Exploiting User Contention to Optimize Proactive Resource Allocation in Future Networks

Vishnu Vardhan Paranthaman
Faculty of Science & Technology
Middlesex University, London

A thesis submitted for the degree of

Doctor of Philosophy

January 2018

I would like to dedicate this thesis to my loving parents, my brother,
and in memory of my grandparents!

Acknowledgements

First and foremost, I am grateful to God for providing me with this opportunity and granting me the capability to proceed successfully. I would like to express my sincere appreciation and gratitude to the following people for helping me complete this thesis.

I am deeply grateful to Dr. Glenford Mapp, my director of studies and supervisor, for giving me the opportunity to work under him. He has guided me and encouraged me to carry on through these years and has contributed to this thesis with a significant impact. I am forever indebted for his enthusiasm, guidance and unrelenting support throughout this process. At many stages, during this research project, I benefitted from his advice, particularly so when exploring innovative ideas. Dr. Mapp has been very supportive and gave me the freedom to pursue various projects without objection. My gratitude for his contribution to my future career is immeasurable.

I would also like to thank my other academic supervisors, Dr. Purav Shah and Dr. Huan X. Nguyen for their invaluable insights and suggestions. Special word of thanks to Dr. Yonal Kirsal who has always given me a hand by spending his valuable time in exploring different ideas and concepts.

I would like to thank Prof. Richard Comley for giving me this wonderful opportunity of pursuing my Ph.D. from the Middlesex University, Terri Demetriou, George Constantinou, and Alex Fogden for providing me administrative support and taking the time in addressing all my queries, and finally the administrative staff at the School of Science and Technology for providing me with all the required facilities and arrangements for doing my research and travelling for conferences.

Middlesex University has provided me with a very stimulating environment in what concerns the extraordinary quality of its academic staff, and that experience will leave a mark beyond this thesis.

I extend my thanks and gratitude to my friends who have always been a significant source of support and encouragement when things would get a bit disappointing: Nishanth Singh, Kamran Ali, Dr. Chennam Vijay Venkatesh, Dr. Arindham Ghosh, Turki Aljrees, Farzad Aria, Dr. Nimai Parmar, and Dr. Yetish Joshi. Thank you, guys, for always being there for me. A special word of thanks goes to all my other friends and colleagues who have provided me with continuous support and encouragement throughout my Ph.D. journey.

This journey would not have been possible without the support of my extended family, Dr. R.S Raajanathan, Dr. Thanusha Selvadurai Raajanathan, Mr. Raj Selvadurai, Mrs. Thayalini Selvadurai, Mr. Suthakaran Mrs. Thushitha Selvadurai Suthakaran, Mr. Ravi Sounthararajan, Mrs. Sunthari Sounthararajan, Shushaanth Sounthararajan, Ashaanth Sounthararajan, Sanjeet Raj, Sanjeev Raj, my nephews Seiyyon Raajanathan, Swaran Suthakaran, and my adorable niece Thara Raajanathan. Thank you all for the constant unconditional support and encouragement you all provided me during this journey.

I would like to dedicate this thesis to my father Paranthaman Narayanaswamy and my mother, Girija Paranthaman. This accomplishment would not have been possible without them. Thank you for always supporting me and believing me. Thank you for teaching me respect, confidence and proper etiquette. I will never truly be able to express my sincere appreciation to the both of you. I would like to express my heartfelt thanks and gratitude to my brother Dr. Ashwan Paranthaman for the excellent example he sets, for always listening when I wanted to talk, for the invaluable support and concern he had for me, my sister-in-law Karthika Raman, and finally my adorable nephews Mouvian Narayanaswamy and Pranav Narayanaswamy.

Finally, to Anagha Lakshmi Jayaprakash, my partner there are no words to express how much you have helped me during these years. Without you, this would have been merely impossible. But the most important part of our chapters are the ones we are yet to write.

Abstract

In order to provide ubiquitous communication, seamless connectivity is now required in all environments including highly mobile networks. By using vertical handover techniques it is possible to provide uninterrupted communication as connections are dynamically switched between wireless networks as users move around. However, in a highly mobile environment, traditional reactive approaches to handover are inadequate. Therefore, proactive handover techniques, in which mobile nodes attempt to determine the best time and place to handover to local networks, are actively being investigated in the context of next generation mobile networks. The Y-Comm Framework which looks at proactive handover techniques has defined two key parameters: Time Before Handover and the Network Dwell Time, for any given network topology. Using this approach, it is possible to enhance resource management in common networks using probabilistic mechanisms because it is now possible to express contention for resources in terms of: No Contention, Partial Contention and Full Contention. As network resources are shared between many users, resource management must be a key part of any communication system as it is needed to provide seamless communication and to ensure that applications and servers receive their required Quality-of-Service. In this thesis, the contention for channel resources being allocated to mobile nodes is analysed. The work presents a new methodology to support proactive resource allocation for emerging future networks such as Vehicular Ad-Hoc Networks (VANETs) by allowing us to calculate the probability of contention based on user demand of network resources. These results are verified using simulation. In addition, this proactive approach is further enhanced by the use of a contention queue to detect contention *between incoming requests and those waiting for service. This thesis also presents a new methodology to support proactive resource allocation for future networks such as Vehicular Ad-Hoc Networks. The proposed approach has been applied to a vehicular testbed and results are presented that show that this approach can improve overall network performance in mobile heterogeneous environments. The results show that the analysis of user contention does provide a proactive mechanism to improve the performance of resource allocation in mobile networks.*

Contents

List of Figures

List of Tables

Notations

5G	5th generation mobile network
AP	Access Point
BS	Base Station
BSM	Basic Safety Message
CAC	Call Admission Control
CAM	Cooperative Awareness Message
CINR	Carrier to Interference Noise Ratio
D2D	Device to Device
D2I	Device to Infrastructure
DSRC	Dedicated Short Range Communication
EDGE	Enhanced Data for GSM Evolution
ETSI	European Telecommunications Standards Institute
FIFO	First In First Out
FSPL	Free Space Path Loss
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile
ICA	Intersection Collision Alert
IP	Internet Protocol
IPv6	Internet Protocol version 6
LAN	Local Area Network
LIFO	Last In First Out
LoS	Line of Sight
LTE	Long Term Evolution
MAC	Media Access Control

MATLAB	Matrix Laboratory
MN	Mobile Node
MT	Mobile Terminal
NDD	Network Dwell Distance
NDT	Network Dwell Time
OMNeT++	Objective Modular Testbed in C++
PHY	Physical Layer
PoA	Point of Attachment
PVD	Probe Vehicle Data
QoE	Quality of Experience
QoS	Quality of Service
RAT	Radio Access Technology
RSA	Road Side Alert
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RSU	Road-side Unit
SINR	Signal to Interference Noise Ratio
SUMO	Simulation for Urban Mobility
TBH	Time Before Handover
TCP	Transmission Control Protocol
TfL	Transport for London
UDP	User Datagram Protocol
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
V2I	Vehicle to Infrastructure
V2P	Vehicle to Pedestrian
V2V	Vehicle to Vehicle
VANET	Vehicular Ad-hoc Network
Veins	Vehicles in Network Simulation
VPN	Virtual Private Network
WAVE	Wireless Access in Vehicular Environments
WLAN	Wireless Local Area Network
WSMP	WAVE Short Messaging Protocol

T_{EH}	Time taken to execute a handover
R_H	Handover radius
ν	Velocity of the mobile node
R_E	Exit radius
Υ	Time before handover
$\aleph$	Network dwell time
$\hbar$	Time to handover
$\mathbb{T}$	Time to get resource
$\mathbb{N}$	Resource hold time
ϱ	Handover prepare time
N	Number of nodes
$\mathbb{Z}$	Delay due to contention
t_c	Current time
t_{det}	Detection Time
t_{con}	Configuration time
t_{reg}	Registration time
t_{adp}	Adaptation time
τ	Reciprocal of $\mathbb{T}$
η	Reciprocal of $\mathbb{N}$
λ	Total arrival rate
λ_{eff}	Effective arrival rate
μ_s	Service rate
S	Number of server/channel
Q	Maximum number of request in queue
U	Total number of users in the system
μ_m	Mobility rate
μ_i	Service rate for state i
ρ	Traffic intensity
P_i	State probability
γ	Throughput
α	Full contention among arrivals around same time
β	Requests leaving the queue due to Full contention
θ	Adjustment/swapping of requests in queue due to contention

Chapter 1

Introduction

We are rapidly moving towards a world in which mobile systems will be operated using the 4A's paradigm of seamless communication: Anytime, Anywhere, Anything and Anyhow. This includes support for seamless connectivity in highly mobile environments. With the rapid development of mobile communication technologies such as WiFi, femtocells, Long-Term Evolution (LTE), Vehicular Ad-Hoc Network (VANET), the integration of various other wireless networks known as Heterogeneous Networking (HetNet) has become necessary to provide Mobile Nodes (MNs) with ubiquitous communication. For example, where, an MN can connect to a nearby LTE, the calls rejected by LTE networks due to lack of radio access can overflow to overlaying WiFi networks, thus reducing the call blocking probability and improving bandwidth utilization in cellular overlay networks. However, this may result in frequent forced handoffs in those areas covered by small cells and leads to extra signalling overheads (Huang et al., 2011). In these environments, traditional handover techniques, which depend on a reactive approach, have been found to be inadequate because of high speeds of vehicles as resources must be quickly allocated and deallocated as the mobile user moves around. Hence, good resource management must be considered as a key enabling functionality to allow seamless connectivity in mobile environment.

This issue needs to be addressed in future network requirements such as 5th generation mobile networks (5G). The 5G standardisation framework will be defined by 2020, and 5G architecture is expected to accommodate a wide range of use cases from the most important vertical sectors namely: Automotive, Energy, Industry 4.0, Health, Media, and Entertainment. These use cases have more demanding network requirements, especially in terms of latency, bandwidth, coverage, and resilience. According to the 5G Infrastructure Public Private Partnership (5G PPP), the most important performance targets that 5G needs to achieve are: latency should be below 5ms, and device density should be up to 100 $devices/m^2$, along with tight constraints on territory and population coverage for supporting all possible services. 5G will also integrate different enabling technologies and networks leading to a heterogeneous environment in which the MN switches seamlessly between networks while maintaining the required Quality of Service (QoS) for its applications (Olwal et al., 2016). Furthermore, it is important that 5G is very efficient in terms of resource allocation because 5G will be deployed on a global scale.

The Internet is currently evolving. Instead of large, global, but individually-managed networks, a core network is being deployed which is fast and getting faster by the use of optical switching. Peripheral wireless networks will be situated at its edges. A core endpoint is an entity which is at the edge of the core network and is used to connect different types of peripheral networks to the core infrastructure as shown in Figure 1.1. In order to ensure consistent and seamless interoperability as the MN moves around, it is crucial to have the ability to anticipate future network conditions with adequate precision. Therefore, one of the main challenges is handover and how to manage the mobility of MN as well as its effects on resource allocation.

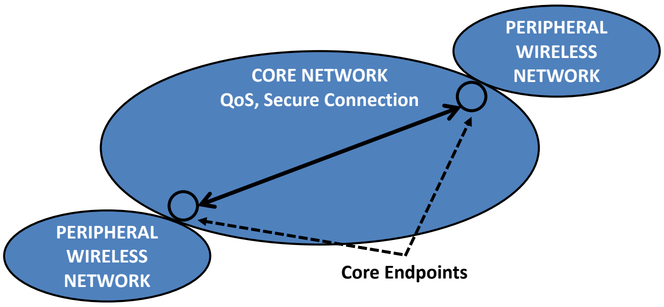


Figure 1.1: Core and Peripheral Networks (Mapp et al., 2016)

Proactive behaviour by systems refers to anticipatory, self-initiated and change-oriented behaviour in different situations. Proactive behaviour involves acting in advance of a future situation, rather than just reacting. It means taking control and making things happen rather than just adjusting to a situation or waiting for something to happen. Therefore, this proactive approach helps any system to perform better than a traditional reactive approach. This applies to wireless networks where the handover and resource allocation currently follow a reactive approach which can be enhanced by using a proactive approach (Saxena and Roy, 2009). It is very important to provide a seamless service to all the users in a network and therefore, these users and their contention for resources have to be looked in detail.

Any proposed solution must also be scalable because in the future the MN may have the ability to handover to hundreds of possible target networks (Ulvan et al., 2010). Knowing the velocity and current position of an MN could help to estimate where the MN is heading. Thus, the next position of MN where handover might be performed can be predicted (Mapp et al., 2016). Hence, using this approach, proactive handover in which the MN actively attempts to decide where and when to handover has been shown to be an efficient handover policy mechanism to minimize packet loss and service disruption as an impending handover can be signalled to the higher layers of the network protocol stack (Mapp et al., 2012). This approach has been shown to be effective in mobile environments (Ghosh, 2016).

In addition to the need for new handover decision mechanisms, there is also a need for better resource reservation mechanisms due to varying traffic characteristics, QoS application requirements and wireless channel conditions at the access point. Efficient resource management is needed to optimize the performance of a wireless network because only a limited number of simultaneous calls can be hosted by a wireless cell. Therefore, incoming handoff calls and new calls should not compromise the quality of the ongoing calls in the cell. Traditionally, user contention has been used to analyse the need for specific resources such as radio channels by mobile users. However, it is possible to use this contention to proactively manage these resources. This thesis explores the use of contention to provide better resource management in highly mobile networks.

One key application area of proactive resource allocation is in Intelligent Transport Systems (ITS) which can be implemented using VANETs. Characteristics of VANET such as high velocity, smaller coverage range and mobility pattern are serious challenges in providing seamless handover, resource allocation and in moving the services from the previous Road-side Unit (RSU) to the new RSU. HetNet for vehicular communication, comprising VANET and LTE (Ucar et al., 2015) have been considered in Cooperative ITS (C-ITS). Vehicular users experience strong data rate fluctuations as their locations change rapidly (Abou-zeid et al., 2013). Besides, available resources are limited in a cell must be used by new calls and handover calls. In addition, the radio resource scarcity and the increasing number and capacity needs of mobile users make it mandatory to utilize the available resources in an efficient way (Boujelben et al., 2014). Therefore, developing proactive handover and resource allocation models for high speed mobile systems would be the best option to achieve a realistic model which can be applied to most of the communication systems.

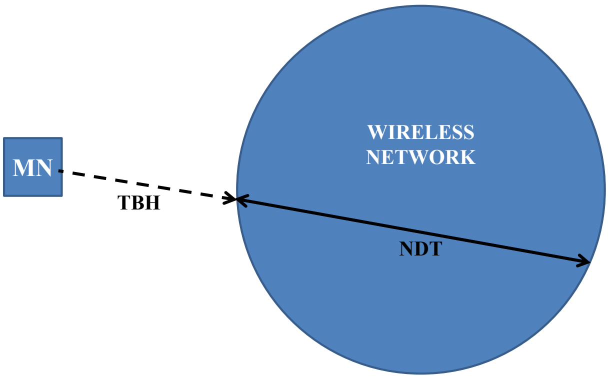


Figure 1.2: Illustrating Time Before Handover & Network Dwell Time (Mapp et al., 2012)

Vertical handover techniques are used to maintain the connection between the MN and the core network via the core endpoint as the user moves around (Mapp et al., 2016). Y-Comm is an architecture that has been designed to build future mobile networks by integrating communications, mobility, QoS, and security. It accomplishes this by dividing the future internet into two frameworks: Core and Peripheral Frameworks. The researchers of Y-Comm have made major contributions in the areas of proactive handover to provide seamless communication by introducing the ability to accurately estimate Time Before Handover (TBH) which is the time after which the handover should occur and Network Dwell Time

(NDT) which is the time the MN will spend in the coverage of the new network as shown in Figure 1.2. These two parameters were used to study seamless handover in mobile environment.

In this thesis, a detailed investigation has been presented on how these two parameters can be used to aid the management of resources in a proactive handover scenario by analysing the contention between mobile users for communication channels in wireless networks, hence, leading to better resource management. These two parameters allow us to determine the time when different nodes will need to acquire and release resources due to mobility. Therefore, it is possible to explore periods of contention for resources which, in turn, will allow us to develop heuristic proactive algorithms to optimise the use of the resources.

1.1 Research Aims and Objectives

The main aim of this thesis is to address the issues highlighted above by looking at developing mechanisms for better resource management to ensure seamless communication in highly mobile environment such as vehicular networks. The primary focus is to study the effects proactive resource allocation in a wireless network. Models are developed to calculate the probability of an MN acquiring resources before it reaches the next coverage network based on different contention types. Further, the application of this approach to a real VANET has been presented.

Considering the scope above, the main research question of this thesis will be outlined in the following subsection.

1.1.1 Key Research Questions

This thesis looks at exploiting user contention to optimize proactive resource allocation in future networks. Some of the key Research Questions that have to be addressed in this thesis are:

1. What are the time parameters that is required to analyse and understand the contention for resources in a highly mobile environment?

2. What are the different types of contention that occur when acquiring a resource based on user mobility?
3. What are the different mathematical conditions that will identify the occurrence of contention?
4. How can these conditions be used to develop an analytical model to calculate the probability of contention for a resource among different MNs?
5. What are the aspects that have to be considered to develop a queuing model for a proactive resource allocation and how well does this approach perform when compared to the classical resource allocation model?
6. How do we build a VANET testbed and acquire the necessary parameters in order to apply the proposed approach to VANETs?

1.1.2 Solution Approach

This thesis provides a more comprehensive analysis of contention for resources. Some of the concepts of Y-Comm architecture such as NDT, TBH and Exit Times have been used to provide a framework to investigate resource allocation issues. Y-Comm has been used to study seamless handover in both homogeneous and heterogeneous networks. This thesis, also explains the overall approach by showing the application to VANET testbed and articulates that in highly mobile environments, it is necessary to consider a proactive approach to handover and resource allocation model which is based on a probabilistic rather than a fixed coverage approach. Furthermore, the developed proactive model is compared with traditional resource allocation approach. Finally, this thesis presents the results from analytical model, simulation and the testbed data collected from the VANET Testbed.

The key contributions of this thesis are therefore as follows:

- A proactive approach has been formulated by analysing the contention among various users in trying to acquire a radio channel in a wireless cell.

- Using probabilistic techniques, the contention among various mobile users in trying to acquire a communication channel in a wireless cell has been analysed. We find the probabilities of No Contention, Partial Contention and Full Contention.
- Two new queuing models are introduced: the first calculates the probability that a MN will never acquire a channel amongst various simultaneous requests for the channel and hence, the MN can be immediately instructed to be handed over to another network.
- A second case proposes a further refinement by introducing a concept of a contention queue which is used to analyse users waiting to acquire the channel before they reach the coverage of the next network.
- Using simulation, this work has shown that these two new approaches significantly improve the overall system performance compared to reactive handover multi-channel queueing models in terms of mean response time and throughput.
- It has been demonstrated that how these new parameters can be calculated for any wireless network.
- Application of this approach to a VANET network has been presented.
- Preliminary results show that this approach can be effective in improving the overall network performance.

1.2 Thesis Outline

The following is an outline of the final thesis:

- Chapter 2: the literature review presents a critical review of the existing solutions and approaches produced by researchers, scientists and groups.
- Chapter 3: details the methods and approaches used for conducting this research. The methods comprises of simulation using matrix laboratory

(MATLAB) tool, analytical modelling using exponential probability distribution function (PDFs) and queueing theory and finally, the proposed approach is applied in the VANET testbed.

- Chapter 4: a more comprehensive analysis of contention and their effects in wired network such as Ethernet for the traditional network access protocols such as Pure ALOHA and Slotted ALOHA were presented. Also, elaborates the contention for resources by MNs in a classical reactive handover scenario. This chapter also details the queueing analysis of the classical approach and highlights the need for better approach.
- Chapter 5: introduces the proactive system and also describes Y-Comm framework. Proactive handover and coverage segmentations of a wireless network is presented. Different contention types involved in acquiring the resources are described.
- Chapter 6: The analytical model to calculate the probability of contention for two and three nodes scenario is presented in this chapter.
- Chapter 7: details the queueing analysis of the two different proactive approach and the results are presented comparing the classical approach.
- Chapter 8: discusses the VANET Testbed in detail and shows the application of proactive approach.
- Chapter 9: concludes this thesis with a summary of the thesis and directions for future work, in order to ensure continual improvement in the current and related field of study.

Chapter 2

Literature Review

Because of the emergence of Smart Cities, there is now an attempt to support seamless communication in highly mobile environments. Several research efforts were carried out looking at routing, security and applications for highly mobile environment but very few addressed handover and resource allocation issues. Recent endeavours in (Almulla et al., 2014) and (Li et al., 2014) clearly show that researchers are interested in proactive handover or predictive handover mechanisms, however these efforts considered parameters like user preferences, user location and application requirements and thus used techniques such as proactive caching but failed to analyse the effects of the lower layers on these parameters as highlighted in (Ghosh et al., 2013, 2014a). This Chapter presents an extensive analysis of various research efforts that looked into handover and resource allocation for mobile networks. Before looking into existing work, a comprehensive understanding of different handover types is necessary which is as shown in the following section.

2.1 Advanced Handover Classification

Handovers in a wireless network can be divided into two advanced type as shown in Figure 2.1. Imperative handovers occur when the MN changes its Point-of-Attachment (PoA) because it has determined by technical analysis based on parameters such as coverage, signal strength, and the QoS offered by the new network. There may be a severe loss of performance or loss of connection if they are not performed and therefore, it is called as an imperative handover. In contrast, alternative handovers occur due to reasons other than technical issues such as preference for a given network based on price or incentives. Hence there is no severe loss of performance or loss of connection if an alternative handover does not occur (Cottingham, 2009).

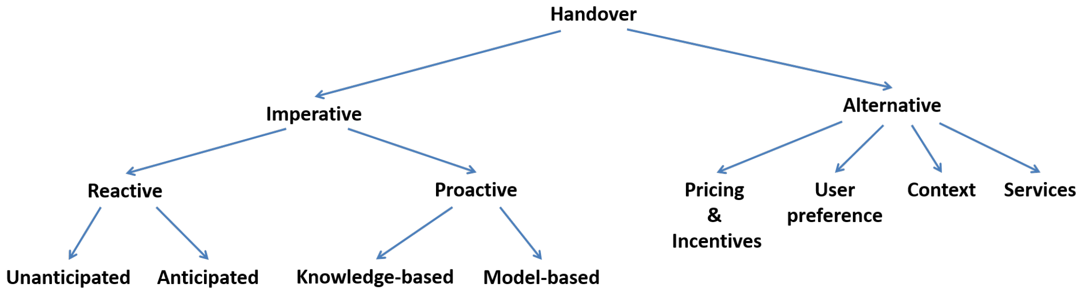


Figure 2.1: Handover Classification (Cottingham, 2009)

Imperative handovers are, in turn, divided into two types that is reactive and proactive handover. As explained previously reactive handover responds to changes in the low-level wireless interfaces as to the availability or non-availability of certain networks and it can be further divided into unanticipated and anticipated handover. Unanticipated handovers are hard handovers where the MN is heading out of range of the current PoA and there is no other access point (AP) or base station (BS) to which to handover. Anticipated handovers are soft handovers which describe the situation where there are alternative BS to which the MN may handover. Proactive handover policies which uses soft handover attempt to know the condition of the various networks at a specific location before the MN reaches that location (Cottingham, 2009).

Proactive policies allow MNs to calculate the time before handover which enables them to minimize packet loss and latency experienced during handovers. Presently, two types of proactive handovers are being explored by various researchers. The first is knowledge based where the signal strengths of available wireless networks is known by measuring beforehand for a given area such as a city. This could involve physically driving around and taking these readings. Due to the effects of seasonal changes on wireless propagation these measurements need to be taken in different seasonal contexts. The second is a proactive policy based on a mathematical model which calculates the point when handover should occur and the time that the MN would take to reach that point based on its direction and velocity. The accuracy of this approach is dependent on various factors including location, network topology, the propagation model used, technology, and specific environments, for example, whether the MN is indoor or outdoor (Mapp et al., 2012).

2.2 Handover Techniques in Highly Mobile Environment

A lot of research has been done to investigate handover in mobile heterogeneous environments in order to guarantee the delivery of acceptable QoS in such environments. A joint optimization was proposed in (Bi et al., 2014) to solve the handover problem in heterogeneous networks, especially in VANETs and cellular technologies, to keep a balanced load across all access points (APs) and to maximize the data rate of overall networks. In addition, in (Kirsal, 2016) an intelligent handover decision approach was proposed to minimize the handover failures and unnecessary handovers whilst maximizing the usage of resources. However, the limitations of current reactive networks were highlighted in (Bastug et al., 2014).

Several studies of the vertical handover procedure, mobility management and common radio resource management schemes in heterogeneous environments have been reported in the literature (Ben Elhadj et al., 2016; Seaman et al., 2012; Vilaplana et al., 2014; Xenakis et al., 2016). These studies show that there are several dynamic factors that must be considered in vertical handover decisions for effective network usage including policies to determine whether or not to handover as well as mechanisms to determine the best network to which to handover. In order to guarantee the delivery of acceptable QoS in such environments; a decision has to be taken on whether or not a handover should be executed based on many factors such as the speed of the user, network conditions, response time of the networks. Service and buffer management issues were explored in (Gemikonakli et al., 2013) to provide integrated voice and data services in single and multi channel wireless networks.

The authors in (Zhang et al., 2011) have attempted to address handovers from macro cells to femto cells for high speed nodes, which may encounter useless handovers as well as increasing signalling overhead based on the mobility state of node. Work in (Hossain et al., 2015) showed the advantages of deploying femtocells and the performance enhancement provided by macrocell/picocell/femtocell integration. The maxmium velocity considered for the mobile user was 36km/h and therefore, did not show the effects of higher velocity on the call dropping probability. Hence, the mobile users moving out of the cell without service due to high mobility was not captured. In (Ulvan et al., 2010) the authors proposed a client-based handover decision procedure, where the MN periodically sends its position relative to the serving cell which maintains a database of all possible target cells to where the handover might be performed. The likely path of the MN was estimated in advance by a probabilistic approach using the remaining time before handover and the handover probability was derived. The authors in (Patanapongpibul et al., 2006) have also proposed a client-based handover for MN in Mobile Internet Protocol version 6 (IPv6) wireless networks supporting both horizontal and vertical handoffs.

Proactive handover policies attempt to know the condition of the various networks at a specific location before the MN reaches that location (Almulla et al., 2014). In addition to this, the Mobile IPv6 Fast Handover (FMIPv6) operates in two modes: Predictive and Reactive. In a Predictive Fast Handover, a MN is able to send an Fast Binding Update (FBU) to the New Access Router (NAR) when it is attached to the Previous Access Router (PAR) even before the MN attaches to the NAR, which then establishes forwarding for its traffic (Schmidt et al., 2014).

An IEEE 802.11 AP driven handover algorithm using network virtualization and high level abstractions provided by the 5G-EmPOWER framework was presented in (Zeljkovic et al., 2017). Proactive and transparent handover steered by a centralized controller which uses RSSI values to estimate the distance of the user from the AP was proposed to reduce the amount of handover and to avoid unnecessary one. The study in (Feirer and Sauter, 2017) shown that a proactive process is necessary to achieve a proper and seamless roaming performance in a dense wireless environment using a IEEE 802.11k amendment. The shortcoming of this approach is that the number of channel scans needed increases as the number of available APs increases. Transport layer handoff issues were highlighted in (Sinky et al., 2015) and the authors proposed a handoff-aware cross-layer assisted multi-path TCP in (Sinky et al., 2016) to improve service continuity during handovers through proactive congestion window adjustments in mobile heterogeneous wireless networks.

A proactive unnecessary handover avoidance scheme was proposed in (Wang et al., 2014) for the LTE-A small cells. Unnecessary handover was avoided by calculating the probability of the active time during the dwell time in one cell and comparing the same to find if pre-defined threshold exceeds certain value. Here, a model to estimate the dwell time in the small cell was not considered. To reduce handoff call dropping rate and maintain acceptable new call blocking rate while providing efficient bandwidth utilization, a distributed bandwidth reservation scheme called the mobility-prediction-aware bandwidth reservation (MPBR) scheme was proposed in (Nadembega et al., 2015a). MPBR consists of a handoff time estimation scheme, an available bandwidth estimation scheme and an efficient call admission control scheme. The results indicated that the proposed scheme had an increased probability of providing acceptable QoS to mobile users.

2.3 Prediction: Mobility and Network Connection Time

In (Yan et al., 2010) a survey of vertical handover decision algorithms in fourth generation heterogeneous wireless networks was presented. Twelve different algorithms were evaluated and grouped based on the criteria they use for making handover decisions. Three algorithms among the twelve used the Network Connection Time estimation technique which is the duration that a Mobile Terminal (MT) remains connected to a PoA. MT lifetime or travel time within the WLAN was proposed in (Zahran et al., 2006) and (Yan et al., 2008). The shortest distance between the point at which handover is initiated and WLAN boundary was presented in (Mohanty and Akyildiz, 2006).

The work in (Gomes et al., 2016) has highlighted the importance of mobility prediction for the cloudified mobile networks in the migration of content-caches located at the edge of network which optimizes their resources and achieves significant bandwidth savings. The authors in (Becvar et al., 2014) have proposed a prediction algorithm for the number of resources required by the users at the target cell after handover. Connectivity management as a service (CMaaS), a new control plane for 5G network architecture was proposed and the performances of reactive and proactive CMaaS for Device to Infrastructure (D2I) links, and reactive CMaaS for the Device to Device (D2D) links were studied. According to the study the proactive CMaaS achieved less than 3 ms round-trip time (RTT) delays but comes at the expense of an increasingly complicated control plane as a larger number of rules need to be computed and communicated by the controller every second compared to reactive CMaaS (Yazc et al., 2014).

Since, mobility results in channel gains which varies with time, Abou-zeid et al. (2013) have shown that the data rates experienced by the mobile user in the future can be predicted by exploiting the user mobility patterns with the correlation of location and received signal strength. The work has presented a long-term predictive resource allocation plans for the BS's by exploiting the radio maps and mobility pattern.

There are several approaches through which Resource Reservation can be implemented. These solutions are QoS routing, Medium Access Control (MAC) scheduling mechanisms and admission control. In the case of high traffic load, there is no admission control available for rejecting traffic so as to support QoS for the existing traffic (Yu et al., 2013). According to (Yu et al., 2013), MAC scheduling mechanisms are essential for managing bandwidth allocation and implementing service differentiation towards distinct types of sessions. The impact of mobility on the infra-structureless IEEE 802.11p MAC performance by investigating the mobility factors such as velocity and number of neighbouring nodes were evaluated in (Alasmary and Zhuang, 2010). The work showed that the relative speed of the MN has a significant impact on channel access at the MAC layer, disregarding the number of communicating nodes. This work was carried out to achieve better fairness in terms of how long each node shares the medium based on the estimated time it spends in the active transmission range.

A mobility prediction scheme for MNs was proposed in (Nadembega et al., 2015b) Probability and DempsterShafer processes were applied to predict the likelihood of the next destination based on the users habits (e.g., frequently visited locations). In addition, a second-order Markov chain process was applied at each road junction for predicting the likelihood of the next road segment transition, given the direction to the destination and the path from the trip origin to that specific road junction. Terminal mobility prediction for ultra dense networks based on Support Vector Machine was proposed in (Challita et al., 2017). A Support Vector Machine was used to obtain the vector index of the mobile terminal and regression algorithms to predict the location of the mobile user and based on the prediction information the proposed scheme make reservations of the needed resources in advance. The approach proposed in (Fazio et al., 2014) showed how a statistical approach based on mobility prediction technique can enhance system performance using In-Advance Multiplexing Call Admission Control (IAM-CAC) scheme for cellular networks.

2.4 Cooperation in Wireless Networks

Cooperative techniques in wireless networks have been recently proposed and are aimed to enhance the efficiency of data dissemination and resource allocation. Potential benefits of cooperation and its challenging issues in wireless communication networks were discussed in (Zhuang and Ismail, 2012). Three cooperation scenarios were examined, namely, cooperation to improve channel reliability through spatial diversity, cooperation to improve throughput through resource aggregation, and cooperation to achieve seamless service provision. The study highlighted that cooperative decision making should be based on cross-layer design among transport, network, MAC and Physical (PHY) layers.

Utilization of roadside wireless local area networks (RS-WLANs) as a network infrastructure was investigated in (Liang and Zhuang, 2012) and shown that cooperative caching at the AP consume more bandwidth and buffer storage resources at the AP, than direct transmission. This demonstrates that in order to make the best use of techniques such as caching and storage it is necessary to know the exact amount of data which can be transmitted to the user in a given time. This helps in reducing the usage for buffer space. However, according to (Ismail and Zhuang, 2012) BS/AP can perform its own resource allocation to support the MT, while at the same time cooperate with other available networks BSs/APs, hence, no central resource manager is required.

2.5 Call Admission Control

Call Admission Control (CAC) is an important part of resource management and hence, a lot of work has been used to analyse CAC schemes, for example, a CAC scheme using a cross layer approach to adapt the transmission power in PHY layer and to optimize the contention window size in the MAC layer on the basis of the vehicle density estimation was proposed in (Bejaoui, 2012, 2014). A priority level computing function was also introduced which is based on vehicular parameters to make the scheduling and resource allocation process fairer. Security, Link Quality, Roadside conditions, Precedence class and waiting times are the network parameters considered in the proposed function.

Carvalho et. al. (Carvalho et al., 2013) proposed a Joint CAC (JCAC) scheme which considers the ratio between the radius of the co-located Radio Access Technologies (RATs) and selects the smallest one for non-real-time service classes and the biggest RAT for real-time service classes. A joint optimization was proposed in (Bi et al., 2014) to solve the handover problem in heterogeneous networks, in particular VANETs and cellular technologies, to keep a balanced load across all access point and to maximize the data rate of overall networks. A mobility-aware CAC algorithm with a first-in-first-out (FIFO) Handoff Queue (HQ) to accommodate more handoff calls if no channels are available and if HQ is not full in mobile hotspots was proposed in (Kim et al., 2013)

Self-Organizing Network (SON) is a new approach, which has self-configuration, self-optimization and self-healing functionalities for minimizing human efforts. A self-configuration algorithm with three self-optimization functions: Admission Control, Handover and Load Balancing for newly added Femto Cell has been proposed in (Boujelben et al., 2014). An algorithm for proactive resource reservation has been proposed in (Ukil and Sen, 2010) based on teletraffic theory and the concepts of adaptive filtering. The limitations of current reactive networks has been highlighted in (Bastug et al., 2014) and the authors proposed a proactive caching networking paradigm where predictable users' peak data traffic demands can be reduced by proactively serving.

Congestion avoidance, which is enforced using CAC algorithms, is also important to avoid degradation in the quality and performance of a network. The question of whether to accept a new call, based on whether this new call can be supported with the desired QoS is dealt by CAC schemes (Bashar et al., 2010). An adaptive JCAC model was proposed in (Kaur and Selvamuthu, 2014) for balancing loads among LTE-UMTS networks. Along with considering the service requirements of an user, the proposed model selects an interface for a new call based on the load conditions of LTE and UMTS interfaces. CAC based on service demands to assign proper network in LTE networks using femtocells was proposed in (Khan et al., 2016). In dense environments the proposed approach produces undesirable amounts of link delay and signalling overhead.

A handoff algorithm based on CAC for heterogeneous networks was proposed in (Kabiri et al., 2016). The proposed algorithm selects a target BS from a list of neighbouring candidate BSs by taking into account multiple criteria such as measured received signal strength (RSS) and signal-to-interference-plus noise ratio (SINR), predicted RSS and SINR, and number of free resource blocks of neighbouring BSs. In addition, logistic smooth transition autoregressive model was used to predict the RSS and SINR samples. An optimized handoff scheme based on mobility prediction, signal strength, and availability of resources for femtocell networks called as OHMPCAC was presented in (Cheikh et al., 2016). The aim was to find the suitable AP and to reduce unnecessary handoffs.

2.6 Network Utility Maximisation and Selection

In 1998, Kelly et al. (Kelly et al., 1998), introduced the notion of Network Utility Maximization (NUM) and formulated resource allocation as an optimization problem for the first time. The authors assumed a wired network consisting of fixed capacity links and a set of users that wanted to transmit data to a set of destination nodes. The path of the traffic is known a priori and does not change during the optimization process. The main assumptions of this framework are that utility functions of the transmission rates are concave in nature and that all links have a fixed capacity.

This work was extended to look at inelastic flows where the utility function was linked to Quality of Experience (QoE) as a function of the bandwidth being delivered to applications such as HTTP, VoIP and IPTV (Liu et al., 2007). Such techniques have also been used to investigate multi-hop wireless networks (Chiang, 2005). This approach was further extended to look at networks with high SINR and derived utility functions for common applications in this environment (Tychogiorgos and Leung, 2014).

Though this approach has been effective, its usefulness is limited in mobile heterogeneous environments for two reasons. The first is handover in which users are continually being switched between different base-stations. The rate of handover is determined by the speed and path taken by the user and therefore has to be modelled for this framework. In (Trestian et al., 2017), the authors looked at resource allocation for multimedia applications in a network of LTE small cells. The work assumed that mobile devices have a Handover (HO) monitor application, which allows them to calculate Time Before Handover and Network Dwell time to ensure seamless connectivity (Shaikh et al., 2007b). So managing handover is now a key part of resource management in modern networks. The second issue is that heterogeneous environments may involve several networks with different qualities of service. This introduces network selection issues based on different criteria depending on if the handover is imperative or alternative (Mapp et al., 2009). In addition, in these environments, we need to simultaneously allocate resources fairly to users, and at the same time ensure the efficient management of several different networks. It should be realised that it may be difficult to optimize these two factors in the face of network selection by mobile users. In (Pervaiz and Bigham, 2009), the authors looked at network selection in competitive networks using a game-theory approach. This effort developed a network selection algorithm based on reputation, degradation, price and availability.

2.7 Proactive and Predictive Resource Allocation

Research into resource allocation in communications networks has a long and distinguished history. The aim of this research is to develop efficient resource algorithms to allocate resources so as to enhance network stability and to enforce a level of fairness amongst users. Knowing the velocity and current position of an MN could help to estimate where the MN is heading, thus the next position of the MN where handover might be performed can be predicted. Proactive handover in which the MN actively attempts to decide when and where to handover has been shown to be an efficient handover policy mechanism to minimize packet loss and service disruption. In addition, an impending handover can be signalled to the higher layers of the network protocol stack (Mapp et al., 2012).

The authors in (El Gamal et al., 2017) proposed a proactive networking paradigm where the network anticipates user demand for networking resources in advance and utilizes this predictive ability to reduce the peak to average ratio of the wireless traffic, and thus yielded savings in the required resources to guarantee certain QoS metrics. The system and method presented focused on the existing cellular architecture and involved the design and analysis of learning algorithms, incentive techniques and predictive resource allocation strategies to maximize the efficiency of proactive cellular networks. In addition, a hierarchical VANET architecture was proposed in (Zhao et al., 2018) which supports content caching at different layers. The authors used the vehicle dataset collected from a VANET testbed deployed in the city of Porto, Portugal. The proposed model supports prefetching mechanism assisted by vehicle mobility prediction. The predicted locations of the vehicles are used to pre-fetch users content before their explicit requests. The proposed mobility prediction solution in (Zhao et al., 2018) is a simple Markov chain-based model, which adaptively selects the first- or the second-order Markov chain model based on the available trace quality.

A user-selective resource allocation scheme was proposed in (Liu et al., 2016) to proactively avoid the problem of interference mitigation for Mobile to Mobile (M2M) communication in LTE networks. The proposed scheme manages the radio resource blocks assignment to the M2M pairs, M2M neighbours, and the far-away users with less interference. Authors in (Guo et al., 2016) investigated the proactive resource allocation method performance gain by exploiting the application level, network level and user level information. A resource allocation planning optimization problem was formulated for pre-downloading the files to be requested to users, which optimizes the transmission duration at each BS along the trajectory of multiple mobile users to minimize the maximal transmission completion time.

The problem of optimal proactive caching was studied in (Tadrous and Eryilmaz, 2016) and the authors suggested that the service provider performs proactive service decisions depending on the degree of uncertainty about future requests to minimize its expected convex cost over time while maintaining on-time delivery of requested content. The proposed approach did not focus on the high mobility and handover issues. Proactive Caching approach for Icnbased VANETs (PeRCeIVE) to place the right content at the right node in-time was proposed in (Grewe et al., 2016). The proposed approach showed that a directed placement of data will improve the performance of the network by distributing the content with a minimal number of replicas one-hop away from the consumer. The approach demonstrated a content distribution mechanism based on geo-location, the current velocity and the heading direction of a vehicle. With the service provider's ability to track and statistically predict future requests of its users, a proactive caching of the peak hour demand ahead during off-peak times was proposed in (Tadrous et al., 2016) to smooth out network traffic. The proposed approach did not consider the high mobility scenarios were statistical prediction might not yield required QoS.

According to (Albasheir and Kadoch, 2016) as more and more innovative mobile services are being introduced, the LTE evolved packet core (EPC) require more resources, and does not have the capabilities to properly utilize the unused bandwidth of the guaranteed bearer when the reserved bandwidth is not fully used by the mobile service, the unused guaranteed bandwidth is considered as wasted resources and consequently the whole LTE/EPC network efficiency gets affected. Therefore, an adaptive technique to enhance the resource reservation was proposed to analyze the ongoing mobile traffic usage, forecast the mobile service resource consumption, provide time-series model, identify the unused/wasted resources, and utilize those resources by other services. Authors in (Mari et al.,

2014) have shown how the system may benefit from elements of future knowledge, even statistical or partial. Flexibility in resource allocation with prediction-based strategies was found to improve the energy efficiency of delay-tolerant transmissions. A quota-based utilization balance algorithm (QUBA) design was proposed in (Chung et al., 2014) to ensure fairness of resource usage and maximize the overall utility of all users in a heterogeneous networking environment. The focus of the proposed approach was to statistically balance traffic loads among the dominant networks with high data rates and to regulate excess traffic loads of extremely heavy users to help in maintaining the fairness among all users.

2.8 Research Gap

Most of the existing work have focused on achieving fairness through techniques like cross layering and modifying the MAC layer. Call admissions and resource allocation were decided based on the cost functions and priority schemes. Though all the efforts discussed above have provided some useful results, what is clearly missing is a method of analysing contention of network resources. This is true in a highly mobile environment where networks resources must be quickly allocated to and de-allocated from mobile devices. Hence there are three possible outcomes: a MN gets the wireless channel and leaves the network after being served; the MN gets the wireless channel but leaves the network due to mobility; the MN leaves the network coverage due to mobility without being served. If these scenarios can be analysed before MNs reach the next wireless network, then it should be possible to signal to the MN that it will not be served and hence it should do a vertical handover to another network.

To the best of my knowledge, there is no work which has focused on assuring that a high speed MN gets its required resources or reserves the resource which will be utilised in an estimated time in a group of available heterogeneous networks based on its mobility. In all these efforts described above, no new analytical model of resource contention was presented, which is the focus of this research.

Recently, proactive handover mechanisms and estimating the network connection time was given importance but a full fledged proactive resource allocation have not been investigated in detail. Therefore, there is a need for new CAC algorithms based on the contention for resources by MNs entering the network. The ability to analyse this scenario in detail should lead to better resource management algorithms as it should reduce the contention for resources.

In this thesis this scenario is analysed using contention analysis to yield No Contention, Partial Contention and Full Contention. In addition, we show how the system can be further extended by adding a contention queue that looks at the contention between incoming requests and requests waiting in the contention queue to be served. Contention analysis has been applied to Pure Aloha and Slotted Aloha as well as to wired and wireless systems. Table 2.1 shows all the mechanisms discussed, the networks to which they can be applied and the goals that each mechanism can achieve. It shows that contention analysis can be applied to all the different types of networks that have been discussed and can help to achieve fairness as well as good overall networking performance.

Table 2.1: Comparison of Different Network Management Mechanisms

✓= Supported   N = Not Supported   P = Partially Supported

Mechanism	**Papers**	**Analytical Method**	**Networks**	**Fairness**	**Performance**
Network Utility Maximization	(Kelly et al., 1998), (Liu et al., 2007), (Chiang, 2005), (Trestian et al., 2017), (Tychogiorgos and Leung, 2014)	Utility Function	Wired + Wireless	✓	N
Network Selection	(Shaikh et al., 2007b), (Mapp et al., 2009), (Pervaiz and Bigham, 2009)	TBH or Game Theory	HetNet	P	N
Reactive Handover	(Bi et al., 2014), (Kirsal, 2016)	Markov Process	HetNet	N	✓
Proactive Handover	(Bastug et al., 2014), (Mapp et al., 2012) (Almulla et al., 2014)	Location or user preferences	HetNet	N	N
Advanced Proactive Handover	(Ghosh et al., 2013), (Ghosh et al., 2014a)	TBH and NDT	VANET	P	N
Predictive Techniques	(El Gamal et al., 2017), (Zhao et al., 2018)	Prefetching + Markov Process	VANET + Cellular	N	✓
Contention Analysis	Proposed Model	TBH, NDT + Markov Process	All Networks	✓	✓

Chapter 3

Research Methodology

In general there are three specific methods to address the issues highlighted in this research. The first is analytical modelling in which an analytical model is developed to capture the key parameters and to generate estimated solutions. The second method is the use of simulation to analyse different outcomes based on scenarios which capture the problems being explored. Finally, the third method is to build a testbed to perform experiments which would give insight into the problem. In this thesis all these three methods are used. Firstly, an analytical model will be explored and this will then be verified using simulation techniques. Finally, a testbed will be used to analyse different scenarios and to gain new insights into the issues.

3.1 Analytical Modelling

Analytical Modelling is a mathematical technique used for exploring, explaining, and making predictions about complex processes. Constructing these mathematical models is an extensive process were a number of parameters affecting the system have to be considered accurate in order to achieve a reliable model (Wittevrongel and Phung-Duc, 2016).

In this study, a tractable analytical model has been proposed to calculate the probability of contention for resources by MNs based on communication range of the AP and the mobility of mobile users in femtocell/macrocell networks.

Therefore, the exponential distribution has been considered, and this is due to the fact the proposed probability model has to take into account several variables and their iteration. Since this a complicated process, it is appropriate to consider a simple distribution first to understand the system rather than choosing a complex distribution. Also, to the best of knowledge of the author, there is no work which has done such an extensive analysis of contention for resources in a highly mobile environment.

Secondly in this study, queueing analysis has been conducted to showcase the benefits of using a proactive resource allocation and the application of the analytical model developed to the queueing system is presented. Also, critical operations involved in such proactive resource allocation queue management are explained in detail and how the relevant parameters are calculated is explained.

3.1.1 Exponential Distribution

The exponential distribution is one of the commonly used distributions to describe continuous random variables. It is often used to model lifetimes of products and times between random events called as interarrival times such as arrivals of customers in a queueing system or arrivals of orders (Balakrishnan, 1996). The distribution has one rate parameter, λ. If our random variable X follows an exponential distribution, then we say

$$X \sim Exp(\lambda) \tag{3.1}$$

Its probability density function (PDF) is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ \\ 0 & otherwise, \end{cases} \tag{3.2}$$

the resultant PDF graph is shown in Figure 3.1 and cumulative probabilities can be calculated using the following equation.

$$P(X \leq x) = P(X < x) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 \\ \\ 0 & otherwise. \end{cases} \tag{3.3}$$

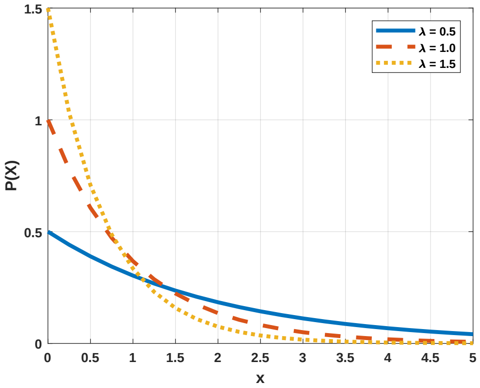


Figure 3.1: Exponential Probability Density Function

The main features of this distribution are:

1. The random variable can only take positive values.
2. Exponential decay - therefore larger values are unlikely.
3. The rate of arrival is fixed by λ and hence, larger values correspond to more rapid decay.

One of the primary uses of the exponential distribution is as a model for the times between events occurring randomly in time. The probabilities for the times between the events are described by the exponential distribution and the probabilities for the number of events taking place in a given period are described by the Poisson distribution. Both of these events occur randomly in time at a constant average rate λ. This is known as a Poisson process (Balakrishnan, 1996) and the Poisson probability is as shown below:

$$P(x;\ \mu) = \frac{(e^{-\mu})(\mu^{x})}{x!} \tag{3.4}$$

where, x is the actual number of successes that result from the experiment, and μ is the mean number of successes.

3.1.2 Queuing Theory

Queueing theory deals with one of the most unpleasant experiences of life, waiting. Erlang was the first who examined congestion problems at the beginning of 20th century, and his work inspired mathematicians and engineers to deal with queueing problems using probabilistic methods. Queueing theory became a field of applied probability, and many of its results have been used in traffic engineering, operations research, reliability theory, computer science, telecommunication, just to mention some (Jain, 2015).

The service discipline and structure of service will tell us the number of servers and capacity, i.e., the maximum number of customers or jobs in the system including the ones that are currently being served. The service discipline determines how the next customer is selected for service. The most commonly used service disciplines are First In First Out (FIFO) in which who comes earlier leaves earlier, Last In First Out (LIFO) in which who comes later leaves earlier, Random Service (RS) where the customer is selected randomly, etc. The service and interarrival times of jobs are usually independent random variables (Jain, 2015).

The aim of analytical models based on queueing theory is to get the performance measures of the system which are probabilistic. The properties such as density function, distribution function, mean and variance of the following random variables: number of customers in the system, utilization of the server/s, busy time of a server, idle time of the server, number of waiting customers, waiting time of a customer and response time of a customer are explored using the queueing theory. The result heavily depends on the assumptions concerning the distribution of interarrival times, service times, number of servers, capacity and service discipline (Jain, 2015).

3.1.2.1 Kendall‘s Notation

A notation originated by Kendall to describe a queueing system is denoted as shown below:

$$A/B/m/K/n/D$$

where

- A: distribution function of the interarrival times
- B: distribution function of the service times,
- m: number of servers,
- K: capacity of the system, the maximum number of customers in the system including the one being serviced,
- n: population size, number of sources of customers,
- D: service discipline.

In this thesis, two types of queueing system are used as explained below:

- $M/M/1$: represents the queue length with infinite size in a system having a single server, where arrivals are determined by a Poisson process and job service times have an exponential distribution.
- $M/M/m/K$: represents the queue length with finite size (K) in a system having $'m'$ number of servers, where arrivals are determined by a Poisson process and job service times have an exponential distribution.

Queuing theory can be explored by modelling, measuring, and analysing the arrival times, wait times, and service times of queuing systems. To do this queueing simulation can be used.

3.1.2.2 Markov Chain Model

A Markov chain is a mathematical model named after A. A. Markov. It is of a random phenomenon evolving with time in a way that the future states are independent of the past and depend only on the present. The time can be discrete (i.e. the integers), continuous (i.e. the real numbers), or, more generally, a totally ordered set. It is not necessary to know how long the process has been in the current state and state time has a memoryless (exponential) distribution (Jain, 2015).

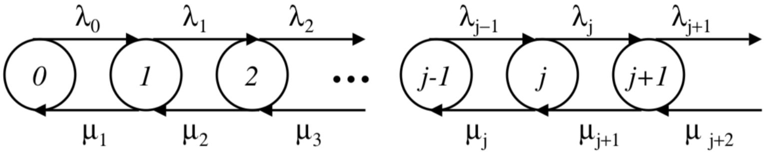


Figure 3.2: Markov Chain Model

The birth-death processes in a system can be represented as a Markov chain as shown in Figure 3.2 and it shows that the transitions are restricted only to the neighbouring states. The state of the system at n can only change to state $n+1$ or $n-1$ (Jain, 2015).

3.2 Simulation

Although we emphasize the fact that the analytical model that has to be constructed must be as realistic as possible, we have to understand that the analytical model itself is never more than a partial representation of reality. Hence, we have to investigate how well the model and its solution perform. Computer-aided techniques are therefore vital for engineers to evaluate performance and optimise designs in a timely, cost-effective and effort-free manner (Banks and Carson, 1984).

OMNeT++ (Objective Modular Network Testbed) is a well known open source simulator with component based C++ simulation library (Varga, 2014) and supports the integration of other frameworks such as Vehicles in Network Simulation (Veins) (Sommer et al., 2011). Veins is an open source framework for running vehicular network simulations which is based on OMNeT++ for event based network simulation and Simulation of Urban MObility (SUMO) for road traffic simulation. Each vehicle represents a mobile wireless node and their movement paths generated by mobility simulator, SUMO (Behrisch et al., 2011), are integrated into the network simulator, OMNeT++, which will manage the communication between the nodes, On-board Units (OBUs), both mobile and infrastructure (RSUs). Veins and SUMO are the normal types of simulation that researchers use to analyse highly mobile environments.

However, this simulation technique, though initially useful, did not support the proactive approach and hence MATLAB was used to develop our own discrete event simulation. To validate the analytical model which has been developed to find the probability of contention, a simulation has been built using MATLAB. The simulation has used an exponential function with chosen mean to generate instantaneous values. In addition, a standard discrete event simulation (Banks and Carson, 1984) in C++ and used in (Kirsal, 2013) has been used to validate the proposed queueing model.

3.2.1 MATLAB

MATLAB is a multi-paradigm numerical computing environment. A proprietary programming language developed by MathWorks, MATLAB allows implementation of algorithms, matrix manipulations, plotting of functions and data, the creation of user interfaces, and can be interfaced with other programs written in languages such as C, C++, C#, Java, Fortran, and Python (Chaturvedi, 2009).

Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD symbolic engine, allowing access to symbolic computing abilities. An additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems (Chaturvedi, 2009).

SimEvents is a package for MATLAB, which provides a discrete-event simulation engine and component library for analyzing event-driven system models and optimizing performance characteristics such as latency, throughput, and packet loss. Queues, servers, switches, and other predefined blocks enable you to model routing, processing delays, and prioritization for scheduling and communication (Chaturvedi, 2009).

3.3 Testbed Experiment

A testbed is a controlled experimentation platform, where solutions can be deployed and tested in an environment that resembles real-world conditions. Testbeds explore untested technologies or existing technologies working together in an untested manner. Testbeds generate requirements and priorities for standards organizations and culminate in new (potentially disruptive) products and services.

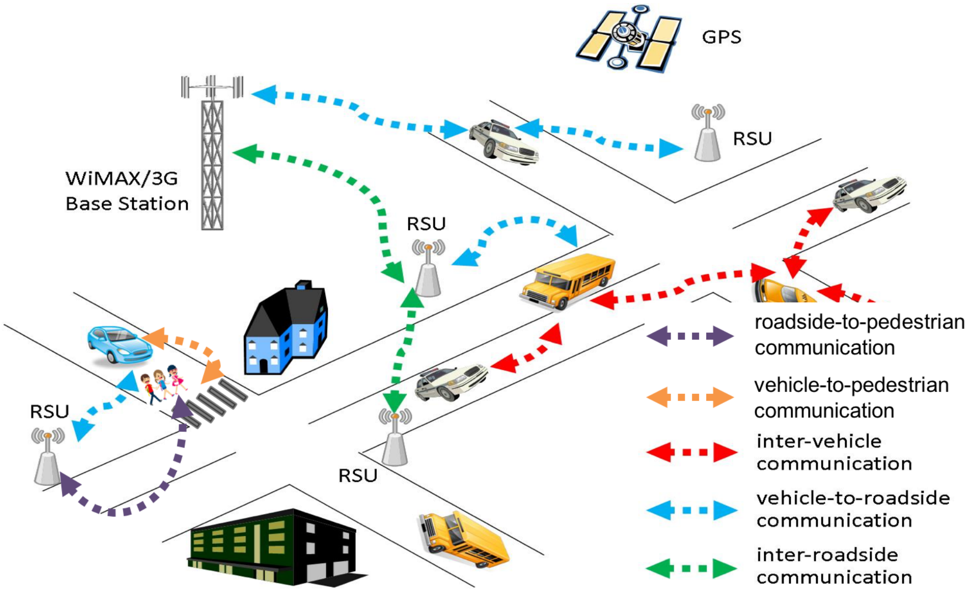


Figure 3.3: VANET Communication

This rapid growth in the number of vehicles on the roads has created a plethora of challenges for road traffic management authorities such as traffic congestion, an increasing number of accidents, air pollution, etc. Over the last decade, significant research efforts from both the automotive industry and academia have been undertaken to accelerate the deployment of a wireless network, Wireless Access in Vehicular Environments (WAVE) standard based on a Dedicated Short Range Communication (DSRC) among moving vehicles (Vehicle-to-Vehicle, V2V) and roadside infrastructure (Vehicle-to-Infrastructure, V2I). This network is called a Vehicular Ad-Hoc Network or VANET and is characterized by high node speed, rapidly changing topologies, and short connection lifetimes as shown in Figure 3.3. VANETs are realised by the deployment of RSUs located along the transport infrastructure and OBUs in the vehicles or worn by pedestrians or cyclists. Road safety and messaging and control require an extremely short latency wireless communication technology which can only be met by DSRC/WAVE.

In VANETs, vehicles periodically broadcast beacons that are essentially status messages used to discover and maintain neighbour relationships. In the U.S., the IEEE 1609 WAVE protocol stack builds on IEEE 802.11p WLAN operating on seven channels reserved in the 5.9 GHz frequency band. The protocol stack of WAVE is designed to provide a multi-channel operation to both emergency and entertainment applications. The components of the WAVE protocol architecture and its associated standards are summarized below and also depicted as shown in Figure 3.4.

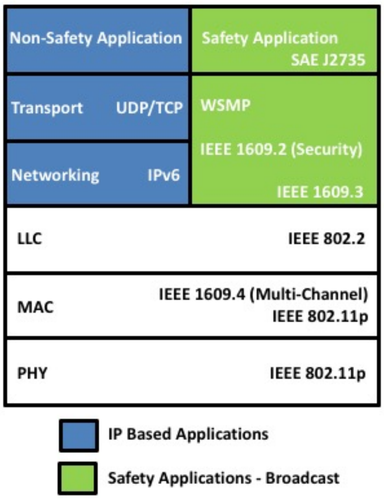
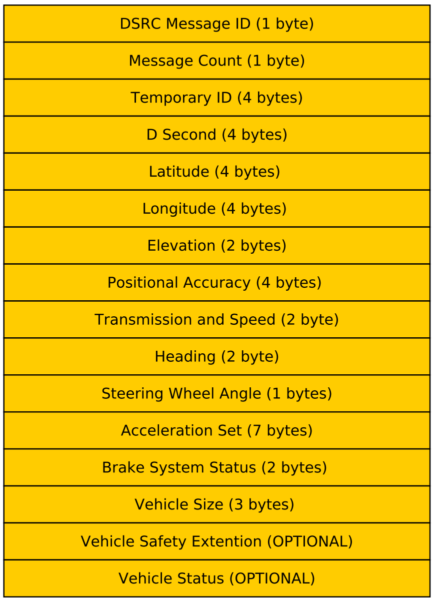


Figure 3.4: WAVE Protocol Stack (Hartenstein and Laberteaux, 2009)

- IEEE P1609.0 Draft Standard for WAVE Architecture
- IEEE 1609.1 Trial Use Standard for WAVE Resource Manager
- IEEE 1609.2 Trail Use Standard for WAVE Security Services for Applications and Management Messages
- IEEE 1609.3 Trail Use Standard for WAVE Networking Services
- IEEE 1609.4 Trail Use Standard for WAVE Multi Channel Operation
- IEEE P1609.11 Over-the-Air Data Exchange Protocol for ITS (Intelligent Transport System)
- IEEE 802.11p Part 11 Wireless LAN MAC (Medium Access Control) and PHY (Physical Layer) specifications - Amendment : WAVE



DSRC Message ID (1 byte)
Message Count (1 byte)
Temporary ID (4 bytes)
D Second (4 bytes)
Latitude (4 bytes)
Longitude (4 bytes)
Elevation (2 bytes)
Positional Accuracy (4 bytes)
Transmission and Speed (2 byte)
Heading (2 byte)
Steering Wheel Angle (1 bytes)
Acceleration Set (7 bytes)
Brake System Status (2 bytes)
Vehicle Size (3 bytes)
Vehicle Safety Extention (OPTIONAL)
Vehicle Status (OPTIONAL)

Figure 3.5: BSM Message Format

Vehicles have communication capabilities which allow them to send and receive network packets. They periodically broadcast traffic safety messages called Basic Safety Messages (BSMs) to all the other vehicles in its communication range with the frequency of 10 Hz, i.e., 10 BSMs/second, and they are called beacons.

The message format of BSM is as shown in Figure 3.5. Beacon messages are generated and broadcasted periodically between the vehicles (V2V) and the RSUs (V2I). These beacons sent from a vehicle will include information such as current speed, Global Positioning System (GPS) location, brake status, heading, etc. Other vehicles that receive these beacons will process them and take appropriate actions such as emergency brake, adjust their speed to avoid a possible collision, give priority to emergency vehicles, etc. With the help of these periodic BSM broadcast throughout the VANET, preventive actions can be taken to avoid traffic congestion and collisions. The European Telecommunication Standards Institute (ETSI) has also defined VANET Protocols and the messages formats for example Cooperative Awareness Message (CAM) (Hartenstein and Laberteaux, 2009). A VANET testbed has been deployed at the Middlesex University, London and surrounding roads along the campus, i.e., A41. The detailed description has been presented in Chapter 8.

3.4 Summary

This chapter has detailed all three methods i.e., analytical modelling, simulation and testbed experiments that are used for this research. It is very significant to use these methods to prove the validity and application of the proposed proactive resource allocation in a real scenario.

Chapter 4

Understanding Contention in Wired Network and Wireless Mobile Networks: Classical Approach

Contention means competition for resources. As a result, contention affects how resources can be used. The term is therefore used in networks to describe the situation where two or more nodes attempt to transmit a message across the same wire at the same time. A type of network protocol that allows nodes to contend for network access. This is true in multi-access networks such as Pure ALOHA, Slotted ALOHA, wired ethernet, etc., These systems have therefore evolved contention protocols that deal with contention such as CSMA/CD and CSMA/CA. Hence, in these systems, a contention protocol defines what happens when this occurs. However, it is important to understand the effects of contention and how that contention limits the optimal use of resources in this environment. Hence, it is very important to study contention and its effects on various types of networks, both, wired and wireless networks. Reducing individual user contention can lead to a tremendous improvement in the performance of network such as improved bandwidth and leads to an effective use of available resources.

In mobile environments it is also necessary to understand why contention should be considered and this can be done by looking at the classical approach to analyse the performance of mobile networks.

Classical handover occurs when a MN changes its PoA/Connection from the current wireless network to another network using a reactive approach i.e., the handover is initiated only after the MN is within the network coverage of the next wireless network. In this approach in order to start the handover the MN should be in the coverage range of the second cell and has to exchange the relative information to start the handover and complete the handover before exiting the first cell for a soft handover (Cottingham, 2009).

There are a number of parameters that need to be known to determine whether a handover is required. The signal strength of the BS with which communication is being made, along with the signal strengths of the surrounding stations. Additionally the availability of channels also needs to be known. The MN is obviously best suited to monitor the strength of the BSs at its location, but only the cellular network knows the status of channel availability and the network makes the decision about when the handover is to take place and to which channel of which cell (Cottingham, 2009).

Accordingly, the MN continually monitors the signal strengths of the BSs it can hear, including the one it is currently using, and it feeds this information back to the BS. When the strength of the signal from the BS that the MN is using starts to fall to a level where action needs to be taken then the cellular network looks at the reported strength of the signals from other cells reported by the MN. It then checks for channel availability, and if one is available it informs this new cell to reserve a channel for the incoming MN. When ready, the current BS passes the relevant information for the new channel to the MN, which then makes the change. Once there the MN sends a message on the new channel to inform that it is now within the coverage of the new network. If this message is successfully sent and received then the network shuts down communication with the MN on the old channel, freeing it up for other users, and all communication takes place on the new channel.

Under some circumstances such as when one base transceiver station is nearing its capacity, the network may decide to hand some MNs off to another base transceiver station they are receiving that has more capacity, and in this way reduces the load on the very busy base transceiver station. Hence, access can be opened to the maximum number of users. In fact channel usage and capacity are very important factors in the design of a cellular network. Therefore, in this chapter we will look into the effects of contention in different network access protocol such as Pure ALOHA and Slotted ALOHA and in the classical handover approach.

4.1 Pure ALOHA

Pure ALOHA is a simple network protocol where a node or station can transmit data i.e., a packet is sent when the node have data to send and while sending the packet if there is any other node also transmitting a packet it will lead to a collision (Abramson, 1970). This approach will lead to loss of both the packets and therefore, both the node has to retransmit the packet later as shown in Figure 4.1.

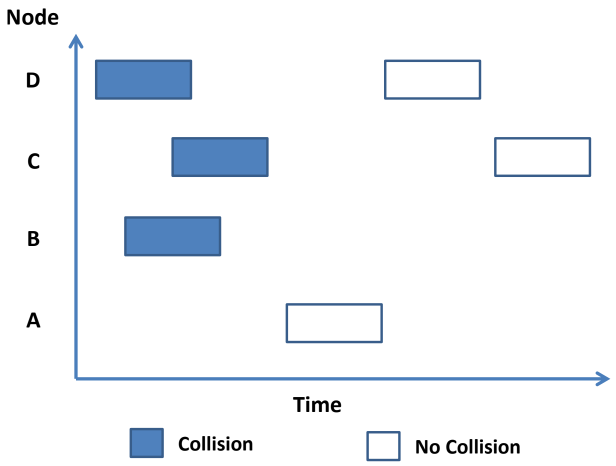


Figure 4.1: Pure ALOHA Protocol

4.1.1 Probability of Success

In the traditional analysis of Pure ALOHA there are N number of nodes transmit packets of length T and each node transmits with a probability P within the time period T. For a successful transmission of a packet there must be exactly one message within a time period of $2T$. This means that the packet of length T can be anywhere between 0 to $2T$ as shown in Figure 4.2.

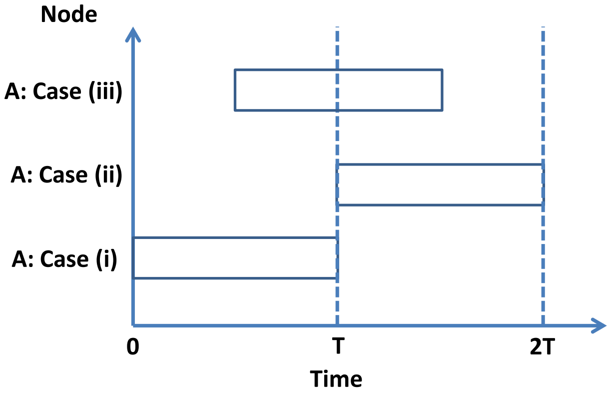


Figure 4.2: Successful Transmission Cases

The above analysis does not consider a case where there is probability that there can be a package transmitted by another node before 0 to T as shown in Figure 4.3.

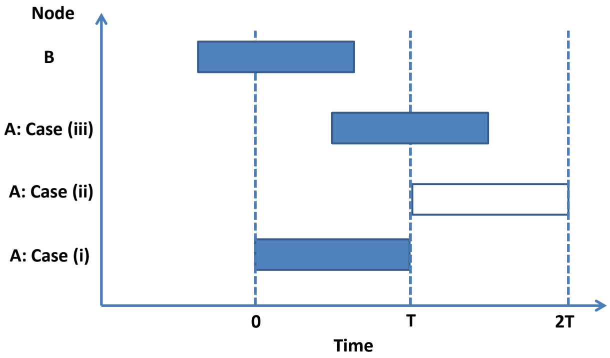


Figure 4.3: Unsuccessful Transmission

Therefore, this study suggests that for a successful packet the transmission exactly at T and should finish at $2T$ and any other node shouldn't start the transmission between 0 to $2T$. This ensures that if there was a packet transmitted before 0 will not collide with the packet transmitted with node A transmitting between 0 to T as shown in Figure 4.4.

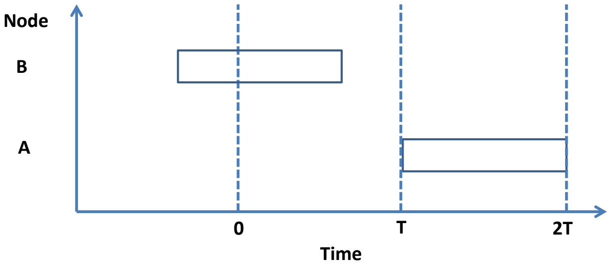


Figure 4.4: Successful Transmission

According to the above analysis the probability of successful transmission is P(Starting transmission at T) and P(No other transmission starting between 0 and $2T$).

$$P(\text{Starting transmission at } T) = \lambda e^{-\lambda T} \tag{4.1}$$

P(No other transmission starting between 0 and $2T$) is $(1 - \int_0^{2T} \lambda e^{-\lambda t} dt)^{(N-1)}$ and therefore,

$$P(\text{No other transmission starting between 0 and } 2T) = e^{-2T\lambda(N-1)} \tag{4.2}$$

Hence, probability of successful transmission for a node is

$$P(\text{Successful transmission}) = \lambda e^{-\lambda T} e^{-2T\lambda(N-1)} = \lambda e^{-\lambda T(2N-1)} \tag{4.3}$$

Therefore, probability of successful transmission for N nodes is

$$P(\text{Successful transmission [N nodes]}) = N\lambda e^{-\lambda T(2N-1)} \tag{4.4}$$

We know that the maximum success is when $\frac{dP(\text{Successful transmission [N nodes]})}{d\lambda} = 0$ where $\frac{dP(\text{Successful transmission [N nodes]})}{d\lambda}$ is shown below,

$$1 + \lambda T(1 - 2N) = 0 \tag{4.5}$$

Therefore, λ for maximum success is given by:

$$\lambda = \frac{1}{T(2N-1)} \tag{4.6}$$

Substituting above λ in Equation 4.4, we get the probability of maximum success for N nodes:

$$P(\text{Maximum Sucess}) = \frac{N}{T(2N-1)} e^{-1} \tag{4.7}$$

Let us consider T as a constant in this case $T = 1$, we get

$$P(\text{Maximum Sucess}) = \frac{N}{(2N-1)} e^{-1} \tag{4.8}$$

Thus, when N goes to ∞,

$$P(\text{Maximum Sucess}) = \frac{1}{2} e^{-1} \tag{4.9}$$

This shows that our analysis complies with the traditional analysis of pure ALOHA (Gao, 2009) (Abramson, 1970). However, our analysis allows us to explore in detail the relationship between probability of success and N the number of users contending for the resources. This is important because N in real systems is not infinite.

4.1.2 Throughput

We know that throughput is $\lambda P(\text{Successful transmission [N nodes]})$ packets/s that is:

$$\text{Throughput} = N\lambda^2 e^{-\lambda T(2N-1)} \tag{4.10}$$

We know that maximum throughput is when $\frac{dThroughput}{d\lambda} = 0$ and therefore, *lambda* shown below,

$$\lambda = \frac{2}{T(2N-1)} \tag{4.11}$$

When we substitute the above λ in the Throughput Equation we get,

$$\frac{4N}{T^2(2N-1)^2} e^{-2} \tag{4.12}$$

Let us consider that T is constant that in this case $T = 1$, the maximum throughput can be given as:

$$\text{Maximum Throughput} = \frac{4N}{(2N-1)^2}e^{-2} \tag{4.13}$$

4.2 Slotted ALOHA

Slotted ALOHA is an improved version of the Pure ALOHA which introduced discrete timeslots as shown in Figure 4.5. Here, the nodes can transmit only at the beginning of a time slot and therefore, if two or more nodes transmitting in the same timeslot will lead to collision. Compared to Pure ALOHA, probability of collision in Slotted ALOHA is less and therefore increases the throughput (Stallings and Manna, 1997).

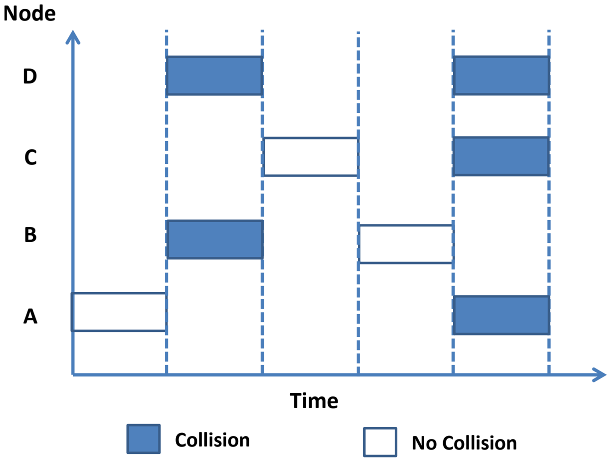


Figure 4.5: Slotted ALOHA Protocol

4.2.1 Probability of Success

Let us consider that the start of time slot is 0 and each slot is of length T, therefore, for a successful transmission at time T no other packet should be generated in the slot 0 to T. Hence, probability of successful transmission comprises (i) P(Packet successfully transmitted in the time slot) and (ii) P(No other node transmitting in the time slot). The respective Equations are shown below:

$$P(\text{Packet successfully transmitted}) = \int_0^T \lambda e^{-\lambda t} dt = (1 - e^{-\lambda T}) \tag{4.14}$$

$$P(\text{No other node transmitting}) = 1 - \int_0^T \lambda e^{-\lambda t(N-1)} dt = e^{-\lambda T(N-1)} \tag{4.15}$$

Therefore, probability of successful transmission for N nodes is shown below:

$$P(\text{Successful transmission [N nodes]}) = N(1 - e^{-\lambda T}) e^{-\lambda T(N-1)} \tag{4.16}$$

Probability of Maximum success is when $\frac{dP(\text{Successful transmission [N nodes]})}{d\lambda} = 0$ and we get,

$$e^{(-\lambda T)N} = \left(\frac{N-1}{N}\right)^N \tag{4.17}$$

Substituting the above equation in the P(Successful transmission [N nodes]) we get,

$$P(\text{Maximum Success}) = \left(\frac{N-1}{N}\right)^{N-1} \tag{4.18}$$

We can write the above equation as:

$$P(\text{Maximum Success}) = \left(1 - \frac{1}{N}\right)^{N-1} \tag{4.19}$$

If N goes to ∞, we can write the above equation as (Joyce, 2012):

$$P(\text{Maximum Success}) = \left(1 - \frac{1}{\infty}\right)^{\infty} = e^{-1} \tag{4.20}$$

This shows that our analysis complies with the traditional analysis of slotted ALOHA (Gao, 2009).

4.2.2 Throughput

We know that throughput is λP(Successful transmission [N nodes]) packets/s that is:

$$\text{Throughput} = \lambda[N(1-e^{-\lambda T})e^{-\lambda T(N-1)}] \tag{4.21}$$

We know that maximum throughput is when $\frac{dThroughput}{d\lambda} = 0$ and therefore,

$$e^{-\lambda T} = \frac{1+\lambda T(1-N)}{1-\lambda TN} \tag{4.22}$$

When we substitute the above Equation in the Throughput Equation we get,

$$N\left(\frac{-\lambda^2 T}{1+\lambda T(1-N)}\right)\left(\frac{1+\lambda T(1-N)}{1-\lambda TN}\right)^N \tag{4.23}$$

Let us consider that T is constant that in this case $T = 1$, the maximum throughput can be given as:

$$\text{Maximum Throughput} = N\left(\frac{-\lambda^2}{1+\lambda(1-N)}\right)\left(\frac{1+\lambda(1-N)}{1-\lambda N}\right)^N \tag{4.24}$$

4.3 Result

Figure 4.6 shows the probability of successful transmission of messages when there are N nodes in the system. We can observe that the probability reduces as the number of nodes increase. Even though the result is same as the traditional analysis of pure and slotted ALOHA, our analysis with the N taken into account shows that there is a need to consider the number of nodes. Especially in the case of mobile users the probability of success depends on various factors such as velocity, time a node is going to spend in the network etc. Therefore, calculating each node's probability of successfully served is very important to effectively serve the mobile users.

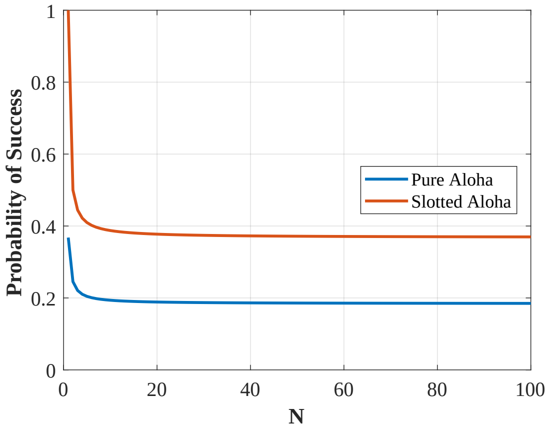


Figure 4.6: Probability of Maximum Success

4.4 Resource Allocation in Classical Handover Approach

This section explains and represents the resource allocation in classical handover approach for wireless and mobile environments where the request from the MN is placed in the queue to be served i.e., waiting for the channel to be served as shown in Figure 4.7.

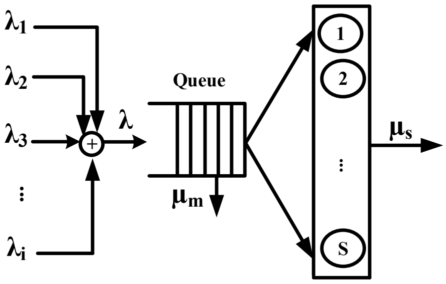


Figure 4.7: Classical Handover Multi-Channel Queueing System

The channel uses a FIFO service and requests are placed in the queue if the channel is busy. The arriving requests may be sent from different users to the system. Hence, the inter-arrival time of consecutive requests follows a Poisson process which can be distributed as an exponential distribution with arrival rate λ_i. Therefore, the total arrival rate can be calculated in Equation (4.25) below which is similarly calculated in (Kirsal et al., 2010).

$$\lambda = \sum_{i=1}^{N} \lambda_i \tag{4.25}$$

In this system, λ is the total arrival rate of the request, μ_s is the rate at which the requests are being served per server/channel, S is the number of servers/channels, the maximum number of requests in the queue is given by Q. In addition, the MN can leave the system due to mobility while waiting for the channel which is denoted as μ_m, as shown in Figure 4.7, where the multi-channel classical model is illustrated. Thus, μ_m can be calculated using Equation (4.26) used in the literature to calculate the dwell time in wireless and mobile systems for handover queuing models (Kirsal et al., 2010, 2013).

$$\mu_m = \frac{E[\nu] \cdot L}{\pi \cdot A} \tag{4.26}$$

Where $E[\nu]$ is the average of expected velocity (ν) of MN, L is the length of the perimeter of cell (a cell with an arbitrary shape is assumed), and A is the area of the cell. Hence, the total channel holding time of a call is exponentially distributed with mean $1/(\mu_s + \mu_m)$. If there are fewer than S requests in the system, $i < S$, only i of the S channels are busy and the combined service rate for the system is $i(\mu_s + \mu_m)$. Hence, μ_i can be calculated as follows:

$$\mu_i = \begin{cases} i(\mu_s + \mu_m) & 0 \leq i < S \\ \\ S\mu_s + i\mu_m & S \leq i \leq S+Q \end{cases} \tag{4.27}$$

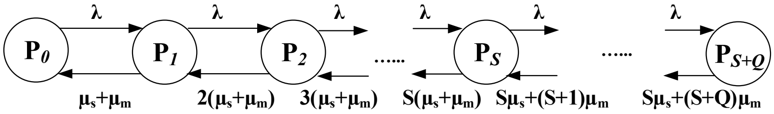


Figure 4.8: Classical Handover Multi-Channel State Diagram

In addition, ρ is the traffic intensity in the system where $\rho = \lambda/\mu$. P_i is the probability that there are i requests in the system as shown in Figure 4.8. Assuming a system in a steady state, the state probabilities, P_i's, can be obtained as in Equation (4.28) based on standard queuing theory (Jain, 2015).

$$P_i = \begin{cases} \frac{\lambda^i P_0}{i!(\mu_s+\mu_m)^i} & 0 \leq i \leq S \\ \frac{\lambda^i P_0}{S!(\mu_s+\mu_m)^S \prod\limits_{j=S+1}^{i} [S\mu_s+j\mu_m]} & S < i \leq S+Q \end{cases} \tag{4.28}$$

The mean queue length (MQL) i.e., the average number of requests in the system can then be calculated as $MQL = \sum\limits_{i=0}^{S+Q} i \cdot P_i$ which gives:

$$MQL = \left[\sum_{i=0}^{S} \frac{i\rho^i}{i!} + \sum_{i=S+1}^{S+Q} \frac{i \cdot \frac{\rho^S}{S!} \cdot \lambda^{i-S}}{\prod\limits_{j=S+1}^{i} \left[S\mu_s + (j-S)\,\mu_m \right]} \right] P_0 \tag{4.29}$$

In addition, the throughput (γ) and mean response time (MRT) of the system can be calculated as follows:

$$\gamma = \sum_{i=0}^{S+Q} i \cdot \mu_i P_i \tag{4.30}$$

$$MRT = \frac{MQL}{\gamma} \tag{4.31}$$

4.5 Enhanced Classical Handover Proposed in (Kirsal, 2013)

Firstly, Kirsal has analysed the horizontal handover in individual networks using performability techniques and then looked at vertical handover between cellular networks using pure performance models. Vertical handover between different networks such as GSM, GPRS, UMTS, EDGE and LTE was modelled to yield performance results such as mean queue length and blocking probability over a range of different conditions. The study found that the common models that were being used were subjected to handover hysteresis resulting from feedback loops in the model. Therefore, a new analytical model was proposed to address this issue. As a practical solution to the observed handover hysteresis, guard channels analysis method was used to give priority to handover in mobile networks.

Kirsal's work took a service oriented approach for MNs, considering that the services will be resumed as soon as the MN moves to a different network. But the work did not focus on the resources available at the AP or BS and the effects of mobility in acquiring this resource. Understanding this effect is very important as the MNs would be waiting to acquire a channel and might move out of the current network due to mobility without service. The focus of this chapter is to understand this effect of mobility in acquiring the resources and services offered.

The problem with the classical handover approach is that the AP/BS does not know in advance the network requirements of MNs heading towards it. Due to this the MN even after entering the network coverage, it will have to wait until the resource becomes free. In a highly mobile environments, MNs will have less time to spend in a network coverage therefore, the AP/BS has to anticipate the network conditions much before based on the MNs about to reach its network in order to have an effective resource utilization.

4.6 Useful Service vs Mobile Service

Since the MN can leave the queuing system due to service (μ_s) or due to mobility (μ_m), it is necessary to distinguish these two events to properly reflect the performance of the system. We therefore define two concepts which are important in a mobile environment. The first is Useful service in classical handover, U_{sc}, where the MN leaves the system after using the channel. When the MN leaves the system due to mobility and is not served by the channel, this is called Mobile Departure or Service, U_{mc}. For the classical case, we can represent these parameters as follows:

$$U_{sc} = \frac{S\mu_s}{S\mu_s + Q\mu_m} \tag{4.32}$$

$$U_{mc} = \frac{Q\mu_m}{S\mu_s + Q\mu_m} \tag{4.33}$$

The system parameters used are mainly taken from (Kirsal et al., 2010, 2013) based on the relevant literature (Gemikonakli et al., 2013; Ghosh et al., 2014b; Kirsal et al., 2010, 2013). The system has a fixed number of identical channels: $S = 12$. Q is the queuing capacity, which represents the number of requests waiting for service and is limited with $Q = 100$. The service rate of the system μ_s is 0.01 requests/sec. The average speed of the MN and the radius of the network are taken as 10km/h to 80km/h and 1000m for all calculations, respectively. The rates are translated into requests per second in order to use consistent values.

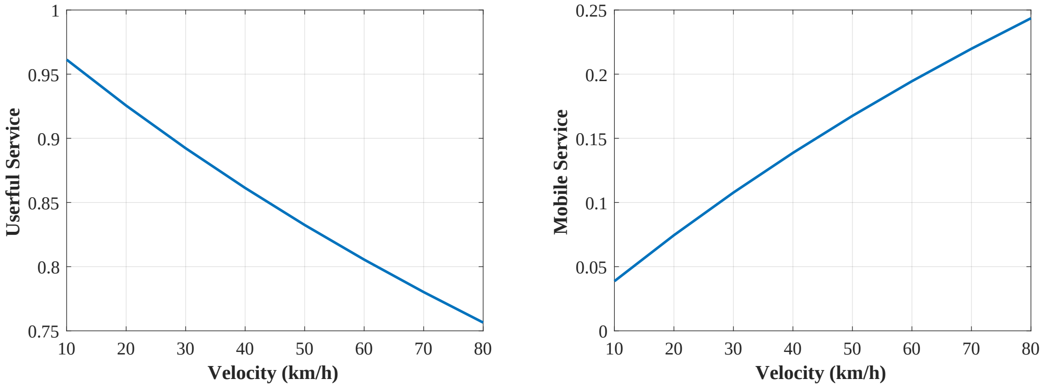


Figure 4.9: Useful Service vs Mobile Service

We can observe from Figure 4.9 that as the velocity of the MN increases, the useful service of the system is significantly reduced and the mobile service is increasing. This means that the number of MNs which are moving out of the network due to mobility without being served is increasing. Hence, in mobile networks it is necessary to take into account the mobility aspects to ensure effective resource management, so that we can optimize the useful service of the channels. In these environments most users will leave the system without being served resulting in poor user experience as well as poor system performance.

The key observation is that because of the reactive approach in the classical model, MNs can queue for a channel with no hope of getting the channel while they are in the coverage area of the network. As a result this is useless waiting and if this inability to obtain a channel can be signalled to the MN before queueing for the channel then it would allow the user to immediately look for an alternative network and hence would improve the Quality of Experience (QoE) as well as the overall system performance. This means that the analysis of contention between different MNs has to be analysed in detail so that only MNs that have a chance of getting the channel should be queued for service.

4.7 Summary

Analysing user contention is very important to understand how contention for resources affect their optimal use and allow us to target specific strategies to improve the overall use of resources. Whereas, it has been shown in this Chapter that the traditional performance analysis of these networks focuses only on overall demand at infinite population levels i.e., only at N goes to ∞ hence, can only be an estimate.

With the analysis presented, we can come to a conclusion that the traditional classical reactive handover approach cannot cope up in high mobility environment. There is a considerable amount of users moving out of the network without useful service due to their mobility. A detailed understanding of contention between mobile users must be explored and will be presented in the following chapter.

Chapter 5

Proactive Systems

The problems faced in the classical approach due to high mobility were presented in the previous Chapter. As mentioned earlier, the proactive approach using user contention analysis should help any system to perform better than a traditional approach. This Chapter will introduce two new parameters that has been derived from Y-Comm framework for analysing the contention; Time to get resource and Resource Hold Time and present necessary conditions to identify contention based on these two parameters when a mobile user is requesting for resources. Before looking into the conditions, it is necessary to understand the background of Y-Comm framework and different parameters proposed in order to achieve a proactive handover.

5.1 Y-Comm Reference Framework

Y-Comm is based around a future vision with key assumptions that the mobile devices will have several wireless interfaces to function in an increasing heterogeneous environment and the deployment of these heterogeneous wireless networks will point to a significant change in the architecture of the Internet. Internet is now evolving with a super-fast core composed of an optical backbone and fast access networks while the end networks dominated by the deployment of wireless technologies. Y-Comm reflects this by dividing the Future Internet into two frameworks as shown in Figure 5.1.

- The Core Framework shows the functionality required in the core network to support the Peripheral Framework
- The Peripheral Framework deals with operations and functions on the MN

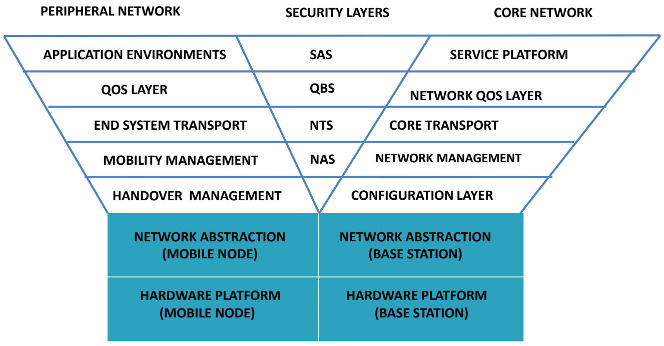


Figure 5.1: YComm Reference Framework (Mapp et al., 2016)

5.1.1 Peripheral Framework

The Hardware Platform and Network Abstraction layers run on the mobile to support various wireless network technologies. All the relevant wireless technologies are classified using the Hardware Platform Layer. These wireless technologies are characterised based on MAC, electromagnetic spectrum, and modulation techniques. Wireless networks are managed and controlled by Network Abstraction Layer using a common interface. The Handover Management Layer executes vertical handover. To perform a handover, this layer does the signalling, acquires the resources, and context transfer for vertical handover. The Mobility Management Layer decides whether and when a handover should occur by analysing parameters such as signal strength and using policy rules to decide both the time and place for doing the handover (Aiash et al., 2013).

The End System Transport Layer is used to provide transport and network functions to make end-to-end connections across the core network for the MN. The QoS Layer supports Upward QoS and Downward QoS mechanisms. Upward QoS is where the application itself tries to adapt to the changing QoS and also monitors the QoS used by the wireless network as a whole to ensure stable operation. Downward QoS is where an application specifies its required QoS to the system and the system attempts to maintain this QoS over varying network channels. The final layer, Applications Environments Layer specifies a set of objects, routines, and functions to build applications which make use of the framework (Aiash et al., 2013).

5.1.2 Core Framework

In the Core Framework, the Hardware Platform and the Network Abstraction layers are used to control the functions of BS or AP of different wireless technologies. The Configuration Layer is a control plane employed to manage key mobile infrastructure such as switches, routers, and other infrastructures using programmable networking techniques. Networking operations in the core are controlled by a management plane called the Network Management Layer. Here, the core is divided into number of networks which are managed in an integrated fashion. It also gathers information on peripheral networks such that it can inform the Mobility Management Layer running on MNs about wireless networks at various locations (Aiash et al., 2013).

The Core Transport Layer is responsible for moving data through the core network and core endpoints with a specified level of security and a given QoS. The point where peripheral networks join the core network is called a core endpoint and several peripheral networks may be attached to a single-core endpoint. QoS issues within the core network especially at the interface between peripheral networks and the core network are handled by the Network QoS Layer. Finally, the Service Platform Layer allows services to be installed on various networks at the same time (Aiash et al., 2013).

5.1.3 Proactive Handovers in Y-Comm

According to Y-Comm Framework, to perform proactive handover using a mathematical model approach, it is necessary to know the topology of these local networks and the Network Management Layer in the Core Framework manages this information. Therefore, the MN polls this layer to obtain information with regard to all nearby wireless networks, their topologies, and QoS characteristics as shown in Figure 5.2. This information is then used by the Mobility Management Layer along with the speed and direction of the MN to determine where and when handover should occur. The Mobility Management Layer calculates time before handover and network dwell time and this information is communicated to the Handover Management Layer which immediately requests resources to do a handover. Even though the resources are reserved early, handover will be initiated only when the time before handover expires.

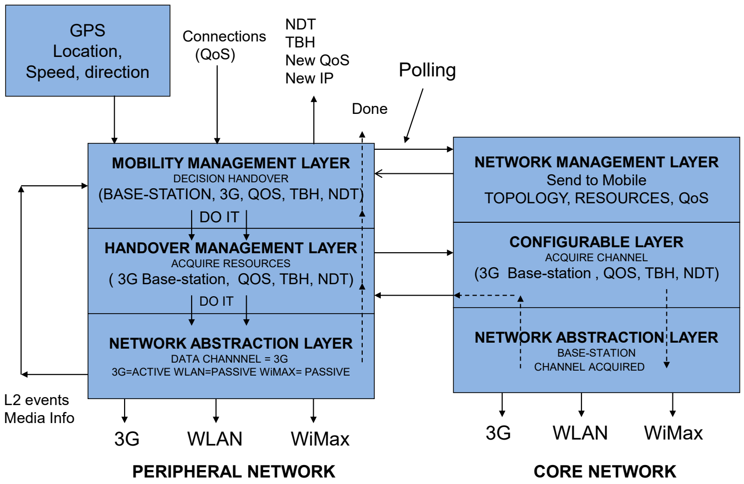


Figure 5.2: Handover in Y-Comm Framework

In addition, once the decision to handover is taken by the Mobility Management Layer, the new QoS, the new IP address, as well as a new time before handover, and estimated network dwell time are communicated to the upper layers. The upper layers with the acquired time before handover are expected to take the necessary steps to avoid any slow adaptation, latency or packet loss. For example, End System Transport Layer may signal an impending change in the QoS on current transport connections and therefore, the packets can be buffered ahead of the handover. After handover, the previous channel used by the MN is released.

5.2 Handover Coverage Parameters Defined in Y-Comm

In this section, we introduce a set of network coverage parameters. The network coverage area is a region with an irregular shape where signals from a given PoA ie., AP or BS can be detected. The signals from the PoA are unreliable at the boundary and beyond the coverage area as the signals from the PoA cannot be detected. For seamless communication, handover should be finished before the coverage boundary is reached. Therefore, a circle known as the handover radius (R_H) and exit radius (R_E) were defined in (Mapp et al., 2012) to ensure smooth handover. The work states that the handover must begin at the exit radius and should be completed before reaching the handover radius boundary as shown in Figure 5.3.

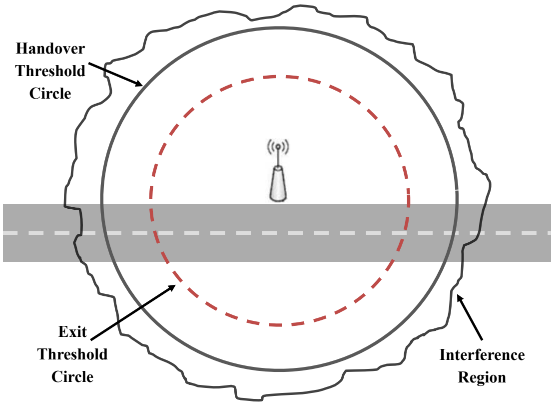


Figure 5.3: Network Coverage (Mapp et al., 2012)

The exit radius will therefore be dependent on the velocity, ν, of the MN. If we represent the time taken to execute a handover by T_{EH}, then:

$$T_{EH} \leq \frac{(R_H - R_E)}{\nu} \tag{5.1}$$

Hence, exit radius can be given as shown in Equation (5.2)

$$R_E \leq R_H - (\nu * T_{EH}) \tag{5.2}$$

So, the faster a MN moves the smaller the R_E at which handover must begin.

Given that we know the time taken to execute a handover, the velocity of the MN and handover radius, then we can calculate the exit radius which is dependent on the handover radius. A good estimation of the handover radius is required for the proposed approach which depends on the propagation models being used (Mapp et al., 2012). The time taken to effect a handover was shown to be dependent by various factors such as:

- Detection Time (t_{det}) is the time taken to detect the first signal of the new network.
- Configuration Time (t_{con}) is the time from detecting the network to the time taken by the MN to get and configure its Care-of-Address (CoA).
- Registration Time (t_{reg}) is the time between delivery of Binding Update to the Home Agent and corresponding nodes and the time a packet arrives on the new interface.
- Adaptation Time (t_{adp}) is the time taken by the MN after a vertical handover to adapt the connection to the new technology at the transport layer by adjusting the TCP state machine parameters (e.g., congestion window size, time-out timers, etc.), due to differences in the link characteristics.

Since reactive handovers respond to network conditions, all four times must be added together. The MN therefore knows nothing beforehand about the characteristics of various networks. For the proactive handover technique, there is no detection time since the MN would know where all the local networks are located. This is particularly valid for vechicular networks as the route is fixed and therefore, the location of the next (target) network is likely to be known. Configuration time is also negligible since the MN will know the IP address of the target network. Registration Time is still valid. In addition, for proactive networks the need for the transport protocol to adapt can be signalled before or during handover and not after the handovers occur. Therefore, it means that the adaptation time can be done in parallel with the registration time. So for proactive handover,

$$T_{EH} = MAX(t_{reg}, t_{adp}) \tag{5.3}$$

As mentioned above a good estimation of the handover radius is required which, in turn, is dependent on the propagation models being used. Propagation models attempt to model the electromagnetic signal's received signal strength at a given distance from where the signal is being transmitted. Propagation models based on mathematical equations called as Empirical Models are well known and widely used. Empirical Models take into account the propagation effects on the electromagnetic signal. Time Dispersive and Non-time Dispersive models are the two types of Empirical models. The former takes into account channel characteristics such as multipath spread and the later predicts the path loss in terms of distance, frequency, and height of antenna. These models are mainly based on measurements as well as observations. Cost-231 and Hata are examples of Non-time Dispersive models (Abhayawardhana et al., 2005).

In order to calculate the exit radius and handover radius, a geometry-based mathematical framework was proposed in (Mapp et al., 2012) for different scenarios such as complete coverage, where a femto cell is placed inside a macro cell, intersecting networks where two cells are placed in an over lapping fashion and non-intersecting networks where two cells are placed apart without an overlap. Even through the proposed approach extensively looked into calculating time before handover and network dwell time using fixed cell size, it has failed to look into other important aspects such as packet size, frequency of packets being exchanged and different velocities of the MN. It has been found in (Ghosh, 2016) that the above mentioned parameters has a significant impact on accurately calculating the handover radius.

5.3 Proactive Handover and Application to VANET Proposed in (Ghosh, 2016)

According to (Ghosh, 2016), due to the high velocity of the vehicles and smaller coverage distances, there are serious challenges in providing seamless handover from one RSU to another and this comes at the cost of overlapping signals of adjacent RSUs. Therefore, to guarantee ubiquitous connectivity a framework was proposed to calculate the regions of overlap in adjacent RSU coverage ranges. The study used the VEINs framework via OMNeT++ for simulations and then used analytical approach to calculate the probability of successful packet reception. Investigation showed that network dwell time and seamless communication was dependent on the velocity of the vehicle, length of the beacon, and the beacon frequency. A detailed analysis and results were presented which show the need for a more probabilistic approach to handover based on cumulative probability of successful packet reception. For convenience, network dwell time is represented as NDT only in this section.

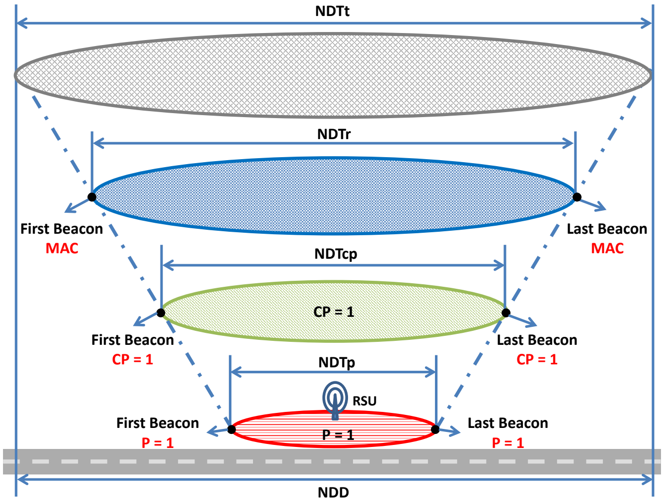


Figure 5.4: Probability Based Network Dwell Distance Types (Ghosh, 2016)

The work in (Ghosh, 2016) presented results from a V2I simulation study, to understand the communication range of a single RSU based on Probability of beacon reception at vehicle. The effects of beaconing on NDT using cumulative probability (CP) and individual beacon reception probability (P) was presented in (Ghosh, 2016). Extensive simulation was carried out with different velocity, different beacon sizes and beacon frequency. Theoretical NDT proposed in (Shaikh et al., 2007a) was presented as NDT_t and NDT was further classified as realistic NDT (NDT_r) i.e. the time between the first & the last beacon reaching the MAC layer, single beacon probability (NDT_P) and cumulative probability NDT (NDT_{CP}) as shown in Figure 5.4. The work also showed that the size of the beacon directly affects the P and CP. However, the frequency of the beacon only affects the CP and showed that the effect of beacon size and beacon frequency are orthogonal to each other with regard to the NDT. It also highlighted that for handover, where predictability is important, maximum beacon size around 600 to 800 bytes (approx.) could give the best chance for seamless communication and a good range of beacon frequency is 10 to 20 Hz.

The main aim was to analyse the effects of Proactive Handover and the Proactive Resource Allocation Mechanisms on the wireless network infrastructure and MNs. The analytical models for the proactive approach are developed using the concepts of Y-Comm framework and validated using simulation and tested in the VANET testbed at Middlesex University.

5.4 Network Coverage Parameters for Mobile Networks

The above mentioned work has shown different coverage parameters and highlighted the importance of segmenting the communication ranges to achieve a seamless handover. In this work we are further exploring and redefining the communication range segments as shown in Figure 5.5 which can be put into effective use for achieving both proactive handover and resource allocation for a highly mobile environment.

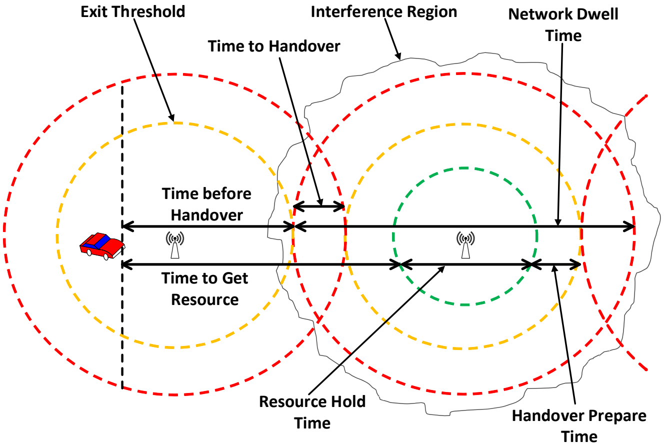


Figure 5.5: Communication Range Segmentation

Figure 5.5 shows a more advanced scenario in which three consecutive overlapping wireless networks are segmented based on various key time variables which can be used to enhance handover and resource allocation. An example of such a scenario has been developed in London in which a Middlesex (MDX) VANET testbed were three RSU were deployed in an overlapping fashion on the A41 road Middlesex University to study and test vehicular communication.

Time before handover (Υ) is the time after which the handover process should start and Time to handover ($\hbar$) is the time before which the handover to next coverage range has to be completed, if not it will result in a hard handover. Network Dwell Time ($\aleph$) as defined in Y-Comm Framework is the time MN will spend in the coverage i.e., the Network Dwell Distance (NDD) of new network. Time to get resource, $\mathbb{T}$ is the time when actual resource requested is available to the requested user i.e., even after entering the network's coverage range with a successful soft handover, the resource required by the MN might not be available, for example, other users might be holding the resource. Resource Hold Time, $\mathbb{N}$ is the resource usage time or when actual exchange of data is taking place. Handover Prepare Time (ϱ) is the time taken to prepare for handover during which the resource usage or data transmission will be paused and will be resumed after successful handover to the new network. Usually $\hbar$ and ϱ are very small compare to the values of other segments and therefore, $\mathbb{T}$ can be approximately equal to Υ if there are resources available in the new network, i.e., if there is no contention.

With the knowledge of these coverage parameters, it is possible to enhance the resource management based on the user contention for resources in a mobile network with proactive handover. A new proactive resource allocation based on the user contention to acquire a wireless channel resource is presented in the following section.

5.5 New Proactive Resource Allocation and Contentions Applied to Acquiring a Channel in a Wireless Network

Resource allocation in the classical handover approach has a persistent danger of a mobile user waiting for the channel and never acquiring it due to mobility resulting in a suboptimal network performance. Therefore, it is necessary to analyse in detail the contention between individual mobile users in order to increase the effective use of communication resources. In this section we look at contention for channel resources using three possible outcomes: No Contention, Partial Contention and Full Contention based on two key parameters i.e., Time to Get Resource ($\mathbb{T}$) and Resource Hold Time ($\mathbb{N}$). Therefore,

$$\mathbb{T} = \hbar + \mathbb{Z} \tag{5.4}$$

Where, $\mathbb{Z}$ is the delay due to contention or queueing effects for the resource. Hence,

$$\mathbb{N} = \aleph - \mathbb{Z} \tag{5.5}$$

Therefore, the aim of this work is to minimize $\mathbb{Z}$ and inorder to do that we need to perform a detailed analysis of contention. We will look into a simple scenario where a network uses a single channel and two MNs are moving at a velocity (υ) towards that network range as shown in Figure 5.6. MN$_A$ and MN$_B$ can request the channel for communication. Assuming that υ, the velocity and t_c, the current time of the node are known; $\mathbb{T}$, the time to get resource and $\mathbb{N}$ which is the estimated resource hold time of the MN in the next network, can both be represented using a probabilistic distribution such as the exponential distribution. Hence, $\mathbb{T}_{MN}$ and $\mathbb{N}_{MN}$ are instantaneous values for $Node_{MN}$ based on their probabilistic distribution at time t_c.

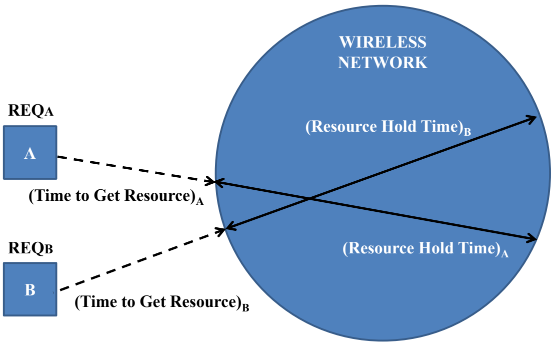


Figure 5.6: Request for Channel Allocation

The time when the channel will be needed for communication and when a MN will release the channel is as shown below:

- MN_A needs channel at $(t_c + \mathbb{T})_A$
- MN_A releases the channel at $(t_c + \mathbb{T} + \mathbb{N})_A$
- MN_B needs channel at $(t_c + \mathbb{T})_B$
- MN_B releases the channel at $(t_c + \mathbb{T} + \mathbb{N})_B$

In our analysis, based on the channel request and holding time of MN_A there are three possible contention outcomes in this scenario i.e., No Contention, Partial Contention and Full Contention. For a two node scenario, all the possible conditions for a given contention outcome can be represented in a tree form as shown in Figure 5.7. In addition, for simplicity, t_c is assumed to be zero. The left branch of the tree shows the condition for MN_A entering the next coverage range first i.e., $\mathbb{T}_A < \mathbb{T}_B$ and right branch shows the conditions for MN_B entering the next coverage range first i.e., $\mathbb{T}_B < \mathbb{T}_A$. To identify the type of contention; the branch which satisfies the condition has to be followed until the last requisite condition is met.

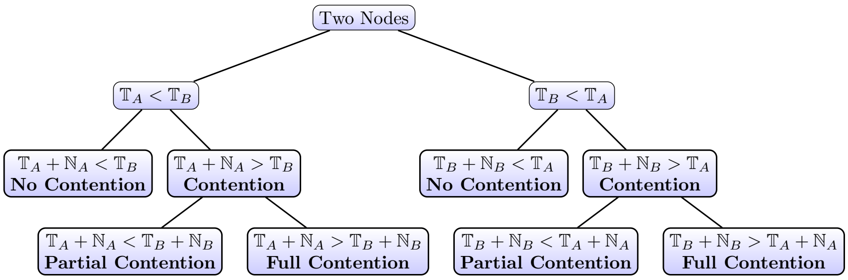


Figure 5.7: Contention Probability Conditions for Two MNs Scenario

In the analysis described below, $\wedge$ represents 'AND'.

5.5.1 No Contention

No Contention occurs when,

$$(\mathbb{T}_A < \mathbb{T}_B) \wedge (\mathbb{T}_A + \mathbb{N}_A < \mathbb{T}_B)$$

If $(\mathbb{T}+\mathbb{N})_A < \mathbb{T}_B$ i.e., channel release time of MN$_A$ is less than the time when MN$_B$ needs the channel. Hence, there is no contention as MN$_B$ needs the channel after MN$_A$ has used the channel.

5.5.2 Partial Contention

Partial Contention occurs when,

$$(\mathbb{T}_A < \mathbb{T}_B) \wedge (\mathbb{T}_A + \mathbb{N}_A > \mathbb{T}_B) \wedge (\mathbb{T}_A + \mathbb{N}_A < \mathbb{T}_B + \mathbb{N}_B)$$

If $\mathbb{T}_A < \mathbb{T}_B$ and $(\mathbb{T}+\mathbb{N})_A < (\mathbb{T}+\mathbb{N})_B$ i.e., channel release time of MN$_A$ is less than the channel release time of MN$_B$. This means that MN$_A$ uses the channel first. However, MN$_A$ releases the channel when MN$_B$ is still within range to make use the channel and hence there is a partial contention.

5.5.3 Full Contention

Full Contention occurs when,

$$(\mathbb{T}_A < \mathbb{T}_B) \wedge (\mathbb{T}_A + \mathbb{N}_A > \mathbb{T}_B) \wedge (\mathbb{T}_A + \mathbb{N}_A > \mathbb{T}_B + \mathbb{N}_B)$$

If $\mathbb{T}_A < \mathbb{T}_B$ and $(\mathbb{T}+\mathbb{N})_A > (\mathbb{T}+\mathbb{N})_B$ i.e., channel release time of MN$_A$ is greater than the channel release time of MN$_B$. In this scenario MN$_A$ uses the channel and releases the channel when MN$_B$ is no longer in range of the next network due to mobility. Hence, MN$_B$ never gets access to the channel, this is called full contention.

Therefore, in the event of a full contention MN$_B$ will not get the channel from the next network range. If this full contention can be identified and MN_B can be notified before it reaches the next network range, then the contention can be signalled and MN$_B$ can use other available networks via vertical handover techniques instead of waiting for the channel which will never be available. For no or partial contention MN$_B$ can be signalled that it will get to use the channel and hence can queue for service. This approach should result in better network performance.

5.6 Summary

A comprehensive background of the Y-Comm reference framework which has aimed to provide ubiquitous communication for future networks and the parameters such as time before handover and network dwell time proposed in the framework has been presented in this Chapter. With the detailed coverage segmentation and the parameters of Y-Comm framework, a new proactive resource allocation based on user contention has been described. Two new parameters has been derived from Y-Comm framework was introduced to analyse the contention; Time to get resource and Resource Hold Time. In addition, the conditions to identify contention based on these two parameters when a mobile user is requesting for resources was presented. A detailed probabilistic model for the proposed user contention conditions is presented in the following Chapter.

Chapter 6

Calculating the Probability of Contention in Mobile Networks using $\mathbb{T}$ and $\mathbb{N}$

A comprehensive study of different user contention encountered when acquiring resources in a wireless mobile system was presented in Chapter 5. It is very important to mathematically analyse those conditions and therefore in this Chapter, a probabilistic model will be developed to find the probability of contention in acquiring channel resources based on the mobility. First, a scenario where two MNs contesting for channel around the same time will be modelled and the approach is then extended to a three MNs scenario. Secondly, the complexity of scaling this approach for N number of MNs will be highlighted.

6.1 Probability Of Contention Between Two MNs Competing for the Same Network

In order to find the probability of different contention scenarios, the respective conditions shown in Figure 5.7 have to be satisfied. In addition, the probabilities of all possible outcomes must sum to one. That is, the sum of the probabilities of No Contention, Partial Contention and Full Contention for MN_A and MN_B is always equal to one.

Now the conditions in each level of tree have to be considered in acquiring the probability of different type of contention. Hence, we define our variables as shown below:

- Z is an exponential random variable of $\mathbb{T}_A$
- Y is an exponential random variable of $\mathbb{N}_A$
- X is an exponential random variable of $\mathbb{T}_B$
- W is an exponential random variable of $\mathbb{N}_B$

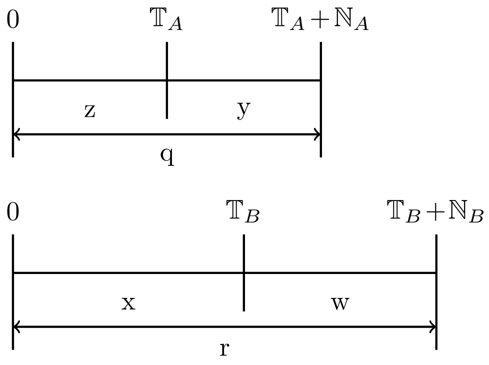


Figure 6.1: Variables for Two Node Scenario

Therefore,

- $P(Z = z) = \tau_A e^{-\tau_A z}$
- $P(Y = y) = \eta_A e^{-\eta_A y}$
- $P(X = x) = \tau_B e^{-\tau_B x}$
- $P(W = w) = \eta_B e^{-\eta_B w}$

where, τ_A and τ_B are the reciprocals of the average time that MN_A and MN_B will take to reach the new network to get the resource, while, η_A and η_B are the reciprocals of the estimated mean resource hold time during dwell time in the new network.

- $\tau_A = \frac{1}{\mathbb{T}_A}$
- $\eta_A = \frac{1}{\mathbb{N}_A}$
- $\tau_B = \frac{1}{\mathbb{T}_B}$
- $\eta_B = \frac{1}{\mathbb{N}_B}$

6.1.1 No Contention

The conditions of No Contention for MN_A can be represented by limits as shown in Figure 6.1. We now define the variable Q which is the sum of $\mathbb{T}_A + \mathbb{N}_A$ can be expressed as $Q = Z + Y$; hence the $P(Q = q)$ is as shown in Equation (6.1).

$$\int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} dz \tag{6.1}$$

The probability of No Contention for MN_A is given by the following conditions and are used to generate the limits for the resulting probabilities:

- $z < x$
- $q < x$

Note if $q < x$, it implies that $z < x$. Therefore, for $P(\text{No Contention}) = P(Q = q)P(x > q)$ where q goes from 0 to ∞ which is given Equation (6.2)

$$\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \tau_B e^{-\tau_B x} dx \Big) dz dq$$

$$= \frac{\tau_A \eta_A}{(\tau_A + \tau_B)(\eta_A + \tau_B)} \tag{6.2}$$

6.1.2 Partial Contention

As shown in Figure 6.1, the conditions of Partial Contention for MN_A can be represented below and are used to generate the limits for the resulting probabilities. The variable R, which is the sum of $\mathbb{T}_B + \mathbb{N}_B$ can be expressed as $R = X + W$.

- $z < x$
- $q > x$
- $q < r$

This implies that $x > z$ and $x < q$ and $r > q$ which is represented by Equation (6.3)

$$\int_q^\infty \int_z^q \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} dx dr \tag{6.3}$$

Therefore, since q goes from 0 to ∞, the probability of Partial Contention for MN_A is given by:

$$\begin{aligned} &\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \\ &\Big(\int_q^\infty \int_z^q \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} dx dr \Big) dz dq \end{aligned}$$

$$= \frac{\tau_A \tau_B \eta_A}{(\tau_A + \tau_B)(\eta_A + \tau_B)(\eta_A + \eta_B)} \tag{6.4}$$

6.1.3 Full Contention

As shown in Figure 6.1 the conditions of Full Contention for MN_A can be represented below and are used to generate the limits for the resulting probabilities.

- $z < x$
- $q > x$
- $q > r$

In Full Contention, x goes from z to r and r goes from z to q as shown in Equation (6.5).

$$\int_z^q \int_z^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} dx dr \tag{6.5}$$

Therefore, the probability of Full Contention for MN_A is given by:

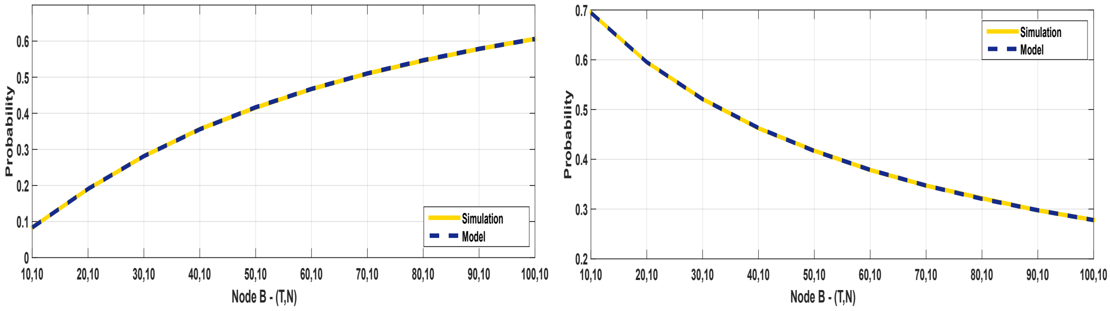
$$\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)}$$
$$\left(\int_z^q \int_z^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} dx dr \right) dz dq$$

$$= \frac{\tau_A \tau_B \eta_B}{(\tau_A + \tau_B)(\eta_A + \tau_B)(\eta_A + \eta_B)} \tag{6.6}$$

Similarly, the Probabilities of No, Partial and Full Contention can be derived for MN_B.

6.1.4 Comparison of Results for Two Nodes

These results were verified using a MATLAB simulation program, where instantaneous values of $\mathbb{T}$ and $\mathbb{N}$ for the various nodes were generated using the exponential function with chosen mean $\mathbb{T}$ and $\mathbb{N}$ parameters. These values were then compared according to our conditions to determine the contention outcome of each set of values to yield No Contention, Partial Contention and Full Contention conditions as shown in Algorithm 1. One million sets of values were considered and the results of different contention outcomes were summed and divided by the total number of events to get the probability of each type of contention which was then compared with the analytical results.The results of the analytical model are in good agreement with the simulation results with discrepancy of less than 2%.



(a) MN_A

(b) MN_B

Figure 6.2: Probability of No Contention - Simulation vs Model

Algorithm 1 Two Node Simulation Logic

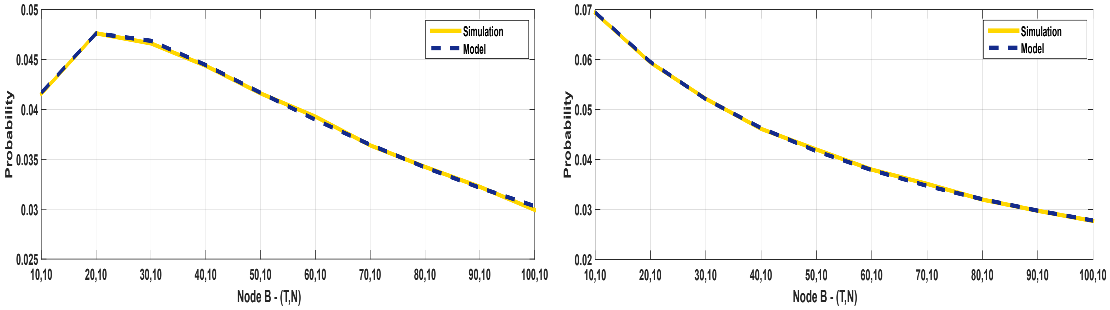
```
Precondition: $\mathbb{T}_A$, $\mathbb{N}_A$, $\mathbb{T}_B$ and $\mathbb{N}_B$ are the exponential random values for a given mean value
  function CONTENTION($\mathbb{T}_A$, $\mathbb{N}_A$, $\mathbb{T}_B$, $\mathbb{N}_B$)
      $\text{No}_{A_\text{Count}}$, $\text{Partial}_{A_\text{Count}}$, $\text{Full}_{A_\text{Count}} \leftarrow 0$
      $\text{No}_{B_\text{Count}}$, $\text{Partial}_{B_\text{Count}}$, $\text{Full}_{B_\text{Count}} \leftarrow 0$
      for $i \leftarrow 1$ to 1000000 do
          if No Contention for $\text{Node}_A$ then
              $\text{No}_{A_\text{Count}} \leftarrow \text{No}_{A_\text{Count}} + 1$
          end if
          if Partial Contention for $\text{Node}_A$ then
              $\text{Partial}_{A_\text{Count}} \leftarrow \text{Partial}_{A_\text{Count}} + 1$
          end if
          if Full Contention for $\text{Node}_A$ then
              $\text{Full}_{A_\text{Count}} \leftarrow \text{Full}_{A_\text{Count}} + 1$
          end if
          if No Contention for $\text{Node}_B$ then
              $\text{No}_{B_\text{Count}} \leftarrow \text{No}_{B_\text{Count}} + 1$
          end if
          if Partial Contention for $\text{Node}_B$ then
              $\text{Partial}_{B_\text{Count}} \leftarrow \text{Partial}_{B_\text{Count}} + 1$
          end if
          if Full Contention for $\text{Node}_B$ then
              $\text{Full}_{B_\text{Count}} \leftarrow \text{Full}_{B_\text{Count}} + 1$
          end if
      end for
      function PROBABILITY(Count)
          return Probability = Count/1000000
      end function
      output Probability
  end function
```

Table 6.1 shows the comparison of probability of different types of contention between simulation and analytical model. Here the MN_A and MN_B's $(\mathbb{T}, \mathbb{N})$ values are $(50, 10)$ and $(10, 10)$ respectively. The table also shows that all the probabilities of MN_A and MN_B sum up to one for both simulation and analytical model. In addition, the probabilities of MN_A and MN_B sum up to one for both the simulation and the analytical model.

Table 6.1: Probability of Contention: Simulation Vs Model; MN_A $(50, 10)$ and MN_B $(10, 10)$.

Contention	**Simulation**		**Model**	
Type	**Node A**	**Node B**	**Node A**	**Node B**
No Contention	0.082815	0.695925	0.083333	0.694444
Partial Contention	0.041925	0.069285	0.041667	0.069444
Full Contention	0.041125	0.068925	0.041668	0.069444
Total	0.165865	0.834135	0.166668	0.833332
Total Probability	Node A+B = 1		Node A+B = 1	

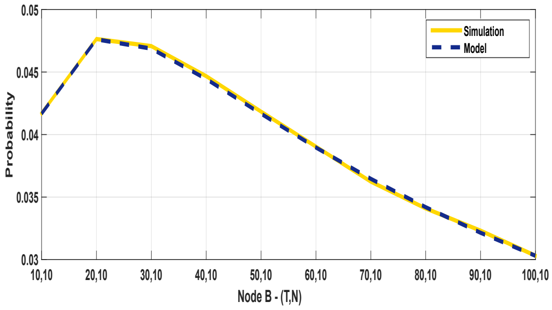
The graphs in Figures 6.2, 6.3 and 6.4 show the comparison of probability of different types of contention between simulation and analytical model for different values of $(\mathbb{T}, \mathbb{N})$ in seconds for MN_B and fixed values for MN_A $(50, 10)$.



(a) MN_A

(b) MN_B

Figure 6.3: Probability of Partial Contention - Simulation vs Model



(a) MN_A (b) MN_B

Figure 6.4: Probability of Full Contention - Simulation vs Model

6.2 Probability Of Contention Between Three MNs Competing for the Same Network

Based on the analysis and results of the Two Node Model, it is possible to extend this approach to the Three Node scenario as disscussed in this section. The probability of contention between three nodes competing for the same network which can be represented in generic tree form as shown in Figure 6.5. The underlying conditions of No Contention, Partial Contention and Full Contention for MN_A are shown in Figure 6.6.

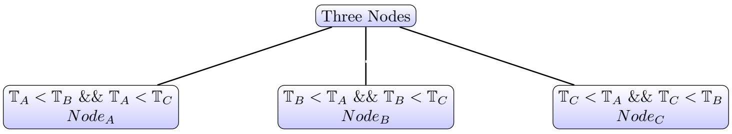


Figure 6.5: Contention Probability Conditions for Three MNs Scenario

For the additional node, MN_C, the distribution of variables is shown in Figure 6.7.

- V is an exponential random variable of $\mathbb{T}_C$
- U is an exponential random variable of $\mathbb{N}_C$

Also, S is the sum of $\mathbb{T}_C$ and $\mathbb{N}_C$ which can be expressed as $S = V + U$. The relevant probability of MN_C is as shown below:

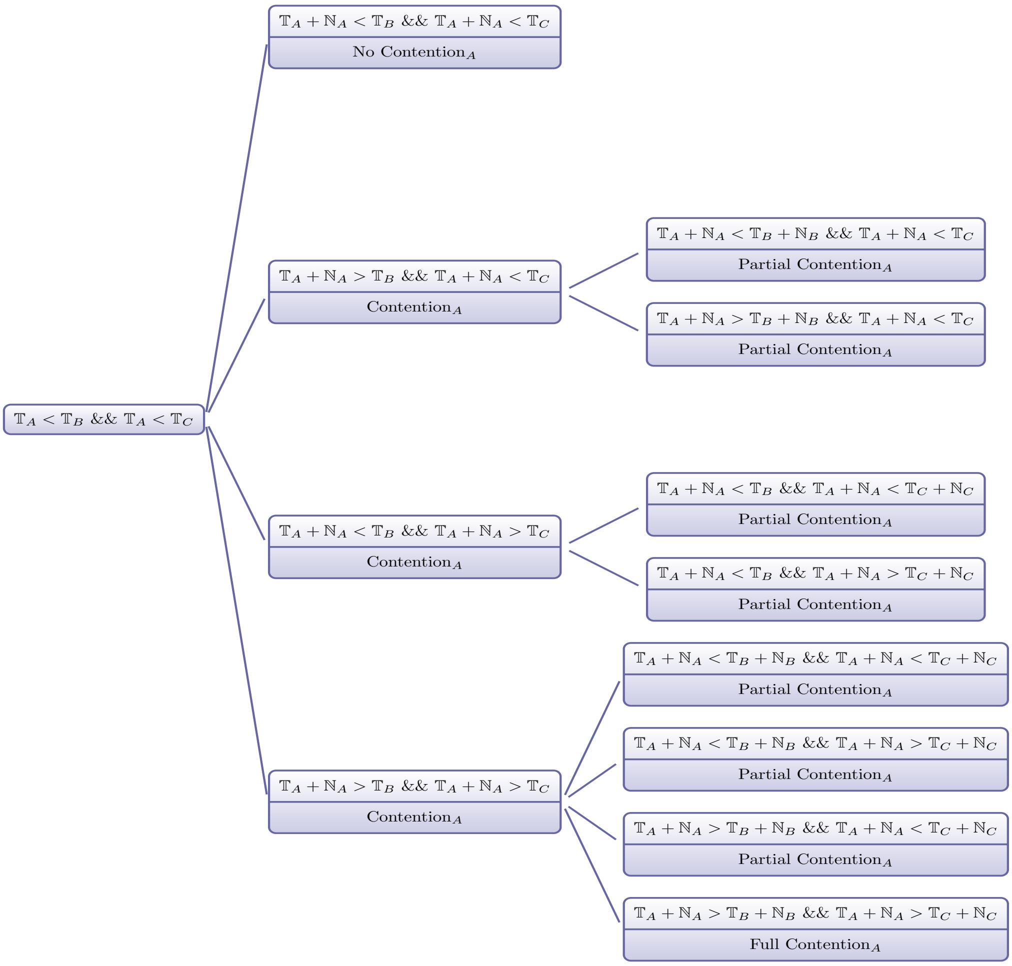


Figure 6.6: Contention Probability Conditions for MN_A

- $P(V = v) = \tau_C e^{-\tau_C v}$
- $P(U = u) = \eta_C e^{-\eta_C u}$

where, τ_C is the reciprocal of the average time that MN_C will take to reach the new network to get the resource and η_C is the reciprocal of the estimated mean resource hold time during dwell time in the new network.

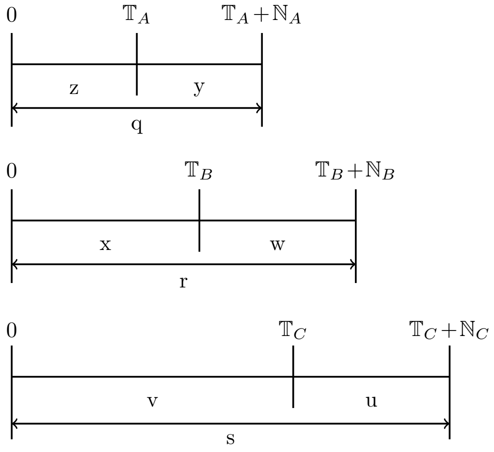


Figure 6.7: Variables for Three Node Scenario

6.2.1 No Contention

According to Figure 6.7, these conditions can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q < x$ and $q < v$

Hence, the probability of No Contention for MN_A is given by:

$$\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \tau_B e^{-\tau_B x} dx \Big) \Big(\int_q^\infty \tau_C e^{-\tau_C v} dv \Big) dz dq \tag{6.7}$$

$$= \frac{\tau_A \eta_A}{(\tau_A + \tau_B + \tau_C)(\eta_A + \tau_B + \tau_C)} \tag{6.8}$$

6.2.2 Partial Contention

For three nodes, there are seven different logical combinations of partial contention conditions as shown in Figure 6.7. All these logical combinations for partial contention are necessary to identify the probability of acquiring any resources during the dwell time of the MN. We will look into all the conditions and derivations to understand the complexity.

6.2.2.1 Partial Contention: Condition 1

According to Figure 6.7, these conditions for Partial Contention Condition 1 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q < v$
- $q < r$ and $q < v$

Hence, the probability of Partial Contention Condition 1 for MN_A is given by:

$$\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \int_z^q \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} \Big(\int_q^\infty \tau_C e^{-\tau_C v} dv \Big) dx dr \Big) dz dq \tag{6.9}$$

$$= \frac{\tau_A \tau_B \eta_A}{(\tau_A + \tau_B + \tau_C)(\eta_A + \tau_B + \tau_C)(\eta_A + \eta_B + \tau_C)} \tag{6.10}$$

6.2.2.2 Partial Contention: Condition 2

According to Figure 6.7, these conditions for Partial Contention Condition 2 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q < v$
- $q > r$ and $q < v$

Hence, the probability of Partial Contention Condition 2 for MN_A is given by:

$$\int_0^\infty \int_0^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} \Big(\int_r^\infty \int_0^x \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \tau_C e^{-\tau_C v} dv \Big) dz dq \Big) dx dr \tag{6.11}$$

$$= \frac{\tau_A \tau_B \eta_A \eta_B}{(\tau_A + \tau_B + \tau_C)(\eta_A + \tau_B + \tau_C)(\eta_A + \eta_B + \tau_C)(\eta_A + \tau_C)} \tag{6.12}$$

6.2.2.3 Partial Contention: Condition 3

According to Figure 6.7, these conditions for Partial Contention Condition 3 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q < x$ and $q > v$
- $q < x$ and $q < s$

Hence, the probability of Partial Contention Condition 3 for MN_A is given by:

$$\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \int_z^q \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C (s-v)} \Big(\int_q^\infty \tau_B e^{-\tau_B x} dx \Big) dv ds \Big) dz dq \tag{6.13}$$

$$= \frac{\tau_A \tau_C \eta_A}{(\tau_A + \tau_B + \tau_C)(\eta_A + \tau_B + \tau_C)(\eta_A + \eta_C + \tau_B)} \tag{6.14}$$

6.2.2.4 Partial Contention: Condition 4

According to Figure 6.7, these conditions for Partial Contention Condition 4 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q < x$ and $q > v$
- $q < x$ and $q > s$

Hence, the probability of Partial Contention Condition 4 for MN_A is given by:

$$\begin{aligned}\int_0^\infty \int_0^s \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C (s-v)} \Big(\int_s^\infty \int_0^v \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \\ \Big(\int_q^\infty \tau_B e^{-\tau_B x} dx \Big) dzdq \Big) dvds\end{aligned} \tag{6.15}$$

$$= \frac{\tau_A \tau_C \eta_A \eta_C}{(\tau_A + \tau_B + \tau_C)(\eta_A + \tau_B + \tau_C)(\eta_A + \eta_C + \tau_B)(\eta_A + \tau_B)} \tag{6.16}$$

6.2.2.5 Partial Contention: Condition 5

According to Figure 6.7, these conditions for Partial Contention Condition 5 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q > v$
- $q < r$ and $q < s$

Hence, the probability of Partial Contention 5 for MN_A is given by:

$$\begin{aligned}\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big(\int_q^\infty \int_z^q \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} \\ \Big(\int_q^\infty \int_z^q \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C (s-v)} dvds \Big) dxdr \Big) dzdq\end{aligned} \tag{6.17}$$

$$= \frac{\tau_A\tau_B\tau_C\eta_A}{(\tau_A+\tau_B+\tau_C)(\eta_A+\tau_B+\tau_C)} \times \frac{(2\eta_A+\eta_B+\eta_C+\tau_B+\tau_C)}{(\eta_A+\eta_B+\tau_C)(\eta_A+\eta_C+\tau_B)(\eta_A+\eta_B+\eta_C)} \tag{6.18}$$

6.2.2.6 Partial Contention: Condition 6

According to Figure 6.7, these conditions for Partial Contention Condition 6 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q > v$
- $q < r$ and $q > s$

Hence, the probability of Partial Contention Condition 6 for MN_A is given by:

$$\begin{gathered}\int_0^\infty \int_0^s \tau_C e^{-\tau_C v}\eta_C e^{-\eta_C(s-v)}\Big(\int_s^\infty \int_0^v \tau_A e^{-\tau_A z}\eta_A e^{-\eta_A(q-z)} \\ \Big(\int_q^\infty \int_z^q \tau_B e^{-\tau_B x}\eta_B e^{-\eta_B(r-x)}dxdr\Big)dzdq\Big)dvds\end{gathered} \tag{6.19}$$

$$\begin{gathered}= \frac{\tau_A\tau_B\tau_C\eta_A\eta_C}{(\eta_A+\tau_B)(\eta_A+\eta_B)}\times \\ \left[\frac{\begin{array}{c}(3\eta_A\tau_B+2\eta_A\tau_C+\eta_B\tau_B+\eta_B\tau_C+\eta_C\tau_B+\eta_C\tau_C \\ +\tau_B\tau_C+3\eta_A^2+\eta_B^2+\tau_B^2+3\eta_A\eta_B+2\eta_A\eta_C+\eta_B\eta_C)\end{array}}{(\eta_A+\eta_B+\eta_C)(\eta_A+\eta_B+\tau_C)(\eta_A+\eta_C+\tau_B)}\right] \\ (\eta_A+\tau_B+\tau_C)(\tau_A+\tau_B+\tau_C)\end{gathered} \tag{6.20}$$

6.2.2.7 Partial Contention: Condition 7

According to Figure 6.7, these conditions for Partial Contention Condition 7 can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q > v$
- $q > r$ and $q < s$

Hence, the probability of Partial Contention Condition 7 for MN_A is given by:

$$\int_0^\infty \int_0^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B(r-x)} \Big(\int_r^\infty \int_0^x \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A(q-z)} \Big(\int_q^\infty \int_z^q \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C(s-v)} dvds \Big) dzdq \Big) dxdr \tag{6.21}$$

$$= \frac{\tau_A\tau_C\tau_B\eta_A\eta_B}{(\eta_A+\tau_C)(\eta_A+\eta_C)} \times \left[\frac{(3\eta_A\tau_C + 2\eta_A\tau_B + \eta_C\tau_C + \eta_C\tau_B + \eta_B\tau_C + \eta_B\tau_B + \tau_C\tau_B + 3\eta_A^2 + \eta_C^2 + \tau_C^2 + 3\eta_A\eta_C + 2\eta_A\eta_B + \eta_C\eta_B)}{(\eta_A+\eta_C+\eta_B)(\eta_A+\eta_C+\tau_B)(\eta_A+\eta_B+\tau_C)(\eta_A+\tau_C+\tau_B)(\tau_A+\tau_C+\tau_B)} \right] \tag{6.22}$$

6.2.3 Full Contention

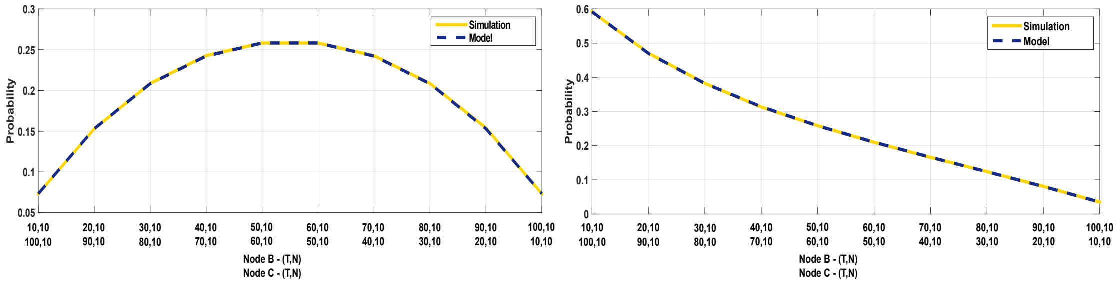
According to Figure 6.7, these conditions can be represented below and are used to generate the limits for the resulting probabilities:

- $z < x$ and $z < v$
- $q > x$ and $q > v$
- $q > r$ and $q > s$

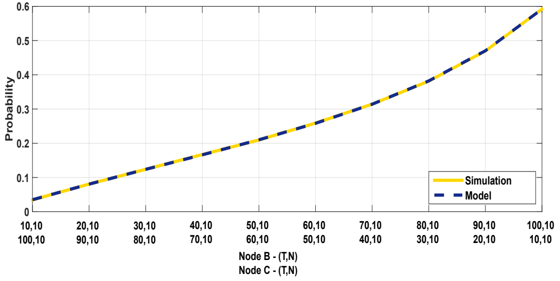
Because, q is both greater than r and s, two conditions must be evaluated and combined: $r > s$ and $s > r$. Hence, the probability of Full Contention for

MN_A is given by:

$$\begin{aligned}
&\int_0^\infty \int_0^q \tau_A e^{-\tau_A z} \eta_A e^{-\eta_A (q-z)} \Big[\Big(\int_z^q \int_z^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} \\
&\Big(\int_z^r \int_z^s \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C (s-v)} dvds \Big) dxdr \Big) + \Big(\int_z^q \int_z^s \tau_C e^{-\tau_C v} \eta_C e^{-\eta_C (s-v)} \\
&\Big(\int_z^s \int_z^r \tau_B e^{-\tau_B x} \eta_B e^{-\eta_B (r-x)} dxdr \Big) dvds \Big) \Big] dzdq
\end{aligned} \tag{6.23}$$



(a) MN_A (b) MN_B



(c) MN_C

Figure 6.8: Probability of No Contention: Three Node - Simulation vs Model

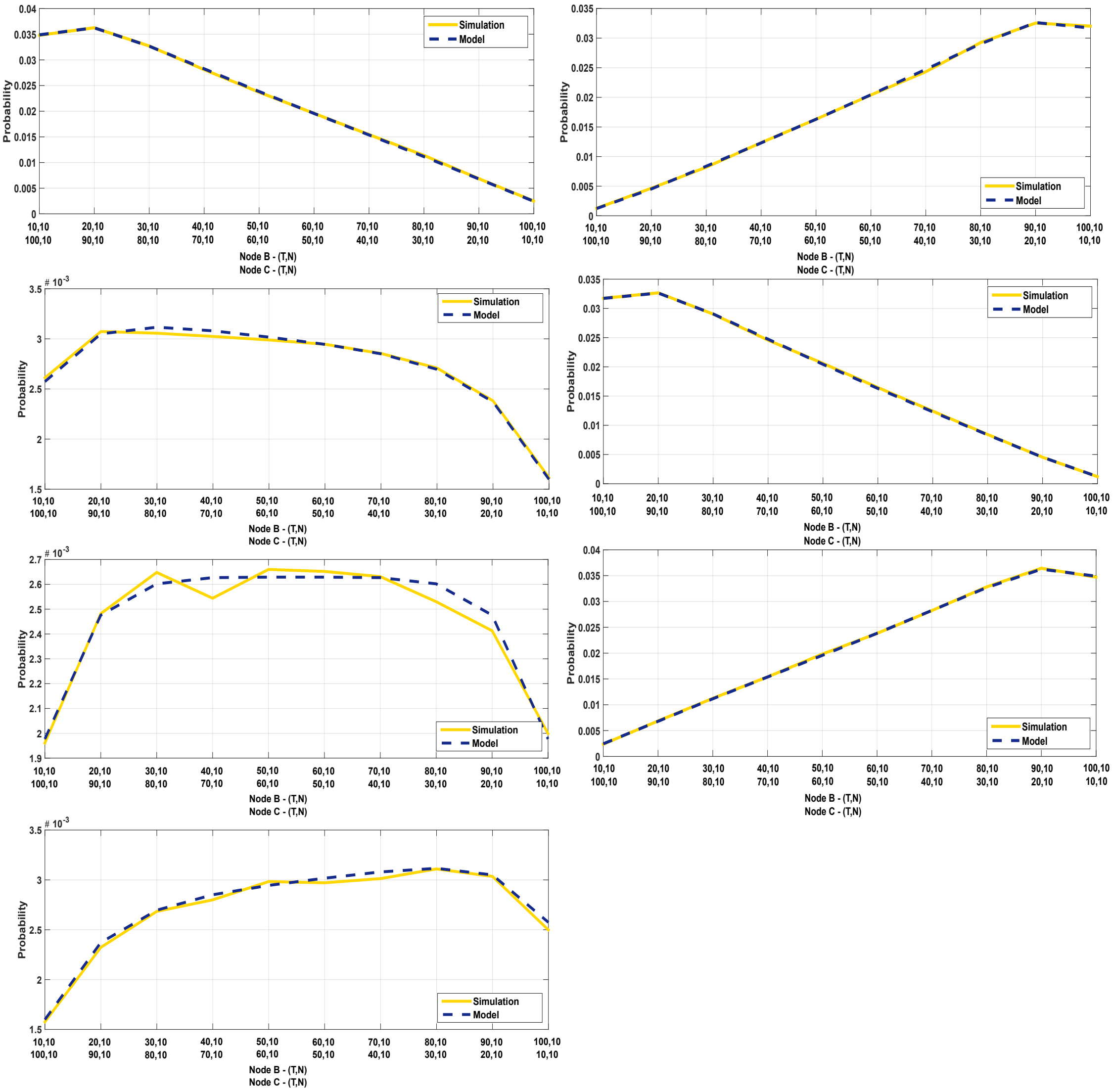
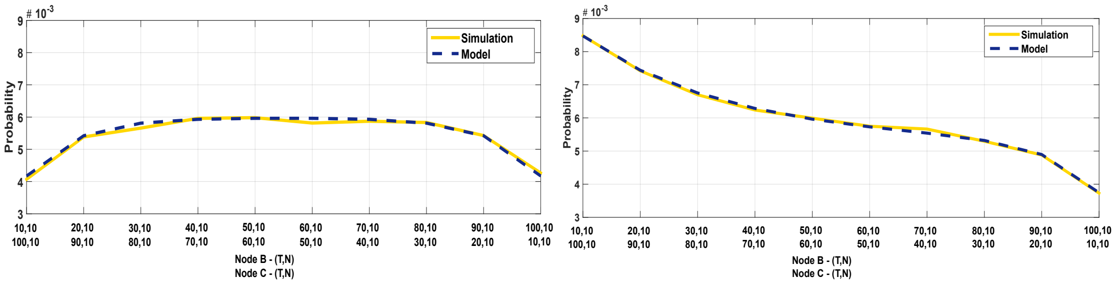
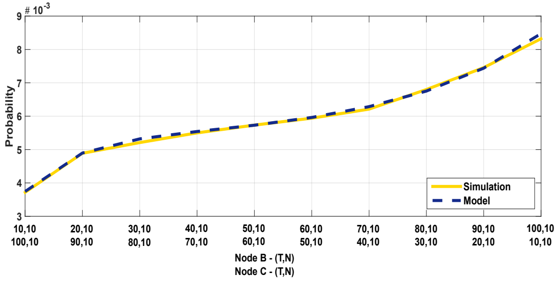


Figure 6.9: Probability of Partial Contention (MN_A): Three Node - Simulation vs Model



(a) MN_A (b) MN_B



(c) MN_C

Figure 6.10: Probability of Full Contention: Three Node - Simulation vs Model

$$\begin{aligned} = \; & \frac{\tau_A \tau_B \tau_C \eta_B \eta_C}{(\eta_A + \eta_B + \eta_C)(\eta_A + \eta_B + \tau_C)(\eta_A + \eta_C + \tau_B)} \times \\ & \qquad (\eta_A + \tau_B + \tau_C)(\tau_A + \tau_B + \tau_C) \\ & \left[\frac{\begin{array}{c}(3\eta_A\tau_B + 2\eta_A\tau_C + \eta_B\tau_B + \eta_B\tau_C + \eta_C\tau_B + \eta_C\tau_C \\ +\tau_B\tau_C + 3\eta_A^2 + \eta_B^2 + \tau_B^2 + 3\eta_A\eta_B + 2\eta_A\eta_C + \eta_B\eta_C)\end{array}}{(\eta_A + \tau_B)(\eta_A + \eta_B)} + \right. \\ & \left. \frac{\begin{array}{c}(3\eta_A\tau_C + 2\eta_A\tau_B + \eta_B\tau_B + \eta_B\tau_C + \eta_C\tau_B + \eta_C\tau_C \\ +\tau_B\tau_C + 3\eta_A^2 + \eta_C^2 + \tau_C^2 + 3\eta_A\eta_C + 2\eta_A\eta_B + \eta_B\eta_C)\end{array}}{(\eta_A + \tau_C)(\eta_A + \eta_C)} \right] \end{aligned} \tag{6.24}$$

Similarly, probabilities of No, Partial & Full Contention can be derived for MN_B and MN_C. The comparison of analytical and simulation models for three-node scenario is as shown in Figures 6.8, 6.9 and 6.10. Results of Partial contention were presented for MN_A. The results of analytical model when compared with the simulation are in good agreement with an error rate of less than 5%.

6.3 Observations From Two and Three Node Scenario

The number of equations or formulas needed to compute the Probability of No, Partial and Full Contention for N number of MNs which total to 1 is:

$$3^{N-1}N \tag{6.25}$$

Furthermore, if the number of possible outcomes is known then the possible number of contention outcome for N nodes is given by:

$$(Number\ of\ Outcomes)^{N-1}N \tag{6.26}$$

A key observation in this model is that by using the process of induction, there appears to be a general formula for calculating the probability of No contention for an arbitrary node MN_x in a network of any number of nodes given by N is as shown in below Equation.

$$P(\text{No Contention}_{\text{MN}_x}) = \frac{\tau_x \eta_x}{\left(\sum_{i=1}^{N} \tau_i\right)\left(\eta_x + \sum_{i=1;\, i \neq x}^{N} \tau_i\right)} \tag{6.27}$$

In the case of Full Contention, for three or more nodes the solution involves a summation of the combinatorial order of the other nodes (For example, MN_B and MN_C) with respect to the node being considered (MN_A) as shown in Equation (6.24). Partial Contention is more complicated as we have to consider all the interactions between the nodes including No Contention, Partial Contention and Full Contention as shown in Figure 6.6. Hence, we can conclude that for N number of nodes, there will be only one No contention and one Full contention for each Node as shown in Figure 6.11. But there are $(3^{N-1}) - 2$ of Partial Contention conditions for each Node i.e., seven Partial Contention conditions for 3 Nodes each. This clearly indicates that the number of Partial Contention conditions rapidly increases with an increasing number of nodes in the system.

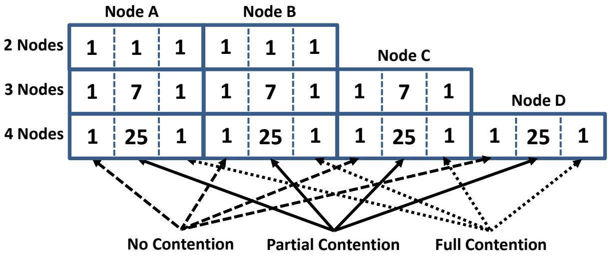


Figure 6.11: Number of Conditions for N Nodes

Even-though as the number of nodes increase the complexity of calculating the probability increases, we have to note that these probabilities are for the nodes competing for resource around the same time. In reality, the number of nodes competing for a resource will not be a lot especially in the case of wireless mobile networks. For example, in a mobile network the number of requests arriving simultaneously at the same time is small, for example, less than five requests. If the algorithm based on the proactive approach will run every 25 milliseconds or 40 times a second, say, then if we can calculate for 4 simultaneous requests, we can serve $40 * 4 = 160$ new calls/connections per second. This would clearly be sufficient for most ingress routers/access points. Therefore, we do not need a model for calculating probability of large number of simultaneous requests to the system. This contention analysis will now be used to develop two new proactive resource allocation queueing models that will take into account the contention between different mobile users.

6.4 Summary

Using probabilistic techniques, the contention among various mobile users in trying to acquire a communication channel in a wireless cell has been modelled for two and three nodes requesting for resource. We find the probabilities of No Contention, Partial Contention and Full Contention. A key observation in this model is that by using the process of induction, there appears to be a general formula for calculating the probability of No contention for an arbitrary node MN_x in a network of any number of nodes given by N. With the probability model to calculate the probability of contention for two and three MNs contending for the resource presented in this Chapter, it now possible to analyse their application on a proactive queueing systems. In order to do so we first need to understand and develop a queueing system for this proactive approach which is detailed in the following Chapter.

Chapter 7

Queuing Models for Proactive Resource Allocation

Queuing models will help us to understand the performance of a wireless system in terms of throughput, mean queue length, blocking probability, etc. Therefore, it is necessary to model a proactive queueing system and compare its performance with the existing classical approach. There are two types of proactive queueing models that have been developed with two key parameters α and β. A detailed analysis of these models are presented and it is shown how these two key parameters can be calculated based on the approach presented in Chapter 7.

7.1 Queuing Model Using Decision System

Proactive resource allocation where the resources will be allocated even before the MN reaches the network will significantly improve the network performance by reducing the contention for resources after reaching the network. In this approach, the decision algorithm (D) based on the contention analysis will decide whether the request for the channel will be admitted to the channel allocation queue. The algorithm will receive the requests and will check for contention among the simultaneous or requests arriving around the same time.

The service rate is, μ_s and the requests in the queue can also leave the system due to mobility, μ_m as shown in Figure 7.1. That is the request in channel queue

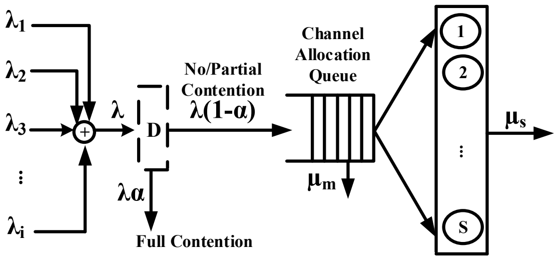


Figure 7.1: Proactive Resource Allocation Queuing System with α

of an AP can move out of the coverage while waiting for the channel due to mobility. α is the percentage of calls dropped due to Full Contention when there are a number of requests around the same time. For example, if two or three nodes are contending for the channel with their respective $\mathbb{T}$ and $\mathbb{N}$ and if there is a full contention among the nodes then a node having a high probability of Full contention with other nodes will be dropped. Therefore, α can be calculated using the Full Contention model presented in the previous Chapter. Hence, the effective arrival rate (λ_{eff}) to the queue is as shown in Equation 7.1. If α is independent of ($\mu_s + \mu_m$) the queue can be treated as a normal M/M/1 queue with finite buffer.

$$\lambda_{eff} = \lambda(1 - \alpha) \tag{7.1}$$

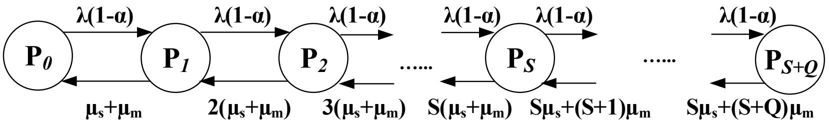


Figure 7.2: Proactive Resource Allocation State Diagram with α

The state diagram of the proposed proactive resource allocation using a decision system is shown in Figure 7.2. The proactive approach is therefore able to capture the circumstances of Full Contention, α, before the MN reaches the network coverage and hence MNs can be signalled that they will not get the channel and thus can immediately do a vertical handover to another network rather than waiting for a channel which they would never be allocated, leading to better overall network performance.

7.2 Queuing Model Using Decision System and Contention Queue

Though the proactive queuing model presented above showed great improvement over the classical approach, it did not fully take into account the different contention scenarios. In the previous model, requests admitted by the decision algorithm cannot be queued in the actual channel allocation queue of the next network as the MN has not reached that network and therefore we need a separate queue. Hence, another queue is introduced before the channel request queue called the contention queue as shown in Figure 7.3. In the contention queue we explore all the possibilities of contention interaction i.e., No Contention, Partial Contention and Full Contention with current and subsequent requests. These interactions may result in a request leaving the contention queue and hence, the overall system due to Full Contention with a subsequent request or the request in queue may be rearranged due to Partial Contention. The contention queue only queues requests before the MNs reach the next network. Once the MN reaches the relevant network, its request in the contention queue will be placed in the channel allocation queue. Therefore, because the contention queue is used before the MNs reach the next network, originating calls in the next network must be placed directly in the channel queue and not the contention queue.

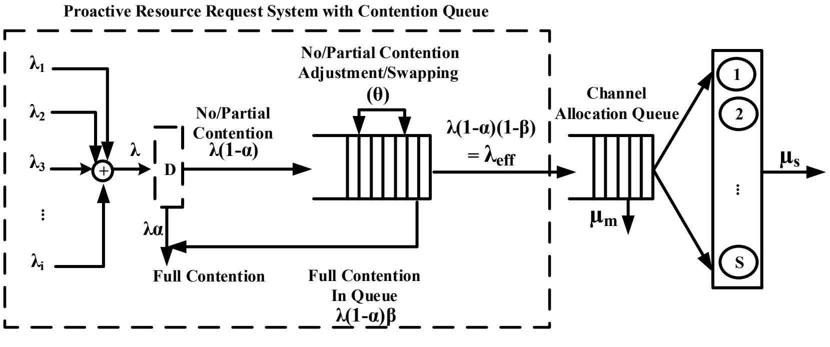


Figure 7.3: Proactive Resource Allocation Queuing System with α and β

In the proposed approach, the decision algorithm decides whether the MN's request will be admitted to the system based on values of $\mathbb{T}$ and $\mathbb{N}$ of all the nodes requesting a channel. There are three possible contention possibilities that affect mobile users in this scenario (No, Partial and Full Contention). Let us consider a simple queue and requests with ($\mathbb{T}$,$\mathbb{N}$) arrive to the queue. If there is no contention with the requests already in the queue then the new request will be placed in the queue. The position of each request will depend on the values of $\mathbb{T}$ and $\mathbb{N}$ i.e., the requests are placed in ascending order with the least $\mathbb{T}$ at first position of the queue as it will be reaching the new network first. On contrary, if the new request encounters a partial contention then the $\mathbb{N}$ value of the requests with the higher $\mathbb{T}$ value will be modified and the request with less $\mathbb{T}$ value will be placed before the modified request. This phenomenon of addition, swapping or re-arrangement in the queue due to No contention or Partial contention is represented as θ as shown in Figure 7.3. Since θ does not involve a request leaving the contention queue therefore, it will have no overall effect with respect to the rate of transfer of requests to the channel allocation queue.

When a new request is being placed in the contention queue, it is compared with each entry in the queue consecutively. If there is a full contention between the new request and a given entry then the one with the higher $\mathbb{T}$ will be dropped, because the higher value of $\mathbb{T}$ means that the MN needs the channel later than the request with less $\mathbb{T}$. Therefore, a new request could be dropped due to Full Contention with the entry in the contention queue or the entry in the contention queue could be dropped due to Full Contention with the incoming request. Both these scenarios result in one of the requests leaving the contention queue and hence, we denote the rate at which requests leave the queue due to Full Contention is denoted as β.

7.2.1 Request Management in the Contention Queue

An example of the above approach is shown in Figure 7.4. In the proposed approach, the decision algorithm decides whether the MN's request will be admitted to the system based on values of $\mathbb{T}$ and $\mathbb{N}$ of all the nodes requesting a channel. There are three possible contention possibilities that affect mobile users in this scenario (No, Partial and Full Contention). Let us consider a simple queue and requests with ($\mathbb{T}$,$\mathbb{N}$) arrive to the queue as shown in Fig. 7.4. Req_A arrives with (10,10) in seconds to the contention queue, the decision algorithm checks the queue and Req_A is queued at the front as the queue is empty.

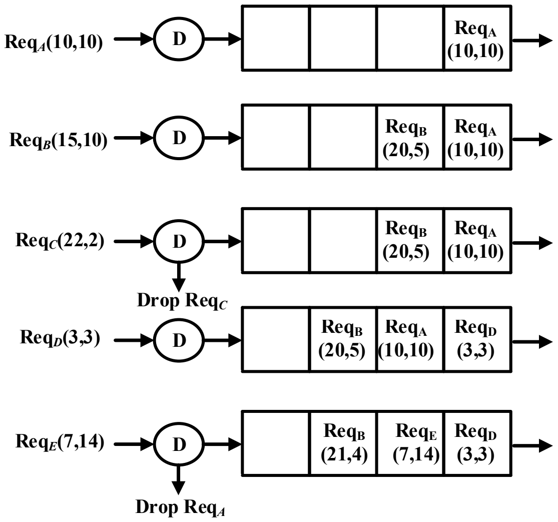


Figure 7.4: Contention Queue Management

Now Req_B arrives with (15,10) and the time when it needs the channel and is when Req_A will still be using the channel resulting in Partial Contention for Req_B. Because, Req_A release the channel at the end of 20. Therefore, Req_B is modified to $\mathbb{T_B} = 20$ and $\mathbb{N_B}$ is modified to $(15 + 10) - 20$ hence, the modified request for Req_B is (20,5).

Req_C now arrives with (22,2) however Req_B will release the channel at 25 which means that Req_C will never get the channel and therefore it is not admitted to the contention queue due to Full Contention. A rejection reply is sent to MN_C causing it to do an immediate handover to another network.

Now Req_D arrives with (3,3), there is No Contention with Req_A or Req_B and therefore, it is placed at the head of the queue since $\mathbb{T_C} + \mathbb{N_C} < \mathbb{T_A}$ i.e., $6 < 10$.

Req_E arrives with (7,14) this results in No Contention with Req_D, Full Contention with Req_A and hence Req_A is ejected from the contention queue because $\mathbb{T_E} < \mathbb{T_A}$ and $\mathbb{T_E} + \mathbb{N_E} > \mathbb{T_A} + \mathbb{N_A}$. Req_B experiences Partial Contention and hence the request is modified accordingly.

7.2.2 New Proactive Markov Queuing Model

The rate at which the incoming request join the contention queue is given by $\lambda(1-\alpha)$ as shown in Figure 7.3. As the requests are rejected from entering the system (α) and requests in the contention queue can also be removed (β) due to Full Contention with subsequent requests. The removing rate from the contention queue is given as λ(1-α)β as shown in Figure 7.3. Thus, scheduling and arrangement of requests take place and the total effective arrival rate to the channel queue is given by:

$$\lambda_{eff} = \lambda(1-\alpha) * (1-\beta) \tag{7.2}$$

Similar to the previous queuing models, the proposed proactive resource allocation queuing model for handover considers S number of channels and can allow i requests at time t as shown in Figure 7.3. Q is the queueing capacity of the proposed system. Figure 7.5 shows the state diagram of the proposed model. Let's define the states i (i=0,1,2,$\cdots$,S+Q) as the number of requests in the system at time t.

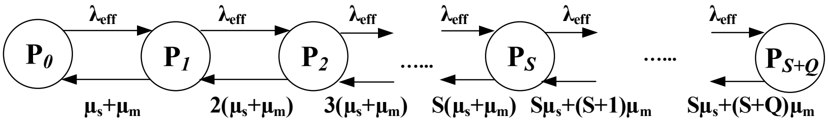


Figure 7.5: Proactive Resource Allocation State Diagram with α and β

In proposed model, $\mathbb{T}$ is exponentially distributed with a mean rate of μ_s is assumed. In addition, $\mathbb{N}$ is also exponentially distributed with a mean rate of μ_m. As explained in Chapter 5, μ_m, the dwell time in wireless and mobile systems can be calculated using Equation (4.26) and μ_i can be calculated using Equation (4.27)

ρ is the traffic intensity in the system, where $\rho = \lambda_{eff}/(\mu_s + \mu_m)$. Assuming a system in a steady state, the state probabilities, P_i's, can be obtained as in

Equation 7.3.

$$P_i = \begin{cases} \frac{\rho^i}{i!} \cdot P_0 & 0 \leq i \leq S \\ \frac{\frac{\rho^S}{S!} \cdot \lambda_{eff}^{i-S} \cdot P_0}{\prod\limits_{j=S+1}^{i} [S\mu_s + (j-S)\mu_m]} & S < i \leq S+Q \end{cases} \tag{7.3}$$

In Equation 7.3, P_i is the probability that there are i calls in the system. P_0 can be defined as follows:

$$P_0 = \left[\sum_{i=0}^{S} \frac{\rho^i}{i!} + \sum_{i=S+1}^{S+Q} \frac{\frac{\rho^S}{S!} \cdot \lambda_{eff}^{i-S}}{\prod\limits_{j=S+1}^{i} [S\mu_s + (j-S)\mu_m]} \right]^{-1} \tag{7.4}$$

The mean queue length (MQL), the throughput (γ) and mean response time (MRT) of the system can be calculated using Equations (4.29), and (4.30) respectively as shown in Chapter 5.

7.3 Results and Discussion

In order to better explore the effects of the new proactive models and to show that they improve resource allocation in highly mobile environments, in this section we present the results of a detailed discrete event simulation with various values of α and β. However, in the next section we show how α and β can be calculated for a real network.

The system parameters used are mainly taken from (Kirsal et al., 2010, 2013) based on the relevant literature (Gemikonakli et al., 2013; Ghosh et al., 2014b; Kirsal et al., 2010, 2013). The system has a fixed number of identical channels: S=12. Q is the queuing capacity, which represents the number of requests waiting for service and is limited with Q=100. The average speed of the mobile user and the radius of the network are taken as 30km/h (≈19Mph) and 1000m for all cal-

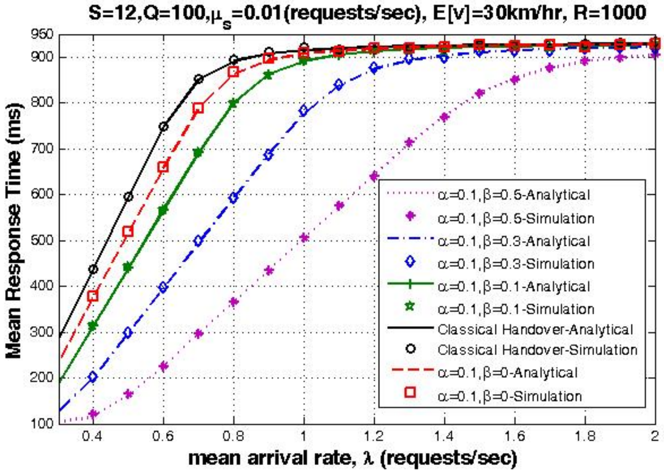


Figure 7.6: MRT results as a function of arrival rates

culations, respectively. The rates are translated into requests per second in order to use consistent values. The service rate of the system μ_s is 0.01 requests/sec.

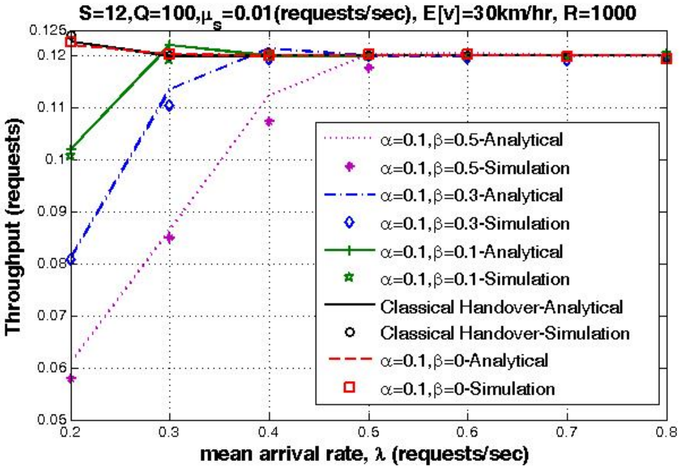


Figure 7.7: Throughput results as a function of arrival rates

The results comparing the classical resource allocation queuing model, the proactive resource allocation queuing model without contention queue (hence, $\beta = 0$) and the proactive resource allocation queuing model with contention queue (with different β) are shown in Figures 7.6 and 7.7. Figure 7.6 shows MRT result as a function of λ with different β values. In heterogeneous mobile environments, response time is another important QoS parameter in order to obtain best QoS. It can be clearly seen that MRT increases rapidly for the classical handover model when λ increases due to the number of unnecessary requests allowed into the system. Such requests, especially handovers will leave the system without being served. Thus, MRT increases. In addition, the MRT results clearly show that the proactive approaches works far better than the classical approach.

For the throughput results shown in Figure 7.7, there is a drop in the throughput of the given network as more requests are removed from the system due to contention, however, this will improve the overall network efficiency as these requests can be dealt with by other networks in a heterogeneous environment. Hence, all results show that, the proactive resource allocation queuing model increases the overall network efficiency in such environments.

7.4 Effects of α and β

The effects of α and β on the system are demonstrated in this section. The parameters taken are the same as described in the previous section with a change of μ_s=0.02(requests/sec) and $E[v]$=50km/s.

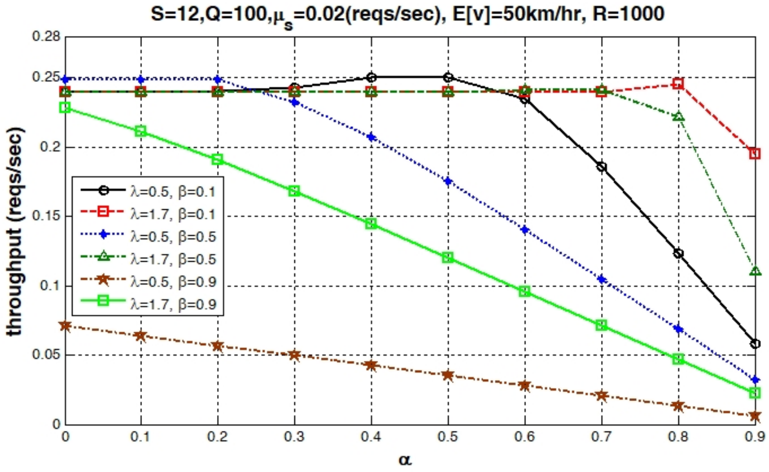


Figure 7.8: THRP results as a function of α for different λ

The throughput results as a function of α with different β are considered in Figure 7.8 for light and heavy traffic loads (e.g., λ=0.5 and 1.7). The results indicate the drop in the throughput of the given both traffic loads since more requests are removed from the system due to contention. In other words, the throughput decreases as MNs leave the system considering the contention probabilities. However, this improves the overall network performance as these requests can be dealt with by other networks.

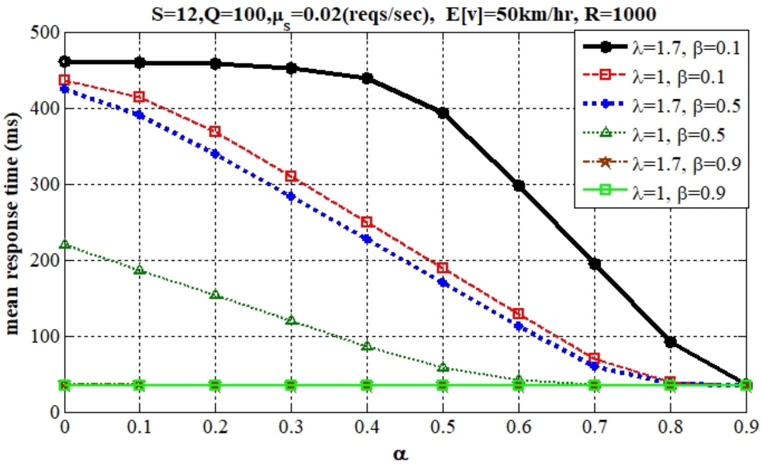
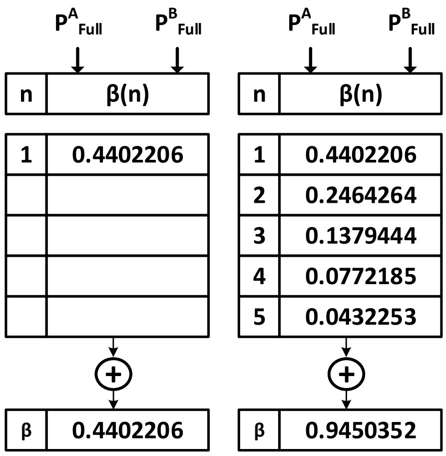


Figure 7.9: MRT results as a function of for different λ

The Figure 7.9 shows MRT results as a function of α considering various full contention probabilities occurring in the queue (β) for different traffic loads (e.g., λ=1, 1.7). As clearly seen from the figure, the best MRT results can be obtained for high value of β. When β=0.9, all MNs can get a channel regardless of α. Even though the system has moderate traffic loads (e.g., λ=1), the Full Contention probability in the queue is an important parameter of getting a resource in such system. It is clear from all the figures that α and β affect the system performance significantly. α and β are important parameters for the decision management of acquiring a resource efficiently.

7.5 Calculating α and β

In the queueing analysis presented in Section 8.3, α and β values are assumed. In order to utilize the proposed proactive approach, a model has to be developed to calculate α and β. To calculate α, we observe that α is the probability of full contention (P_{Full}) and it is important to understand that for α we are dealing with the number of simultaneous requests before entering the contention queue. As explained in the previous Chapter, in a mobile network the number of request arriving simultaneously at the same time is relatively small.Therefore, for α we do not need a model for calculating probability of a large number of simultaneous requests to the system.



n	β(n)
1	0.4402206
β	0.4402206

n	β(n)
1	0.4402206
2	0.2464264
3	0.1379444
4	0.0772185
5	0.0432253
β	0.9450352

Figure 7.10: β Calculation

To calculate β, we must use the two-node model because β represents the probability of the request leaving the contention queue due to contention with the incoming request or vice-versa. So, the first request A, is an entry that is already in the queue and the second request is the incoming request B. We need to consecutively compare each entry in the table with the incoming request. We are interested if the nth entry in the queue or the incoming request is ejected due to full contention. This is given by $P^A_{Full} + P^B_{Full}$ and the probability that either occurs at the nth entry is:

$$\beta(n) = \{1 - (P^A_{Full} + P^B_{Full})\}^{n-1} \times \{P^A_{Full} + P^B_{Full}\} \tag{7.5}$$

Therefore, β is the summation of all $\beta(n)$ probabilities from n = 1 to N, where N is the number of requests in the queue as shown in Figure 7.10.

For example, let us consider $\mathbb{T}$ for MN_A is 10 and MN_B is 60. $\mathbb{N}$ for MN_A is 20 and MN_B is 30. The resulting probability of Full Contention based on our two node model: $P^A_{Full} = 0.085714$ and $P^B_{Full} = 0.064286$. β value for $N = 5$ is 0.556, $N = 20$ is 0.961 and $N = 80$ is 0.999. Hence, the two node model is sufficient to yield significant results. The application of β will be demonstrated on the results acquired from the testbed in the following Chapter.

7.6 Probability of a Request Staying in the Contention Queue

We know from the standard queueing analysis for a system of k nodes:

$$P_n = \rho^n \frac{1-\rho}{1-\rho^{k+1}} \tag{7.6}$$

Where, ρ is the utilisation (λ/μ), k is the size of queue and n is the number of request in the system.

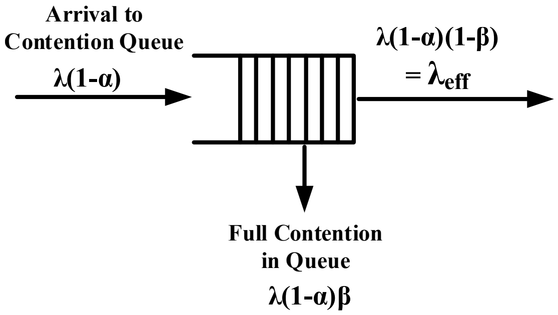


Figure 7.11: Proactive Contention Queue

We can assume that the proactive contention queue as a simple M/M/1/k queue. Here, we are interested only to find out the probability of a given request staying in the contention queue and not be affected due to β. Here the arrival rate to the contention queue is $\lambda(1-\alpha)$ and service rate is β as we are only interested in the requests affected by β as shown in Figure 7.11. Therefore, ρ can be represented as:

$$\rho = \frac{\lambda(1-\alpha)}{\beta} \tag{7.7}$$

Therefore,

$$P_{n-1} = \rho^{n-1} \frac{1-\rho}{1-\rho^{k+1}} \tag{7.8}$$

Here, the number of requests in the system becomes $n - 1$ as there are no entries in the server.

In order to find the probability of a given request staying in the contention queue without being affected by β, A request needs to satisfy three conditions:

- The new incoming request occupied the N^{th} position in the queue, i.e., last position of the queue.
- New incoming request is having a full contention with an entry at the N^{th} position and pushing that entry our of the queue.
- New request after occupying a place in the queue is not affected by the subsequent new incoming request.

Here, let us assume that P^A_{Full} is the probability of full contention for incoming request A. P^B_{Full} is the probability of full contention for the requests in contention queue and here an average velocity is considered. P^C_{Full} is the probability of full contention for a new subsequent incoming request C after A.

- $(1 - P^B_{Full})^{N-1}$ is the probability that request A occupies the N^{th} position in the contention queue and it is not colliding with any other request before it.
- $(1 - P^A_{Full})^{N-1} P^A_{Full}$ is the probability that request A do not collide with the request in front of N^{th} position but does collide with the request already occupying the N^{th} position.
- $1 - [(1 - P^C_{Full})^{N-1} P^C_{Full}]$ is the probability of the request A not getting booted out by a new subsequent request entering the queue i,e C.

Therefore, all these three conditions can be represented as shown in Equation (7.9) to calculate the probability of a request staying in the queue and not been affected by β.

$$\sum_{n=1}^{n=k}\Big\{\Big[\Big(\frac{(\rho^{n-1})(1-\rho)}{1-\rho^{k+1}}(1-P^{B}_{Full})^{n-1}\Big)+ \Big(\sum_{m=n}^{m=k}\frac{\rho^{m}(1-\rho)}{1-\rho^{k+1}}(1-P^{A}_{Full})^{m-1}P^{A}_{Full}\Big)\Big]\times \Big(1-(1-P^{C}_{Full})^{n-1}P^{C}_{Full}\Big)\Big\}\times(1-P^{B}_{Full}) \tag{7.9}$$

For example, let us consider $\mathbb{T}$ for MN_A is 10 and MN_B is 60. $\mathbb{N}$ for MN_A is 20 and MN_B is 30. The resulting probability of Full Contention based on our two node model: $P^{A}_{Full} = 0.085714$, $P^{B}_{Full} = 0.064286$ and let us assume $P^{C}_{Full} = 0.2$. Probability of a request not affected by β for $k = 1$ is 0.68635, $k = 20$ is 0.45405 and $N = 80$ is 0.75405. The application of this model will be demonstrated on the results acquired from the testbed in the following Chapter.

7.7 Summary

A thorough analysis of queueing models for a proactive system where the users contend for resources has been presented in this Chapter. These models have yielded two key parameters, α and β, that have to be calculated to effectively use the proposed proactive approach. It has been shown that if these two parameters are identified then we can ensure that each user in the system will be effectively served by at least one of the available network. This can be achieved by using vertical handover to alternative network if an user is about to experience a full contention in the target network. To the best of our knowledge there is no study that has modelled a proactive model in the context of users contending for resources based on their mobility. With the results showing that the proposed approach is outperforming the classical model, it is now necessary to understand the application of proposed approach in a highly mobile real network such as VANET.

Chapter 8

Application to VANET systems

The main objective of this Chapter is to illustrate the designing and implementation of VANET Testbed at Middlesex University, London. In addition, to demonstrate the application of the proactive models presented in this work in a real network, such as the VANET testbed at Middlesex University, London. The coverage range, viability and performance of long range transmissions on the 5.9 GHz frequency spectrum are also presented.

8.1 Designing and Implementation of VANET Testbed

In this section, a step by step account detailing of how the testbed deployment at Middlesex University was carried out, is presented. The testbed has seven RSUs; four on the university campus buildings and three along the A41 road behind the university. RSUs and OBUs were manufactured by Lear Corp. (ARA, 2016) in compliance with the IEEE 802.11p (WAVE) standard specifications and the maximum output power used was 200mW or +23dBm. The operating frequency of the RSUs is 5.9 GHz and the channel being used to send broadcasts is CH172. The table below shows the location names and their GPS coordinates.

Table 8.1: RSU Location Information.

RSU	Location	Latitude	Longitude
1	Grove Building	51.588837	-0.230769
2	Sheppard Building	51.590808	-0.229672
3	Williams Building	51.590494	-0.228594
4	Hatchcroft Building	51.589073	-0.228374
5	Lamp Post (No.6)	51.586145	-0.231452
6	Lamp Post (No.42WW)	51.591779	-0.234701
7	Lamp Post (No.87WW)	51.599255	-0.233245

8.1.1 MDX Deployment

In order to determine the best location for the RSU on the Middlesex University buildings, it was important to minimise the distance between the RSU and the router elements in the university network. This enabled us to directly backhaul data from the RSU to the central MDX VANET Server located at the basement of Sheppard library using the university network as shown in Figure 8.1. Four RSUs have been deployed on top of the Hatchcroft building, Williams building, Sheppard library building and Grove building to cover the roads around the campus and to support the movement of pedestrians within the campus, hence enabling the development of Vehicle-to-Pedestrian (V2P) applications.

8.1.2 A41 Deployment

Working with Transport for London (TfL), three RSUs were mounted on lampposts along Watford Way along the A41 road behind the university. This was deployed to substantially improve the coverage of the Middlesex testbed, which is being used to develop better propagation models in an attempt to understand the best way of deploying a citywide VANET system. Due to the lack of any existing communication network on the lamppost, the data received by the RSU was backhauled using an LTE router. In addition, due to security restrictions of the Middlesex network a Virtual Private Network (VPN) tunnel service provided by Mobius network via Vodafone was used as shown in Figure 8.1.

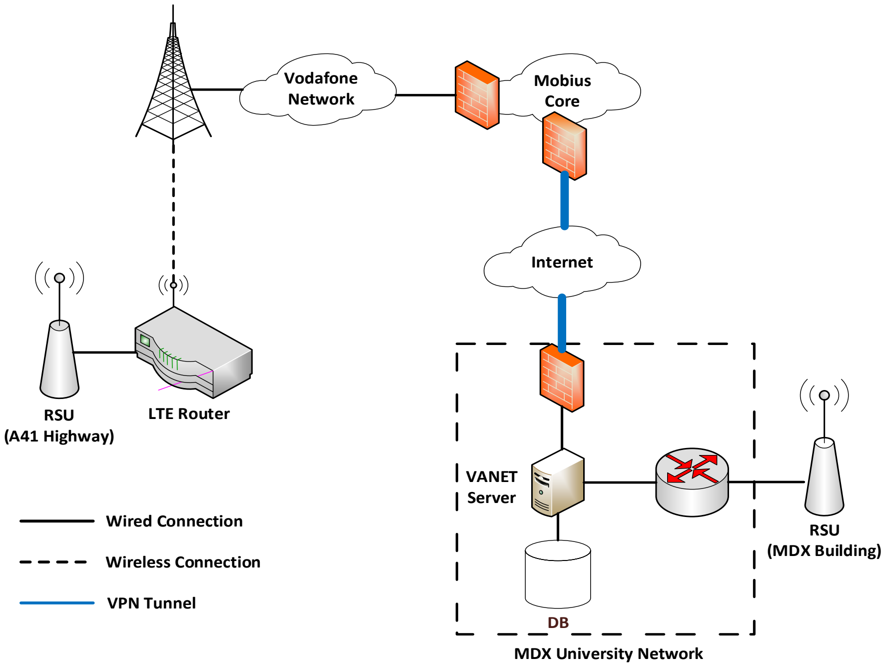


Figure 8.1: Network Diagram.

8.1.3 Application Description

There are different types of messages used to communicate various items of information between OBUs and RSUs such as CAMs, BSMs, Road Side Alerts (RSAs), Intersection Collision Alerts (ICAs) and Probe Vehicle Data (PVD). Since we are interested in safety applications, the BSM message was used as the first message to be collected for analysis. BSMs are periodically broadcast (10 Hz) from the OBU to the RSU and a BSM (41 bytes of length) contains information regarding position, motion, time, and general status of the vehicle as shown in the Figure 8.2 (SAEInternational, 2010). At the present time, we are able to only use the Message ID and 3D Position i.e., Latitude, Longitude and Elevation parameters of the BSM message. This is because all other fields such as steering wheel angle, acceleration set, brake system etc., have to be gathered from the vehicle via sensors or other mechanical devices and then added to the BSM packet and broadcast. This is planned for the next phase of the project and it is not in the scope of this thesis.

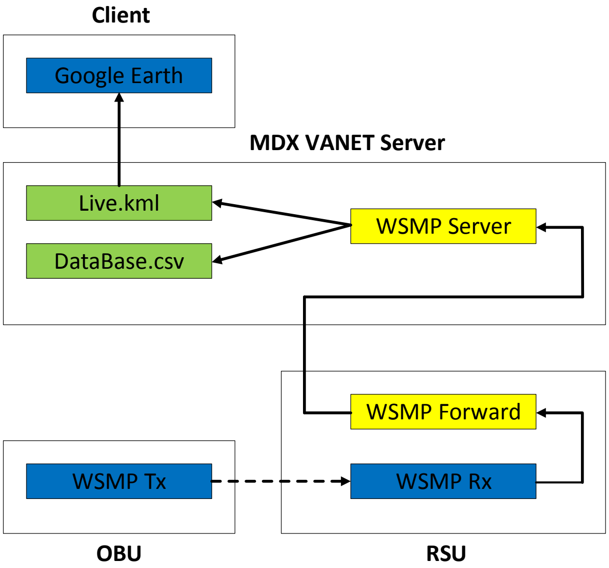


Figure 8.2: Packet forwarding from RSU to Server.

Wave Short Message Protocol (WSMP) Tx is an application used by the OBU to broadcast the BSM packets and the RSU receives these packets using WSMP Rx application. The received packets are forwarded to the server using the WSMP Forward application via an IPv6 address of the server as shown in Figure 8.2. The MDX VANET Server uses a WSMP Server application to receive the packets and save the data. At this stage, additional information such as a timestamp and the RSUs IP address are stored along with the message received. The received data was saved in two different files: Live.kml and Database.csv, where, live.kml contains the live or current positions of each OBU through the packets received from those OBU and this file is saved in the Apache Web Server space for remote access. Using Google Earth and by adding a network link to the live.kml file, the live tracking of the OBUs was achieved. The second file Database.csv contains the most of the available information in the packets such as the OBUs MAC address, the received signal strength indicator (RSSI) value of the received packet, GPS coordinates along with the time stamp of the packet and IP address of the RSU by which the packet has been forwarded. Every day the Database.csv file was backed up for analysis through MySQL.

8.2 MDX and A41 Coverage Map Result

Analysis and comparison of Free Space Path Loss (FSPL) propagations model were performed to understand the signal propagation behaviour. As explained in previous chapters a concrete Propagation modelling is necessary to accurately estimate the NDD which inturn will be used to estimate different times which is required to design proactive handover and resource allocation. We know that performance of several propagation models depends on the environment and characteristics of the antenna itself. Since, the scope of this research is not about the propagation models, a simple FSPL model is considered to showcase the importance and necessity of exploring other propagation models that will suit VANET scenarios. This will also enable us to develop an efficient deployment strategy for RSU.

Figure 8.3 shows the unique GPS coordinates from the packets received by the MDX VANET Server, sent by the OBUs which were placed in a car. Figure 8.3 displays the trial data for a car driving on the roads around the Middlesex University campus to map the coverage which was collected on 17th May 2017. The coverage map shows the individual coverage achieved by the RSU located on each building and the RSUs located along the A41 road with different colour dots.

The first four (1-4) RSUs cover the Hendon Campus and surrounding roads. This covers an area of around 0.7 miles or 1.1 kms. The A41 is covered by the other (5-7) RSUs. The coverage runs from the entrance of the Great Northern Way (top of brown line) to Hendon Central Tube Station (bottom of the blue line). This is a distance of 2 miles or 3.2 kms. Hence the total coverage of the testbed is 2.7 miles or 4.31 kms. This Middlesex testbed is interesting because it encompasses motorway as well as urban roads and thus can allow us to do a comprehensive study of signal propagation for RSUs deployed along the road-side on the lampost and on buildings i.e., MDX campus buildings.

We can observe that the more coverage is achieved for the RSUs deployed at higher heights and also which have clear Line-of-Sight (LoS) in relation to the intended roads (Gozalvez et al., 2012). The coverage was better than anticipated but this was mainly because of the height of the RSU deployment. Some blind

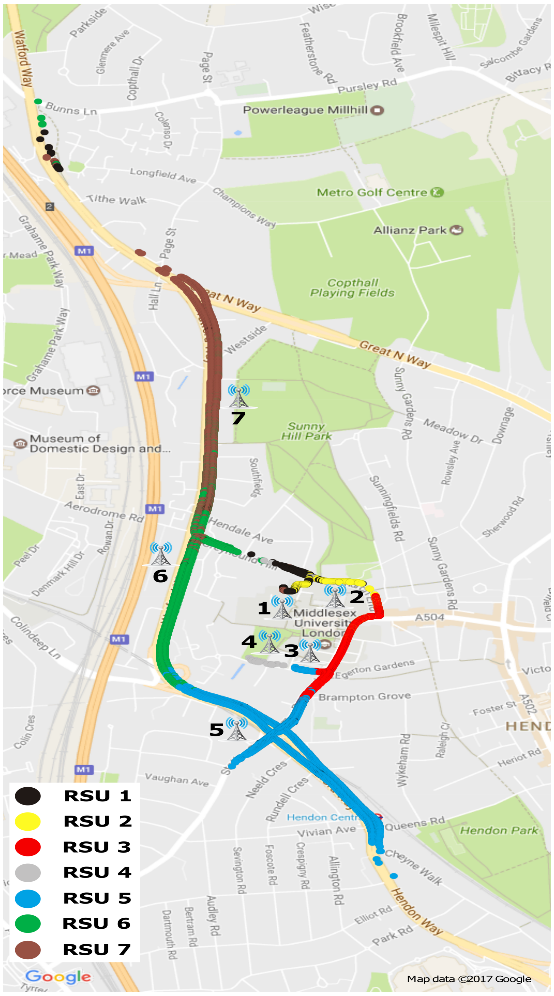
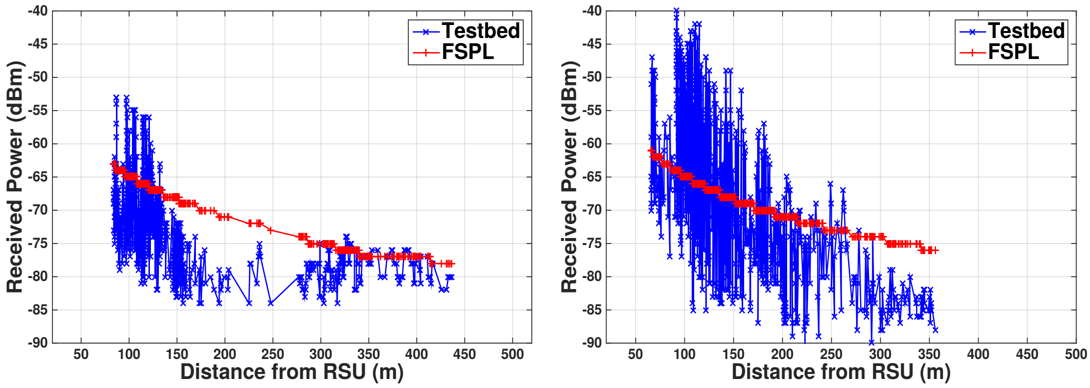
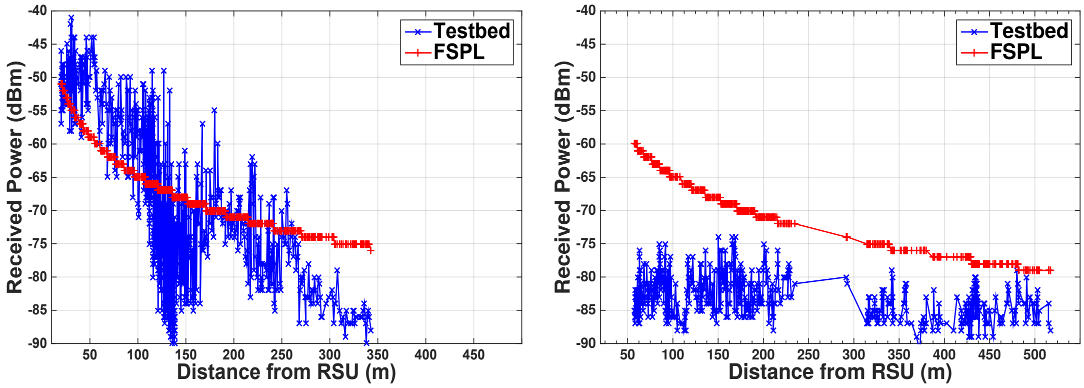


Figure 8.3: Coverage Map.

spots can be observed; these were purely due to effects of surrounding buildings. From these observations it is clear that we need more RSUs alongside the road to be deployed in an urban area. However, from our results, the deployment of RSU on high buildings and cellular masts should also be explored. The dense lines indicate very reliable communication and single spots indicate only few packets were received and hence there is no continuous communication in these regions. For example, the coverage map of RSU deployed at Grove building has almost no coverage on the A504 due to the blockade of signal propagation because of the nearby buildings. The farthest point from where the packets were sent by



(a) RSU 1 - Williams Building (b) RSU 2 - Sheppard Building



(c) RSU 3 - Hatchcroft Building (d) RSU 4 - Grove Building

Figure 8.4: Power Received for Each Packet - MDX Testbed vs FSPL.

the vehicles and successfully received by the RSU was approximately 1.15 Km from the Williams Building RSU to a point on the East side of Watford way close to where the A41 and M1 divide. This was achieved purely due to the very high elevation of the RSU on the Williams Building, hence allowing LoS communication over a great distance.

These readings are omnidirectional with respect to the RSU, so only the radius of the RSU was considered, so that better readings were due to the LoS and multi-path. For Figure 8.4 the detailed analysis as follows.

Figure 8.4a RSU 1:- Here, we again see that there is a good coverage close to the RSU, however there is a large drop due to the steep decline in the height of the road going down the Greyhound Hill. In comparison, more readings are obtained at the traffic lights further down the Greyhound Hill and along the Watford Way

(A41).

Figure 8.4b RSU 2:- This gave the most consistent readings relative to the FSPL model. Therefore, signals closer to the RSU generally gave very strong readings. But there is a sharp drop due to LoS issues, further away from the RSU.

Figure 8.4c RSU 3:- The readings showed that in front of the Hatchcroft Building, very good coverage was received up to 120m. However, after this the road starts to turn left towards the Church End Road resulting in less effective coverage, but as we go down the Greyhound Hill more readings are obtained.

Figure 8.4d RSU 4:- The RSU did not really cover the surrounding roads as its position relative to the extended roads was blocked by surrounding buildings. Hence, most of the readings for the extended coverage were lower than the FSPL calculations. However, the RSU allowed most of the readings to be covered in an east-to-west direction.

For Figure 8.5 the detailed analysis as follows:

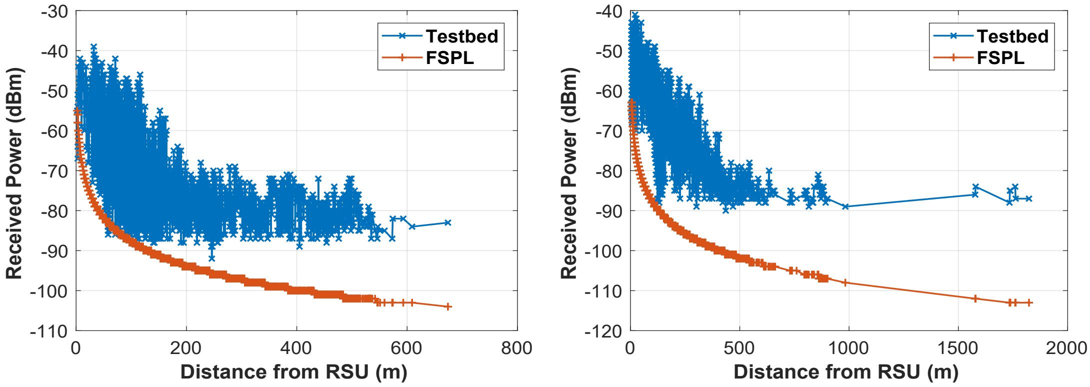
Figure 8.5a RSU 5:- Here, we can see that the readings are very consistent and there is a good coverage compared to the RSUs deployed at MDX campus buildings. However, when compared to the other RSUs deployed along the A41 the distance covered is considerably less and this is due to a bend and decreasing road elevation towards the other RSUs.

Figure 8.5b RSU 6:- The readings were consistent similar to other RSUs deployed along the A41 and the distance covered is considerably the highest among other RSUs due to LoS.

Figure 8.5c RSU 7:- Very consistent readings for a very long distance due to LoS. We can observe the received power for all the RSUs are in line with the FSPL for only a short distance. The different between the readings and FSPL is increasing considerably as the distances increase.

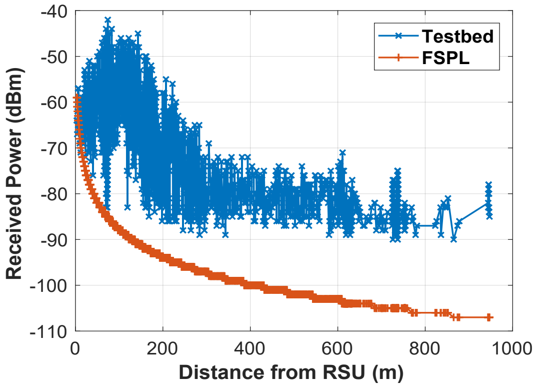
8.3 Application of Contention Probabilities

In this section, we are using three RSUs deployed on the A41 to reflect the scenario shown in Figure 5.5. The NDD of three RSUs deployed on A41 highway and also the overlapping coverage are shown in Figure 8.6. The distance of this coverage



(a) RSU 5

(b) RSU 6



(c) RSU 7

Figure 8.5: Power Received for Each Packet - A41 Testbed vs FSPL.

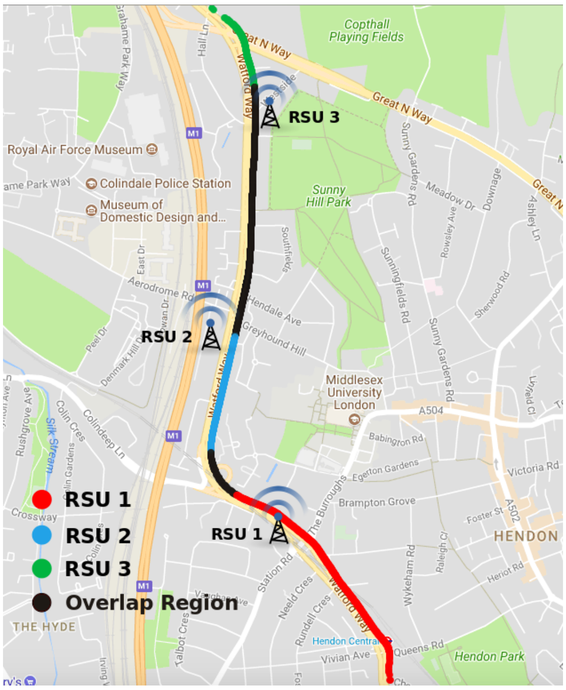


Figure 8.6: Full Coverage and Overlapping Map for A41, Watford Way, Hendon, London

is shown in Table 8.2. The speed limit on A41 Watford Way is 50 Mph (22.352 meters per second) and assuming that there is no speeding, dwell times can be calculated with the known dwell distance as shown in Table 8.2. The table also shows the overlapping distance between (RSU 1, RSU 2) and (RSU2, RSU3). This dwell time for the overlapping distance will be the Time to Handover as that will be the maximum time a MN will have with the given velocity to have a successful soft handover for seamless communication. An important observation with this coverage readings is that the NDD and the overlapping distance will not be same and it will change based on the deployment of the RSUs. The deployment in-turn will be dependent of the geographical features. Therefore, infrastructure has to be dynamic and intelligent enough to know its coverage area in any given scenario and use this information to achieve a proactive intelligent edge infrastructure to support proactive handover and resource allocation.

In order to demonstrate the calculation of contention probabilities for the NDDs of the RSUs, let us consider two vehicles MN_A and MN_B. Since the NDD

Table 8.2: Communication Coverage Segmentation Distance and Time.

RSU No.	**NDD**	$\aleph$	$\hbar$	$\mathbb{N}$
RSU 1	974 m	43.57 s	4 s	39.57 s
RSU 2	1390 m	62.19 s	4 s	58.19 s
RSU 3	1140 m	51.00 s	4 s	47.00 s
Overlapping Distance				
RSU 1 & RSU 2		173 m	7.74 s	
RSU 2 & RSU 3		828 m	37.04 s	

of the RSUs are known, the average $\aleph$ for both MNs for all three RSU can be calculated and in turn $\mathbb{N}$ can be calculated as shown in Table 8.2. We know from (Mapp et al., 2012) that the handover execution time is 4s and let us assume that the wait time of a request in the queue as 6s so that the request does not queue up long before it needs the resource from next network. Here, we assume that $\mathbb{T}$ is 20s (Figure 8.7) in order to make sure that if there is a Full Contention, the MN gets enough time to contest for resource in an alternative network. During contention, there are possibilities that $\mathbb{N}$ can be very small and therefore choose a suitable network to avoid frequent handover.

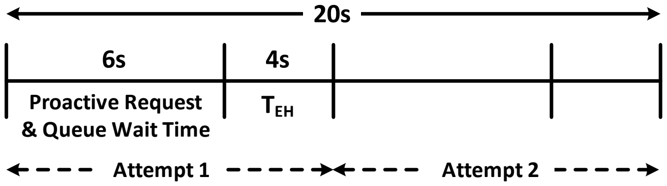


Figure 8.7: $\mathbb{T}$ Split up

8.3.1 Results and Discussion

All the results shown in this section is computed for RSU 1's NDD. Figure 8.8 shows the probability of having a Full Contention for MN_A i.e., α_A before entering the proactive resource allocation queue. The probability is calculated for two cases: keeping velocity of MN_A constant at 30 Mph and changing the average velocity of other incoming requests represented by MN_B from 10 to 70 Mph which is shown as the black line in Figure 8.8. In this case, α_A increases as the velocity of other incoming requests are increased because the higher velocity ensures that these requests will reach the queue more quickly and hence there is a higher possibility that MN_A will be rejected due to Full Contention with these faster requests. In the second scenario where the average velocity of the incoming requests i.e., MN_B is kept constant at 30 Mph while the velocity of MN_A is varied from 10 to 70 Mph which is shown as the red line in Figure 8.8. Here, α_A decreases due to the increase in velocity of MN_As relative to other incoming requests. This ensures that MN_As request will reach the queue faster. Hence, α_A decreases as the velocity of MN_A increases. So MN_A has a better chance of reaching the contention queue. In addition, α_A for MN_A and MN_B are equal at 30 Mph as the velocity of both the nodes become equal which shows that the probability of Full contention becomes equal for both requests.

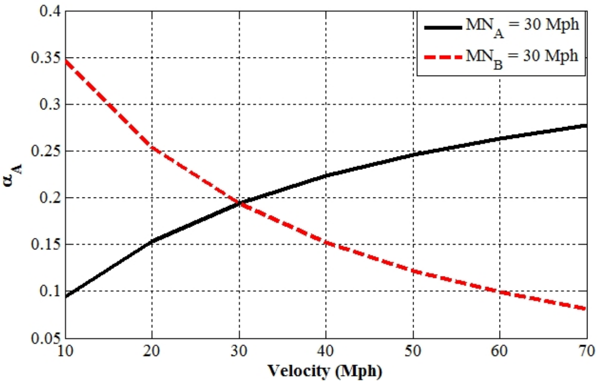


Figure 8.8: Probability of α

Figure 8.9 shows the probability of having a Full Contention in the proactive contention queue (β) for an incoming request, in this case let us consider MN_A. In the first case for calculating β, an average velocity (30 Mph) was considered for all the requests in proactive contention queue to avoid complexity and to demonstrate the importance of β. The probability is calculated for 10 to 70 Mph for MN_A which is shown as green line in Figure 8.9. We can observe that the β value is decreasing exponentially as the velocity increases which means that β decreases as $\mathbb{N}$ decreases. This is due to the fact that the time spent by MN_A in the contention queue will be smaller as the velocity increases relative to the average velocity of the requests in the contention queue. Therefore, the effect of

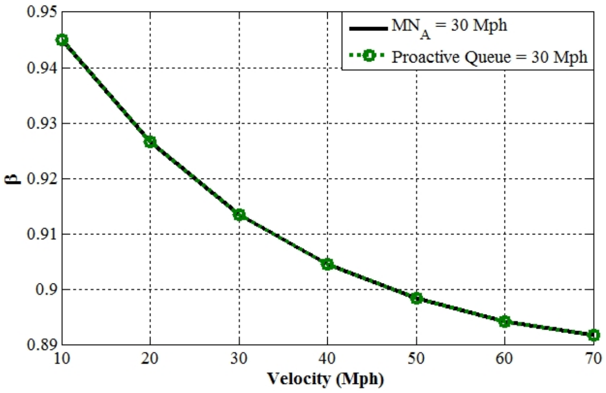


Figure 8.9: Probability of β

β of MN_A decreases as its velocity increases.

In the second case, where a velocity 30 Mph was considered for the incoming request MN_A and the probability is calculated for 10 to 70 Mph which is the average velocities for the requests in the contention queue, shown as black line in Figure 8.9. The effect of β is the same as previous case and this is because β is the summation of both incoming request and the requests in the queue i.e., $30 + 20 == 20 + 30$.

Calculating β accurately is very important so that the requests leaving contention queue due to Full Contention can be served by an alternative network. Since, β is dependent on the values of $\mathbb{N}$ and $\mathbb{T}$ of each request in the queue and new incoming request, therefore, a detailed analysis is required to accurately model β. By using this approach, it is possible to achieve seamless communication.

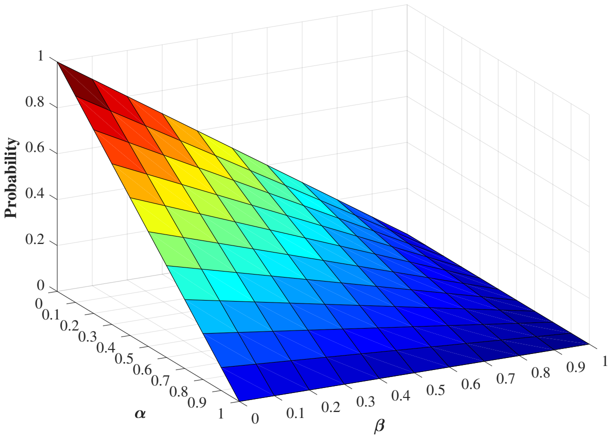


Figure 8.10: General Probability of a Request Reaching the Channel Queue

$$(1-\alpha)\times(1-\beta) \tag{8.1}$$

Equation (8.1) shows the general probability of a request reaching the channel queue and the resulting graph is shown in Figure 8.10 for all α and β value combinations from 0 to 1. The resultant probability is a 3D plane which shows effect of α and β in acquiring a resource. This can be further explored to find an optimal working space to build a proactive resource allocation algorithm.

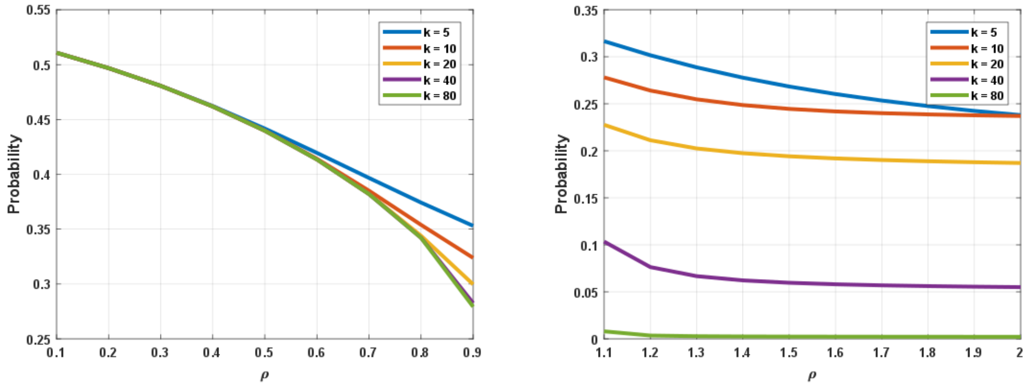


Figure 8.11: Probability of a Request Not Being Affected by β

Let us consider $\mathbb{T}_A$ and $\mathbb{T}_B$ are both 20. MN_A is moving at a velocity of 30 Mph and average velocity of the requests in the contention queue as 10Mph. Therefore, for RSU 1 the P^{A}_{Full} is 0.094055 and P^{B}_{Full} which is the probability of full contention for the requests with average velocity in the queue is 0.346165. Let us assume P_{Full}^{C} as 0.2. Therefore, we can find the probability of request A staying in the queue using Equation (7.9) and the resulting graph for different ρ values from 0.1 to 0.9 and ρ values from 1.1 to 2 is shown in Figure 8.11. From the result we can observe that as ρ increases the probability of a request being affected by β, resulting in not reaching the channel queue increases. In other words, probability of a request not being affected by β deceases as shown in Figure 8.11.

8.4 Summary

This Chapter has demonstrated the application of the proactive model in a VANET environment. The analysis was performed using static data collected from the testbed but it is possible to build an extensive real-time algorithm based on the proposed model to efficiently allocate the resources for users pro-actively.

Chapter 9

Conclusion and Future Work

In this chapter, a summary of the work done in this thesis is given where the major contributions are highlighted. This is followed by conclusions resulting from this work and a discussion on the directions for future research is presented.

9.1 Contribution of the Thesis

The contributions of the thesis can be summarised as follows:

Chapter 1 - In this thesis, we began by motivating the need for an efficient resource management techniques in a highly mobile environments such as vehicular networks. In order to do so it has was proposed that a proactive handover should be exploited to develop a proactive resource allocation based on user contention for resources.

Chapter 2 - A critical review of related literature on several areas like proactive handover, call admission control, proactive resource management, mobility prediction, etc., was presented in this chapter. In all these efforts reviewed, no new analytical model for user based resource contention was presented, which was the focus of this research.

Chapter 3 - This chapter described the significance of three methods i.e., analytical modelling, simulation and testbed experiments which has been used for conducting this research.

Chapter 4 - A detailed analysis of user contention for sending a message in traditional protocols like Pure ALOHA and Slotted ALOHA has been presented. It has been found that the traditional performance analysis of these networks focuses only on overall demand at infinite population levels i.e., only at N goes to ∞. Therefore, the Chapter highlighted the importance of looking into individual user performance in the network rather than just as a whole system. It was also shown that the proposed approach can also be applied to mobile systems. In addition, a comprehensive analysis of the traditional classical reactive handover approach was also presented and the analysis proved that it cannot cope in high mobility environment. This is due to the fact that there is a considerable amount of users moving out of the network without useful service due to their mobility. Therefore, has highlighted the need for a new proactive approach in order to address this issue.

Chapter 5 - Background of the Y-Comm reference framework and the key parameters such as time before handover and network dwell time proposed in the framework has been briefed. Based on a new coverage segmentation proposed in this chapter and the parameters of Y-Comm framework, a new condition for contention for resources has been described which can be used to develop proactive resource allocation.

Chapter 6 - A detailed mathematical description of the model to calculate the probability of contention for two and three MNs contending for the resource was presented. The analytical results were validated with simulation and found to be highly accurate. Also, it has been shown why we do not need a model to calculate the probability for large number of nodes contending for resources around the same time.

Chapter 7 - In this Chapter, using a thorough analysis, two different queueing models for a proactive system where the users contend for resources were presented. These models yielded two key parameters α i.e., full contention before entering the proactive queue and β i.e., full contention encountered in the proactive queue. It has been shown that the proposed model have outperformed the classical model. Based on the contention probability model it has been shown how these two key parameters can be calculated.

Chapter 8 - The usefulness of the proposed model has been demonstrated with the help of a VANET testbed. The analysis was performed using static data collected from the testbed and was shown that a heuristic algorithm can be designed to achieve a real time proactive resource allocation based on user mobility.

9.2 Conclusion

This thesis has explored a new proactive resource allocation approach by analysing the contention among various users trying to acquire a radio channel in a wireless network using two key parameters; Time to get Resource and Resource Hold Time. We introduced two proactive queuing models, the first calculates the probability that a Mobile Node will never acquire a channel amongst various simultaneous requests and so they can be instructed to do a handover to another network. A second case added a further refinement by introducing the concept of a proactive contention queue which was used to analyse users waiting to acquire the channel before they reach the coverage of the next network. The results show that the proactive approach leads to better management of network resources and significantly improves the overall system performance in terms of mean response time and throughput. Methods to calculate α and β have been explored and the application of this approach has been demonstrated in a Vehicular Ad-hoc Network testbed.

9.3 Future Work

9.3.1 Exploring an Operational Space

For future work, an operational space based on α, β and probability of a request reaching the channel queue can be explored to effectively use the proactive resource allocation system.

9.3.2 Developing Proactive Handover and Resource Allocation Algorithm for VANET Testbed

A proactive handover and resource allocation algorithm can be developed for the VANET testbed based on the above analysis and the initial thought about such an algorithm is shown in Figure 9.1. The RSUs will have the knowledge of the network coverage provided through the feedback from the MNs. This is performed by the Network Management Layer in Y-Comm framework. A database containing the coverage rage (GPS coordinate) and the RSSI value will be stored at the RSU and will be updated at a regular interval and sent to the MDX Cloud. The MDX Cloud System will share this information with other adjacent RSUs which, in turn will broadcast the information to the MN. This will help in predicting the time before handover and network dwell time for the next RSU range. Therefore, the application in the MN will calculate the $\mathbb{T}$ and $\mathbb{N}$ based on the received coverage information from the RSU with its current velocity.

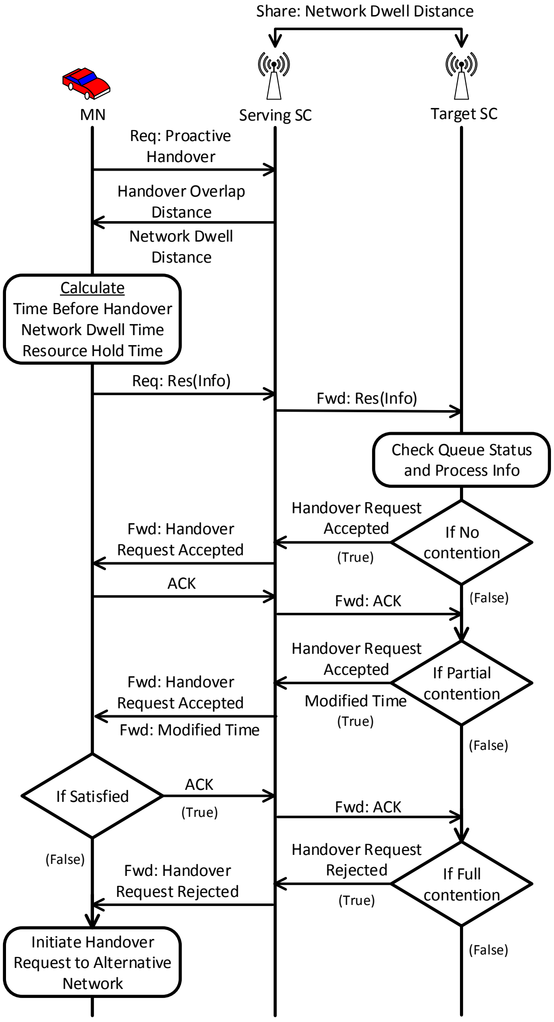


Figure 9.1: Proactive Handover and Resource Allocation in VANET Testbed

References

Arada systems, Feb 2016. URL http://www.aradasystems.com. 99

VS Abhayawardhana, IJ Wassell, D Crosby, MP Sellars, and MG Brown. Comparison of empirical propagation path loss models for fixed wireless access systems. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 1, pages 73–77. IEEE, 2005. 56

H. Abou-zeid, H.S. Hassanein, and S. Valentin. Optimal predictive resource allocation: Exploiting mobility patterns and radio maps. In *Global Communications Conference (GLOBECOM), 2013 IEEE*, pages 4877–4882, Dec 2013. doi: 10.1109/GLOCOMW.2013.6855723. 4, 14

Norman Abramson. The aloha system: another alternative for computer communications. In *Proceedings of the November 17-19, 1970, fall joint computer conference*, pages 281–285. ACM, 1970. 37, 40

Mahdi Aiash, Glenford Mapp, Aboubaker Lasebae, and Ameer Al-Nemrat. *Supporting LTE Networks in Heterogeneous Environment Using the Y-Comm Framework*, pages 125–136. Springer New York, New York, NY, 2013. ISBN 978-1-4614-6154-8. doi: 10.1007/978-1-4614-6154-8_13. URL https://doi.org/10.1007/978-1-4614-6154-8_13. 51, 52

W. Alasmary and Weihua Zhuang. The mobility impact in ieee 802.11p infrastructureless vehicular networks. In *Vehicular Technology Conference Fall (VTC 2010-Fall), 2010 IEEE 72nd*, pages 1–5, Sept 2010. doi: 10.1109/VETECF.2010.5594542. 14

S. Albasheir and M. Kadoch. Enhanced control for adaptive resource reservation of guaranteed services in lte networks. *IEEE Internet of Things Journal*, 3(2): 179–189, April 2016. ISSN 2327-4662. doi: 10.1109/JIOT.2015.2471105. 20

M. Almulla, Yikun Wang, A. Boukerche, and Zhenxia Zhang. Design of a fast location-based handoff scheme for ieee 802.11 vehicular networks. *Vehicular Technology, IEEE Transactions on*, 63(8):3853–3866, Oct 2014. ISSN 0018-9545. doi: 10.1109/TVT.2014.2309677. 9, 12, 23

K. Balakrishnan. *Exponential Distribution: Theory, Methods and Applications*. Taylor & Francis, 1996. ISBN 9782884491921. URL `https://books.google.co.uk/books?id=sQfpNOKzDjcC`. 25, 26

J. Banks and J.S. Carson. *Discrete-event system simulation*. Prentice-Hall international series in industrial and systems engineering. Prentice-Hall, 1984. ISBN 9780132155823. URL `https://books.google.co.uk/books?id=_mJRAAAAMAAJ`. 29, 30

Abul Bashar, Gerard Parr, Sally McClean, Bryan Scotney, and Detlef Nauck. *Learning-Based Call Admission Control Framework for QoS Management in Heterogeneous Networks*, pages 99–111. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-14306-9. doi: 10.1007/978-3-642-14306-9_11. 17

E. Bastug, M. Bennis, and M. Debbah. Living on the edge: The role of proactive caching in 5g wireless networks. *IEEE Communications Magazine*, 52(8):82–89, Aug 2014. ISSN 0163-6804. doi: 10.1109/MCOM.2014.6871674. 11, 17, 23

Z. Becvar, P. Mach, and E. C. Strinati. Prediction of channel quality after handover for mobility management in 5g. In *5G for Ubiquitous Connectivity (5GU), 2014 1st International Conference on*, pages 35–39, Nov 2014. doi: 10.4108/icst.5gu.2014.258140. 14

Michael Behrisch, Laura Bieker, Jakob Erdmann, and Daniel Krajzewicz. Sumo–simulation of urban mobility: an overview. In *Proceedings of SIMUL 2011,* *The Third International Conference on Advances in System Simulation*. ThinkMind, 2011. 30

T. Bejaoui. Vehicle density-based resource allocation policy for qos provisioning in vehicular communication networks. In *Innovations in Information Technology (IIT), 2012 International Conference on*, pages 310–315, March 2012. doi: 10.1109/INNOVATIONS.2012.6207756. 16

Tarek Bejaoui. Qos-oriented high dynamic resource allocation in vehicular communication networks. *The Scientific World Journal*, vol. 2014(Article ID 718698):9 pages, 2014. doi: 10.1155/2014/718698. 16

Hadda Ben Elhadj, Jocelyne Elias, Lamia Chaari, and Lotfi Kamoun. Multi-attribute decision making handover algorithm for wireless body area networks. *Comput. Commun.*, 81(C):97–108, May 2016. ISSN 0140-3664. doi: 10.1016/j.comcom.2016.01.007. URL `http://dx.doi.org/10.1016/j.comcom.2016.01.007`. 11

Shumin Bi, Cailian Chen, Rong Du, and Xinping Guan. Proper handover between vanet and cellular network improves internet access. In *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*, pages 1–5, Sept 2014. doi: 10.1109/VTCFall.2014.6966137. 11, 16, 23

M. Boujelben, S. Ben Rejeb, and S. Tabbane. A novel self-organizing scheme for 4g advanced networks and beyond. In *Networks, Computers and Communications, The 2014 International Symposium on*, pages 1–5, June 2014. doi: 10.1109/SNCC.2014.6866534. 4, 16

Glaucio H.S. Carvalho, Isaac Woungang, Alagan Anpalagan, Rodolfo W.L. Coutinho, and Joo C.W.A. Costa. A semi-markov decision process-based joint call admission control for inter-rat cell re-selection in next generation wireless networks. *Computer Networks*, 57(17):3545 – 3562, 2013. ISSN 1389-1286. doi: http://dx.doi.org/10.1016/j.comnet.2013.07.029. URL `http://www.sciencedirect.com/science/article/pii/S1389128613002600`. 16

U. Challita, L. Dong, and W. Saad. Deep learning for proactive resource allocation in lte-u networks. In *European Wireless 2017; 23th European Wireless Conference*, pages 1–6, May 2017. 15

Devendra K. Chaturvedi. *Modeling and Simulation of Systems Using MATLAB and Simulink*. CRC Press, Inc., Boca Raton, FL, USA, 1st edition, 2009. ISBN 1439806721, 9781439806722. 30, 31

A. Ben Cheikh, M. Ayari, R. Langar, and L. A. Saidane. Ohmp-cac: Optimized handoff scheme based on mobility prediction and qos constraints for femtocell networks. In *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 936–941, Sept 2016. doi: 10.1109/IWCMC.2016.7577184. 17

M Chiang. Balancing transport and physical layers in wireless multihop networks: jointly optimal congestion control and power control. *IEEE Journal of Selected Areas in Communications*, 23(1):104–116, 2005. 18, 23

Y. L. Chung, Z. Tsai, and C. H. Yang. A study of quota-based dynamic network selection for multimode terminal users. *IEEE Systems Journal*, 8(3):759–768, Sept 2014. ISSN 1932-8184. doi: 10.1109/JSYST.2012.2228121. 21

David N. Cottingham. Vehicular wireless communication. Technical Report UCAM-CL-TR-741, University of Cambridge, Computer Laboratory, January 2009. URL `http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-741.pdf`. xi, 10, 36

Hesham El Gamal, Atilla Eryilmaz, Giuseppe Caire, SHA Fei, and Margaret McLaughlin. Predictive network system and method, June 13 2017. US Patent 9,680,766. 19, 23

P. Fazio, M. Tropea, C. Sottile, S. Marano, M. Voznak, and F. Strangis. Mobility prediction in wireless cellular networks for the optimization of call admission control schemes. In *2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–5, May 2014. doi: 10.1109/CCECE.2014.6900981. 15

S. Feirer and T. Sauter. Seamless handover in industrial wlan using ieee 802.11k. In *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*, pages 1234–1239, June 2017. doi: 10.1109/ISIE.2017.8001421. 12

Jie Gao. Analysis of aloha and slotted aloha, 2009. URL `http://www3.cs.stonybrook.edu/~jgao/CSE590-fall09/aloha-analysis.pdf`. 40, 42

E. Gemikonakli, G. Mapp, O. Gemikonakli, and E. Ever. Exploring service and buffer management issues to provide integrated voice and data services in single and multi-channel wireless networks. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, pages 1056–1063, March 2013. doi: 10.1109/AINA.2013.57. 11, 48, 91

A. Ghosh, V. Vardhan, G. Mapp, O. Gemikonakli, and J. Loo. Providing ubiquitous communication using road-side units in vanet systems: Unveiling the challenges. In *ITS Telecommunications (ITST), 2013 13th International Conference on*, pages 74–79, Nov 2013. doi: 10.1109/ITST.2013.6685524. 9, 23

A. Ghosh, V.V. Paranthaman, G. Mapp, and O. Gemikonakli. Providing ubiquitous communication using handover techniques in vanet systems. In *Ad Hoc Networking Workshop (MED-HOC-NET), 2014 13th Annual Mediterranean*, pages 195–202, June 2014a. doi: 10.1109/MedHocNet.2014.6849124. 9, 23

Arindam Ghosh. *Investigating seamless handover in VANET systems*. PhD thesis, Middlesex University, 2016. viii, xi, 3, 56, 57

Arindam Ghosh, Vishnu Paranthaman, Glenford Mapp, and Orhan Gemikonakli. Exploring efficient seamless handover in vanet systems using network dwell time. *EURASIP Journal on Wireless Communications and Networking*, 2014 (1):227, 2014b. ISSN 1687-1499. doi: 10.1186/1687-1499-2014-227. 48, 91

Andre S. Gomes, Bruno Sousa, David Palma, Vitor Fonseca, Zhongliang Zhao, Edmundo Monteiro, Torsten Braun, Paulo Simoes, and Luis Cordeiro. Edge caching with mobility prediction in virtualized {LTE} mobile networks. *Future Generation Computer Systems*, pages –, 2016. ISSN 0167-739X. doi: http://dx.doi.org/10.1016/j.future.2016.06.022. URL `http://www.sciencedirect.com/science/article/pii/S0167739X16302072`. 14

J. Gozalvez, M. Sepulcre, and R. Bauza. Ieee 802.11p vehicle to infrastructure communications in urban environments. *IEEE Communications Magazine*, 50 (5):176–183, May 2012. ISSN 0163-6804. doi: 10.1109/MCOM.2012.6194400. 103

D. Grewe, M. Wagner, and H. Frey. Perceive: Proactive caching in icn-based vanets. In *2016 IEEE Vehicular Networking Conference (VNC)*, pages 1–8, Dec 2016. doi: 10.1109/VNC.2016.7835962. 20

J. Guo, C. Yao, and C. Yang. Proactive resource allocation planning with three-levels of context information. In *2016 IEEE/CIC International Conference on Communications in China (ICCC)*, pages 1–6, July 2016. doi: 10.1109/ICCChina.2016.7636825. 20

H. Hartenstein and K. Laberteaux. *VANET: Vehicular Applications and Inter-Networking Technologies.* Intelligent Transport Systems. Wiley, 2009. ISBN 9780470740620. URL `https://books.google.co.uk/books?id=VNbkpbIg1EoC`. xi, 32, 34

M. R. Hossain, A. Z. M. A. H. Real, and A. Rahman. Handover management in heterogeneous network. In *2015 18th International Conference on Computer and Information Technology (ICCIT)*, pages 87–92, Dec 2015. doi: 10.1109/ICCITechn.2015.7488048. 12

Qian Huang, Yue-Cai Huang, King-Tim Ko, and V.B. Iversen. Loss performance modeling for hierarchical heterogeneous wireless networks with speed-sensitive call admission control. *Vehicular Technology, IEEE Transactions on*, 60(5): 2209–2223, Jun 2011. ISSN 0018-9545. doi: 10.1109/TVT.2011.2142203. 1

M. Ismail and Weihua Zhuang. A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium. *Selected Areas in Communications, IEEE Journal*, 30(2):425–432, February 2012. ISSN 0733-8716. doi: 10.1109/JSAC.2012.120222. 16

R. Jain. *Art of Computer Systems Performance Analysis: Techniques For Experimental Design Measurements Simulation and Modeling.* Wiley,

2015. ISBN 9781118858424. URL `https://books.google.co.uk/books?id=sqqCoAEACAAJ`. 27, 29, 46

David E. Joyce. e as the limit, 2012. URL `https://mathcs.clarku.edu/~djoyce/ma122/elimit.pdf`. 42

S. Kabiri, H. Kalbkhani, T. Lotfollahzadeh, M. G. Shayesteh, and V. Solouk. Technique for order of preference by similarity to ideal solution based predictive handoff for heterogeneous networks. *IET Communications*, 10(13):1682–1690, 2016. ISSN 1751-8628. doi: 10.1049/iet-com.2015.0598. 17

S. Kaur and D. Selvamuthu. Adaptive joint call admission control scheme in lte-umts networks. In *2014 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pages 63–70, Nov 2014. doi: 10.1109/COMNETSAT.2014.7050527. 17

F P Kelly, A Maulloo, and D Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, pages 247–252, 1998. URL `http://www.statslab.cam.ac.uk/~frank/rate.pdf`. 17, 23

S. Khan, A. Ahmed, I. Ullah, and S. M. Zubair. Call admission control based femtocell handover in lte networks. In *2016 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube)*, pages 196–201, April 2016. doi: 10.1109/ICECUBE.2016.7495223. 17

Younghyun Kim, Haneul Ko, Sangheon Pack, Wonjun Lee, and Xuemin Shen. Mobility-aware call admission control algorithm with handoff queue in mobile hotspots. *Vehicular Technology, IEEE Transactions on*, 62(8):3903–3912, Oct 2013. ISSN 0018-9545. doi: 10.1109/TVT.2013.2245928. 16

Y. Kirsal, E. Gemikonakli, E. Ever, G. Mapp, and O. Gemikonakli. An analytical approach for performance analysis of handoffs in the next generation integrated cellular networks and wlans. In *Computer Communications and Networks (ICCCN), 2010 Proceedings of 19th International Conference on*, pages 1–6, Aug 2010. doi: 10.1109/ICCCN.2010.5560093. 45, 48, 91

Y. Kirsal, E. Ever, G. Mapp, and O. Gemikonakli. Enhancing the modelling of vertical handover in integrated cellular/wlan environments. In *Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference on*, pages 924–930, March 2013. doi: 10.1109/AINA.2013.37. 45, 48, 91

Yonal Kirsal. *Modelling and performance evaluation of wireless and mobile communication systems in heterogeneous environments.* PhD thesis, Middlesex University, 2013. viii, 30, 46

Yonal Kirsal. Analytical modelling of a new handover algorithm to improve allocation of resources in highly mobile environments. *International Journal of Computers Communications & Control*, 11(6):789–803, 2016. 11, 23

Juntao Li, Tao Luo, and Lei Ding. A location aware based handoff algorithm in v2i system of railway environment. In *Global Information Infrastructure and Networking Symposium (GIIS), 2014*, pages 1–6, Sept 2014. doi: 10.1109/GIIS.2014.6934273. 9

Hao Liang and Weihua Zhuang. Cooperative data dissemination via roadside wlans. *Communications Magazine, IEEE*, 50(4):68–74, April 2012. ISSN 0163-6804. doi: 10.1109/MCOM.2012.6178836. 15

C Liu, L Shi, and B Liu. Utility-based bandwidth allocation for triple-play services. In *Fourth European Conference on Universal Multiservice Networks, ECUMN '07*, pages 327–336, 2007. 18, 23

Q. Liu, H. Yu, and C. W. Chen. Proactive interference avoidance for mobile-to-mobile communication in lte networks. *IEEE Transactions on Vehicular Technology*, 65(9):7064–7077, Sept 2016. ISSN 0018-9545. doi: 10.1109/TVT.2015.2492240. 19

G. Mapp, F. Shaikh, M. Aiash, R. P. Vanni, M. Augusto, and E. Moreira. Exploring efficient imperative handover mechanisms for heterogeneous wireless networks. In *2009 International Conference on Network-Based Information Systems*, pages 286–291, Aug 2009. doi: 10.1109/NBiS.2009.95. 18, 23

G Mapp, F Katsriku, M Aiash N., Chinnam, R. Lopes, E. Moreira, R. M. Porto Vanni, and M. Augusto. Exploiting location and contextual information to develop a comprehensive framework for proactive handover in heterogeneous environments. *Journal of Computer Networks and Communications*, 2012(Article ID 748163):1–17, 2012. doi: 10.1155/2012/748163. xi, 3, 4, 11, 19, 23, 54, 55, 56, 109

G. Mapp, F. Sardis, and J. Crowcroft. Developing an implementation framework for the future internet using the y-comm architecture, sdn and nfv. In *2016 IEEE NetSoft Conference and Workshops (NetSoft)*, pages 43–47, June 2016. doi: 10.1109/NETSOFT.2016.7502440. xi, 2, 3, 4, 51

M. De Mari, E. Calvanese Strinati, and M. Debbah. Energy-efficiency and future knowledge tradeoff in small cells prediction-based strategies. In *2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 43–49, May 2014. doi: 10.1109/WIOPT.2014.6850277. 20

S. Mohanty and I.F. Akyildiz. A cross-layer (layer 2 + 3) handoff management protocol for next-generation wireless systems. *Mobile Computing, IEEE Transactions on*, 5(10):1347–1360, Oct 2006. ISSN 1536-1233. doi: 10.1109/TMC.2006.142. 13

A. Nadembega, A. Hafid, and T. Taleb. Mobility-prediction-aware bandwidth reservation scheme for mobile networks. *IEEE Transactions on Vehicular Technology*, 64(6):2561–2576, June 2015a. ISSN 0018-9545. doi: 10.1109/TVT.2014.2345255. 13

A. Nadembega, A. Hafid, and T. Taleb. A destination and mobility path prediction scheme for mobile networks. *IEEE Transactions on Vehicular Technology*, 64(6):2577–2590, June 2015b. ISSN 0018-9545. doi: 10.1109/TVT.2014.2345263. 15

T. O. Olwal, K. Djouani, and A. M. Kurien. A survey of resource management toward 5g radio access networks. *IEEE Communications Surveys Tutorials*, 18(3):1656–1686, thirdquarter 2016. ISSN 1553-877X. doi: 10.1109/COMST.2016.2550765. 2

Leo Patanapongpibul, Glenford Mapp, and Andy Hopper. An end-system approach to mobility management for 4g networks and its application to thin-client computing. *SIGMOBILE Mob. Comput. Commun. Rev.*, 10(3):13–33, July 2006. ISSN 1559-1662. doi: 10.1145/1148094.1148097. URL `http://doi.acm.org/10.1145/1148094.1148097`. 12

Haris Pervaiz and John Bigham. Game theoretical formulation of network selection in competing wirless networks: An analytic hierarchy process model. In *International Conference on Next Generation Mobile Applications, Services and Technologies*, 2009. doi: 10.1109/NGMAST.2009.28. 18, 23

SAEInternational. Dsrc implementation guide - a guide to users of sae j2735 message sets over dsrc, Feb 2010. URL `http://www.sae.org/standardsdev/dsrc/DSRCImplementationGuide.pdf`. 101

N. Saxena and A. Roy. Experimental framework of proactive handover with qos over wlans. *Electronics Letters*, 45(25):1313–1315, December 2009. ISSN 0013-5194. doi: 10.1049/el.2009.1246. 3

Thomas C. Schmidt, Matthias Waehlisch, Rajeev Koodli, Godred Fairhurst, and Dapeng Liu. Multicast Listener Extensions for Mobile IPv6 (MIPv6) and Proxy Mobile IPv6 (PMIPv6) Fast Handovers. RFC 7411, IEFT, November 2014. URL `https://tools.ietf.org/rfc/rfc7411.txt`. 12

Carolyn Seaman, Yuepu Guo, Clemente Izurieta, Yuanfang Cai, Nico Zazworka, Forrest Shull, and Antonio Vetrò. Using technical debt data in decision making: Potential decision approaches. In *Proceedings of the Third International Workshop on Managing Technical Debt*, MTD '12, pages 45–48, Piscataway, NJ, USA, 2012. IEEE Press. ISBN 978-1-4673-1749-8. URL `http://dl.acm.org/citation.cfm?id=2666036.2666044`. 11

F. Shaikh, G. Mapp, and A. Lasebae. Proactive policy management using tbvh mechanism in heterogeneous networks. In *Next Generation Mobile Applications, Services and Technologies, 2007. NGMAST '07. The 2007 International*

Conference on, pages 151–157, 2007a. doi: 10.1109/NGMAST.2007.4343414. 57

F Shaikh, G Mapp, and A Lasebae. Proactive policy management using tbvh mechanism in heterogeneous networks. In *International Conference on Next Generation Mobile Applications, Services and Technologies*, pages 151–157, 2007b. 18, 23

H. Sinky, B. Hamdaoui, and M. Guizani. Handoff-aware cross-layer assisted multi-path tcp for proactive congestion control in mobile heterogeneous wireless networks. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–7, Dec 2015. doi: 10.1109/GLOCOM.2015.7417490. 13

H. Sinky, B. Hamdaoui, and M. Guizani. Proactive multipath tcp for seamless handoff in heterogeneous wireless access networks. *IEEE Transactions on Wireless Communications*, 15(7):4754–4764, July 2016. ISSN 1536-1276. doi: 10.1109/TWC.2016.2545656. 13

Christoph Sommer, Reinhard German, and Falko Dressler. Bidirectionally coupled network and road traffic simulation for improved ivc analysis. *IEEE Transactions on Mobile Computing*, 10(1):3–15, 2011. 29

William Stallings and Moumita Mitra Manna. *Data and computer communications*, volume 6. Prentice hall Englewood Cliffs, NJ, 1997. 41

J. Tadrous and A. Eryilmaz. On optimal proactive caching for mobile networks with demand uncertainties. *IEEE/ACM Transactions on Networking*, 24(5): 2715–2727, October 2016. ISSN 1063-6692. doi: 10.1109/TNET.2015.2478476. 20

J. Tadrous, A. Eryilmaz, and H. El Gamal. Joint smart pricing and proactive content caching for mobile services. *IEEE/ACM Transactions on Networking*, 24 (4):2357–2371, Aug 2016. ISSN 1063-6692. doi: 10.1109/TNET.2015.2453793. 20

R Trestian, Q T Vien, P Shah, and G Mapp. Uefa-m: Utility-based energy efficient adaptive multimedia mechanism over lte hetnet small cells. In *In-*

ternational Symposium on Wireless Communication Systems (ISWCS 2017), pages 408–413, 2017. 18, 23

G Tychogiorgos and K K Leung. Optimization-based resource allocation in communication networks. *Computer Networks, Elsevier*, 66:32–45, 2014. 18, 23

S. Ucar, S. Coleri Ergen, and O. Ozkasap. Multi-hop cluster based ieee 802.11p and lte hybrid architecture for vanet safety message dissemination. *Vehicular Technology, IEEE Transactions on*, PP(99):1–1, 2015. ISSN 0018-9545. doi: 10.1109/TVT.2015.2421277. 4

A. Ukil and J. Sen. Proactive resource reservation in next-generation wireless networks. In *Communications (NCC), 2010 National Conference on*, pages 1–5, Jan 2010. doi: 10.1109/NCC.2010.5430157. 16

A. Ulvan, R. Bestak, and M. Ulvan. The study of handover procedure in lte-based femtocell network. In *Wireless and Mobile Networking Conference (WMNC), 2010 Third Joint IFIP*, pages 1–6, Oct 2010. doi: 10.1109/WMNC.2010.5678766. 3, 12

Andras Varga. Omnet++: an extensible, modular, component-based c++ network simulation, 2014. 29

Jordi Vilaplana, Francesc Solsona, Ivan Teixidó, Jordi Mateo, Francesc Abella, and Josep Rius. A queuing theory model for cloud computing. *J. Supercomput.*, 69(1):492–507, July 2014. ISSN 0920-8542. doi: 10.1007/s11227-014-1177-y. URL `http://dx.doi.org/10.1007/s11227-014-1177-y`. 11

W. Wang, H. Xu, and H. Zhou. Proactive unnecessary handover avoidance scheme in lte-a small cells. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2214–2218, April 2014. doi: 10.1109/WCNC.2014.6952673. 13

Sabine Wittevrongel and Tuan Phung-Duc, editors. *Analytical and Stochastic Modelling Techniques and Applications*. 9845. Springer International Publishing, 2016. doi: 10.1007/978-3-319-43904-4. URL `http://www.springer.com/gb/book/9783319439037`. 24

D. Xenakis, L. Merakos, N. Passas, and C. Verikoukis. Handover Decision for Small Cells: Algorithms, Lessons Learned and Simulation Study. *Communication Networks, Elsevier*, 100:64–74, May 2016. 11

Xiaohuan Yan, Nallasamy Mani, and Y.Ahmet Sekercioglu. A traveling distance prediction based method to minimize unnecessary handovers from cellular networks to wlans. *Communications Letters, IEEE*, 12(1):14–16, January 2008. ISSN 1089-7798. doi: 10.1109/LCOMM.2008.071430. 13

Xiaohuan Yan, Y. Ahmet ekerciolu, and Sathya Narayanan. A survey of vertical handover decision algorithms in fourth generation heterogeneous wireless networks. *Computer Networks*, 54(11):1848 – 1863, 2010. ISSN 1389-1286. doi: http://dx.doi.org/10.1016/j.comnet.2010.02.006. URL `http://www.sciencedirect.com/science/article/pii/S1389128610000502`. 13

V. Yazc, U. C. Kozat, and M. O. Sunay. A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management. *IEEE Communications Magazine*, 52(11):76–85, Nov 2014. ISSN 0163-6804. doi: 10.1109/MCOM.2014.6957146. 14

Xiaobo Yu, P. Navaratnam, and K. Moessner. Resource reservation schemes for ieee 802.11-based wireless networks: A survey. *Communications Surveys Tutorials, IEEE*, 15(3):1042–1061, Third 2013. ISSN 1553-877X. doi: 10.1109/SURV.2012.111412.00029. 14

Ahmed H. Zahran, Ben Liang, and Aladdin Saleh. Signal threshold adaptation for vertical handoff in heterogeneous wireless networks. *Mob. Netw. Appl.*, 11(4):625–640, August 2006. ISSN 1383-469X. doi: 10.1007/s11036-006-7326-7. URL `http://dx.doi.org/10.1007/s11036-006-7326-7`. 13

E. Zeljkovic, R. Riggio, and S. Latre. Exploiting distance information for transparent access point driven wi-fi handovers. In *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pages 1–6, June 2017. doi: 10.1109/WoWMoM.2017.7974344. 12

Haijun Zhang, Wenmin Ma, Wei Li, Wei Zheng, Xiangming Wen, and Chunxiao Jiang. Signalling cost evaluation of handover management schemes in lte-advanced femtocell. In *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pages 1–5, May 2011. doi: 10.1109/VETECS.2011.5956481. 12

Z Zhao, M Karimzadeh L Guardalben, J Silva, and T Braun. Mobility prediction-assisted over-the-top edge prefetching for hierarchical vanets. *IEEE Journal of Selected Areas of Communications*, 36(8):1786–1801, 2018. doi: 10.1109/JSAC.2018.2844681. 19, 23

Weihua Zhuang and M. Ismail. Cooperation in wireless communication networks. *Wireless Communications, IEEE*, 19(2):10–20, April 2012. ISSN 1536-1284. doi: 10.1109/MWC.2012.6189408. 15

Appendix A - List of Publications

1. V. V. Paranthaman, Y. Kirsal, G. Mapp, P. Shah, and H. X. Nguyen,“Exploiting Resource Contention in Highly Mobile Environments and its Application to Vehicular Ad-Hoc Networks”, *IEEE Transactions on Vehicular Technology (TVT), February 2019*

2. V. V. Paranthaman, Y. Kirsal, G. Mapp, P. Shah, and H. X. Nguyen,“Exploring a New Proactive Algorithm for Resource Management and Its Application to Wireless Mobile Environments”, *2017 The 42nd IEEE Int. Conf. on Local Computer Networks (LCN)*

3. G. Hanson, S. Phull, G. Mapp, A. Ghosh and V. V. Paranthaman,“Building a Connected Vehicle Testbed to study the development and deployment of C-ITS in the UK”, *Intelligent Transport Systems (ITS) UK Review Article, July 2017*

4. V. V. Paranthaman, A. Ghosh, G. Mapp, V. Iniovosa, P. Shah, H. X.Nguyen, O. Gemikonakli, and S. Rahman,“Building a Prototype VANET Testbed to Explore Communication Dynamics in Highly Mobile Environments”, *2016 Springer Int. Conf. on Testbeds and Research Infrastructures for the Development of Networks & Communities (TRIDENTCOM)*

5. V. V. Paranthaman, G. Mapp, P. Shah, H. X. Nguyen, and A. Ghosh,“Exploring Markov Models for the Allocation of Resources for Proactive Handover in a Mobile Environments”, *2015 IEEE Int. Work. in Conj. with Conf. on Local Computer Networks (LCN)*

6. A. Ghosh, V. V. Paranthaman, G. Mapp, O. Gemikonakli, and J. Loo,

---

“Enabling Seamless V2I communications Towards Developing Cooperative Automotive Applications in VANET Systems”, *IEEE Communication Magazine on Advances in V2X Connectivity*