Hashimoto, N. Kato, and Y. Nemoto, "A Stable Routing Protocol to Support ITS Services in VANET Networks," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 3337 –3347, Nov. 2007.

[37] M. Zhang and R. Wolff, "Routing protocols for vehicular Ad Hoc networks in rural areas," *IEEE Communications Magazine*, vol. 46, pp. 126 –131, Nov. 2008.

[38] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET Routing on City Roads Using Real-Time Vehicular Traffic Information," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 3609 –3626, Sep. 2009.

[39] Q. Mussabbir, W. Yao, Z. Niu, and X. Fu, "Optimized FMIPv6 Using IEEE 802.21 MIH Services in Vehicular Networks," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 3397 –3407, Nov. 2007.

[40] Jaeok Park and Van Der Schaar, M., "Adaptive MAC Protocols Using Memory for Networks With Critical Traffic," *IEEE Trans. Signal Process.*, vol. 59, pp. 1269 –1279, March 2011.

[41] H. Luan, L. X. Cai, and X. Shen, "Impact of Network Dynamics on User's Video Quality: Analytical Framework and QoS Provision," *IEEE Trans. on Multimedia,*, vol. 12, pp. 64 –78, Jan. 2010.

[42] X. Tong, Y. Andreopoulos, and M. Van Der Schaar, "Distortion-Driven Video Streaming over Multihop Wireless Networks with Path Diversity," *IEEE Trans. Mobile Comput.*, vol. 6, pp. 1343–1356, Dec. 2007.

[43] B. Qazi, H. Alshaer, and J. Elmirghani, "Analysis and Design of a MAC Protocol and Vehicular Traffic Simulator for Multimedia Communication on Motorways," *IEEE Trans. on Veh. Technol.*, vol. 59, pp. 734 –741, Feb. 2010.

[44] G. Korkmaz, E. Ekici, and F. Özgüner, "Supporting Real-Time Traffic in Multihop Vehicle-to-Infrastructure Networks," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 3, pp. 376 – 392, 2010.

[45] J. Zhao and G. Cao, "VADD: Vehicle-Assisted Data Delivery in Vehicular Ad Hoc Networks," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 1910 –1922, May 2008.

[46] F. Li, G. Liu, and L. He, "Application-driven cross-layer approaches to video transmission over downlink OFDMA networks," *In Proc. of IEEE GLOBECOM Workshops, Honolulu, Hawaii, USA*, pp. 1 –6, Dec. 2009.

[47] "Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle SystemsŮ5GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2007.

[48] "IEEE Trial-Use Standard for Wireless Access in Vehicular Environments (WAVE) - Multi-Channel Operation," *IEEE Std. 1609.4*, pp. c1 –74, 2006.

[49] S. Ou, K. Yang, H. Chen, and A. Galis, "A Selective Downlink Scheduling Algorithm to Enhance Quality of VOD Services for WAVE Networks," *EURASIP Journal of Wireless Communications and Networking*, pp. 1–12, Jan. 2009.

[50] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity," in *In Proc. of the 37th IEEE Conference on Decision and Control*, vol. 1, pp. 1106–1110, 1998.

[51] M. Desai and D. Manjunath, "On the Connectivity in Finite Ad Hoc Networks," *IEEE Commun. Lett.*, vol. 6, pp. 437 – 439, Oct. 2002.

[52] J. Li, L. Andrew, C. H. Foh, M. Zukerman, and M. Neuts, "Meeting Connectivity Requirements in a Wireless Multihop Network," *IEEE Commun. Lett.*, vol. 10, pp. 19 – 21, Jan. 2006.

[53] A. Ghasemi and S. Nader-Esfahani, "Exact Probability of Connectivity One-Dimensional Ad-Hoc Wireless Networks," *IEEE Commun. Lett.*, vol. 10, pp. 251 – 253, Apr. 2006.

[54] G. Farhadi and N. Beaulieu, "On the connectivity and average delay of mobile ad hoc networks," *In Proc. of IEEE ICC, Istanbul, Turkey*, vol. 8, pp. 3868 –3872, June 2006.

[55] G. Han and A. M. Makowski, "A Very Strong Zero-One Law for Connectivity in One-Dimensional Geometric Random Graphs," *IEEE Commun. Lett.*, vol. 11, pp. 55 –57, Jan. 2007.

[56] M. Khabazian and M. Ali, "A Performance Modeling of Connectivity in Vehicular Ad Hoc Networks," *IEEE Trans. Veh. Technol.*, vol. 57, pp. 2440 –2450, July 2008.

[57] S. Panichpapiboon and W. Pattara-atikom, "Connectivity Requirements for Self-Organizing Traffic Information Systems," *IEEE Trans. on Veh. Technol.*, vol. 57, pp. 3333 –3340, Nov. 2008.

[58] S. Yousefi, E. Altman, R. El-Azouzi, and M. Fathy, "Analytical Model for Connectivity in Vehicular Ad Hoc Networks," *IEEE Trans. on Veh. Technol.*, vol. 57, pp. 3341 –3356, Nov. 2008.

[59] D. Miorandi and E. Altman, "Connectivity in One-Dimensional Ad-Hoc Networks: A Queueing Theoretical Approach," *Wireless Networks*, vol. 12, pp. 573–587, 2006.

[60] J. Wu, "Connectivity analysis of a mobile vehicular ad hoc network with dynamic node population," *In Proc. IEEE GLOBECOM Workshops*, pp. 1–8, Dec. 2008.

[61] W.-L. Jin and W. Recker, "An Analytical Model of Multihop Connectivity of Inter-Vehicle Communication Systems," *IEEE Trans. on Wireless Commun.*, vol. 9, pp. 106 –112, Jan. 2010.

[62] F. Bai, S. Narayanan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for ad hoc networks," *In Proc. of IEEE INFOCOM, San Francisco, CA, USA*, vol. 2, pp. 825 – 835, March-April.

[63] G. Mohimani, F. Ashtiani, A. Javanmard, and M. Hamdi, "Mobility Modeling, Spatial Traffic Distribution, and Probability of Connectivity for Sparse and Dense Vehicular Ad Hoc Networks," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1998–2007, May 2009.

[64] M. Asefi, J. Mark, and X. Shen, "A cross-layer path selection scheme for video streaming over vehicular ad-hoc networks," *In Proc. of IEEE 72nd VTC-Fall*, Sep. 2010.

[65] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. Tonguz, "Routing in Sparse Vehicular Ad Hoc Wireless Networks," *IEEE JSAC*, vol. 25, pp. 1538 –1556, Oct. 2007.

[66] S. C. Ng, W. Zhang, Y. Zhang, Y. Yang, and G. Mao, "Analysis of Access and Connectivity Probabilities in Vehicular Relay Networks," *IEEE JSAC*, vol. 29, pp. 140 –150, Jan. 2011.

[67] Y. Chang, M. Lee, and J. Copeland, "An adaptive on-demand channel estimation for vehicular ad hoc networks," *In Proc. IEEE CCNC, Las Vegas, NV, USA*, pp. 1 –5, Jan. 2009.

[68] Q. Li and M. Van Der Schaar, "Providing Adaptive QoS to Layered Video Over Wireless Local Area Networks Through Real-Time Retry Limit Adaptation," *IEEE Trans. Multimedia*, vol. 6, pp. 278–290, April 2004.

[69] S. Kompella, M. Shiwen, Y. T. Hou, and H. T. Sherali, "On path selection and rate allocation for video in wireless mesh networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 212–224, 2009.

[70] H. Zhu, R. Lu, X. Shen, and X. Lin, "Security in Service-Oriented Vehicular Networks - [Service-Oriented Broadband Wireless Network Architecture]," *IEEE Wireless Communication*, vol. 16, pp. 16–22, Oct. 2009.

[71] K. Stuhlmuller, M. Link, and B. Girod, "Analysis of Video Transmission over Lossy Channels," *IEEE JSAC*, vol. 18, no. 6, pp. 1012–1032, 2000.

[72] M. Lelarge, Z. Liu, and C. H. Xia, "Asymptotic tail distribution of end-to-end delay in networks of queues with self-similar cross traffic," *In Proc. of IEEE INFOCOM, Atlanta, USA*, pp. 2352–2363, 2004.

[73] I. Norros, "On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks," *IEEE JSAC*, vol. 13, no. 6, pp. 953–961, 1995.

[74] G. D. Van Der Auwera and M. Reisslein, "Traffic and Quality Characterization of Single Layer Video Streams Encoded with H.264/MPEG-4 Advanced Video Coding Standard and Scalable Video Coding Extension," *IEEE Trans. on Broadcas.*, vol. 54, pp. 698–718, 2008.

[75] I. Reljin, A. Samčović, and B. Reljin, "H.264/AVC Video Compressed Traces: Multifractal and Fractal Analysis," *EURASIP J. Appl. Signal Process.*, pp. 123–123, 2006.

[76] M. Taqqu, V. Teverovsky, and W. Willinger, "Estimators for Long-Range Dependence: An Empirical Study," *Fractals*, vol. 3, pp. 785–798, 1995.

[77] C. J. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Disc.*, vol. 2, pp. 121–167, 1998.

[78] H. Shiang and M. Van Der Schaar, "Multi-User Video Streaming over Multi-Hop Wireless Networks: A Distributed, Cross-Layer Approach Based on Ariority Queuing," *IEEE JSAC*, vol. 25, no. 4, pp. 770–785, 2007.

[79] X. Chao, M. Miyazawa, and M. Pinedo, *Queuing Networks*. Hoboken, NJ: WILEY, 2000.

[80] G. Bolch, S. Greiner, H. d Meer, and K. S. Trivedi, *Queuing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Scienece Applications, 2nd Edition*. WILEY, 2006.

[81] D. Cox, "A use of complex probabilities in the theory of stochastic processes," *In Proc. Cambridge Philosophical Society*, vol. 51, pp. 313–319, 1955.

[82] S. Olariue and M. C. Weigle, *Vehicuar Networks, From Theory to Practice*. Chapman and Hall/CRC, 2009.

[83] A. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, "Proxy Mobile IPv6," *IETF RFC 5213*, August 2008.

[84] R. Baldessari, C. J. Bernardos, and M. Calderon, "GeoSAC - Scalable Address Autoconfiguration for VANET Using Geographic Networking Concepts," *In Proc. of 19th IEEE PIMRC, 2008.*, vol. 2, pp. 1–7.

[85] J. Choi, Y. Khaled, M. Tsukada, and T. Ernst, "IPv6 support for VANET with geographical routing," *In Proc. of IEEE 8th ITST*, Oct. 2008.

[86] T. Narten, E. Nordmark, W. Simpson, and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)," *IETF RFC 4861*, September 2007.

[87] R. Baldessari, A. Festag, and J. Abeille, "NEMO meets VANET: A Deployability Analysis of Network Mobility in Vehicular Communication," *In Proc. of IEEE 7th International Conference on ITS*, pp. 375–380, June 2007.

[88] I. Butun, A. Cagatay Talay, D. Turgay Altilar, M. Khalid, and R. Sankar, "Impact of mobility prediction on the performance of cognitive radio networks," *In Proc. of IEEE Wireless Telecommunications Symposium (WTS), Tampa, FL, USA*, pp. 1–5, April 2010.

[89] H. Si, Y. Wang, J. Yuan, and X. Shan, "Mobility prediction in cellular network using hidden markov model," *In Proc. of 7th IEEE CCNC, Las Vegas, NV, USA*, pp. 1–5, Jan. 2010.

[90] H. Abu-Ghazaleh and A. Alfa, "Application of Mobility Prediction in Wireless Networks Using Markov Renewal Theory," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 788 –802, Feb. 2010.

[91] F. Baskett, K. M. Chandy, R. R. Muntz, and F. G. Palacios, "Open, Closed, and Mixed Networks of Queues with Different Classes of Customers," *Journal of the ACM*, vol. 22, pp. 248–260, April 1975.

[92] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and other Kernel-based Learning Methods*. Cambridge University Press, 2000.